Alex Faron

EEN 540
Digital Speech & Audio Processing

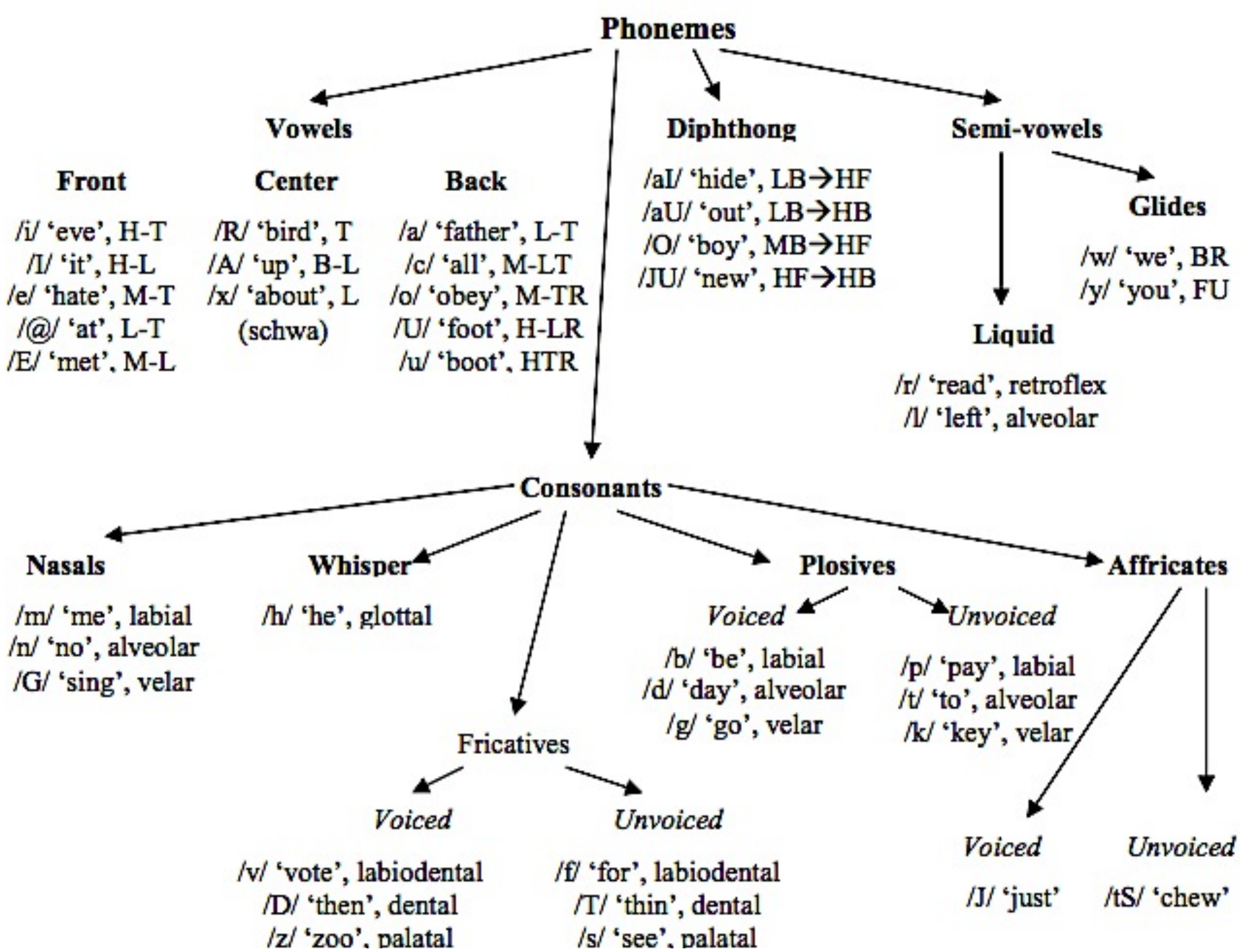
Dr. Scordilis
University of Miami

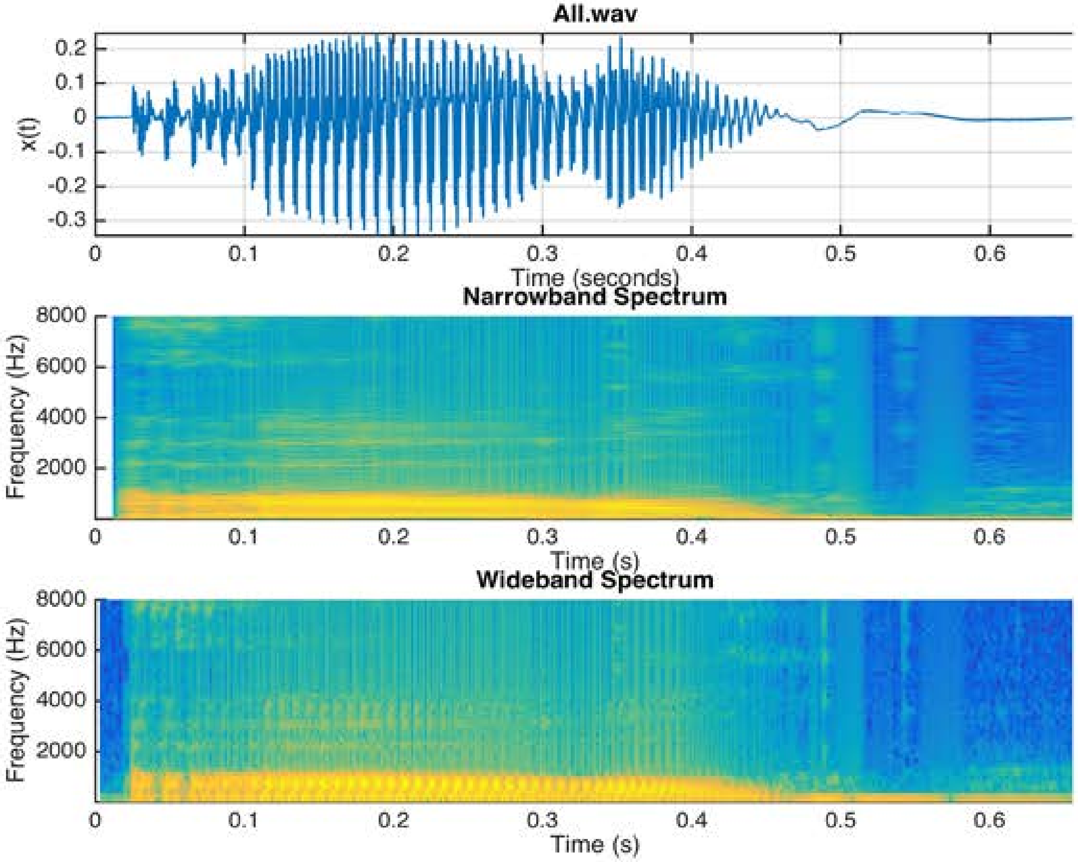
Project 1
Spectrographic Analysis of English Phonemes

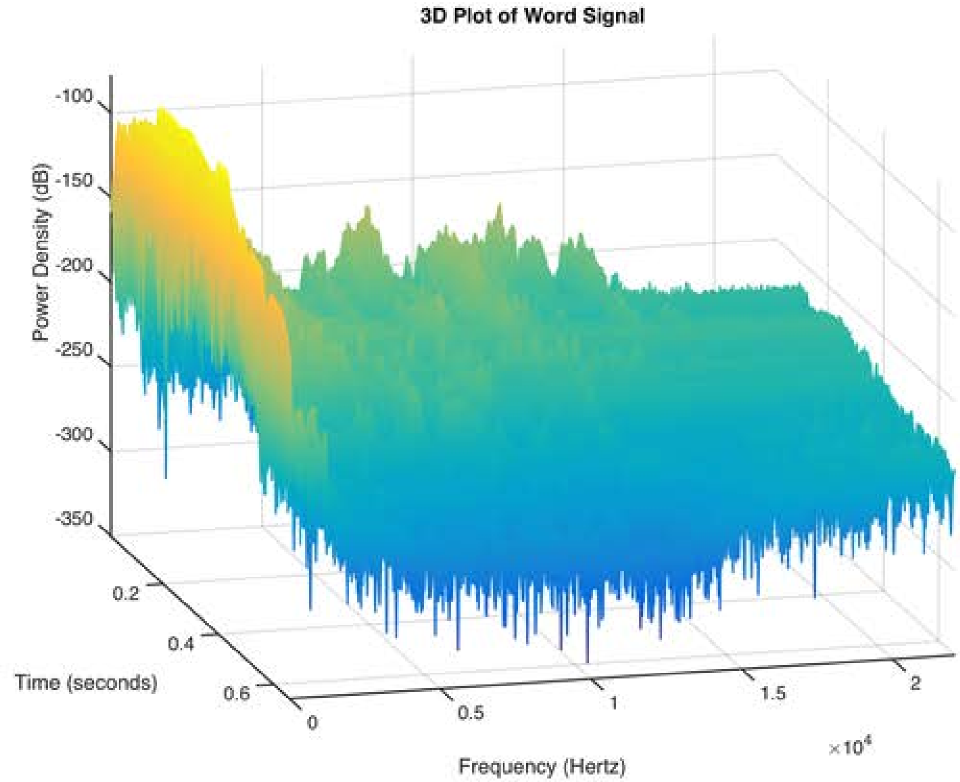
Introduction:

Phonemes are the building blocks of words. They are the smallest units that can change the meaning of a word. Kiss and kill differ by only one phoneme but their meanings vary greatly. As the chart below shows, phonemes are classified in a multitude of ways. Some phonemes are close to periodic in nature while others are noisy. In addition, the different filters of the vocal tract and even the transition between phonemes can change the meaning of a word. This project's goal was to realize a library of the forty English phonemes and analyze the waveforms of each. The results are organized online and in this PDF. This PDF contains the MATLAB plots and code while the webpage hosts a diagram of the phoneme tree and hyperlinks to all forty specified phonemes and words. The plots in the top left corner show the waveform and spectrographs (narrowband and wideband) of the word containing the phoneme of interest. The three-dimensional plots in the top right corners show the relationships between frequency, time, and power density of the word. The plots in the lower region show the waveform and linear prediction spectral envelope of the phoneme of interest.

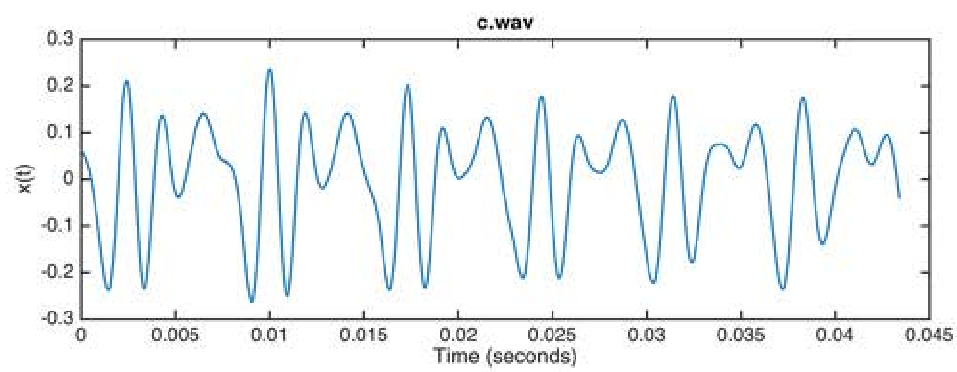
English phonemes and their features

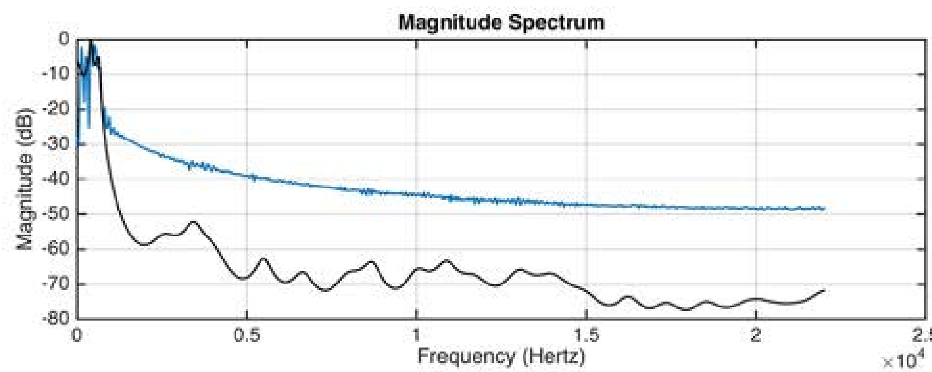


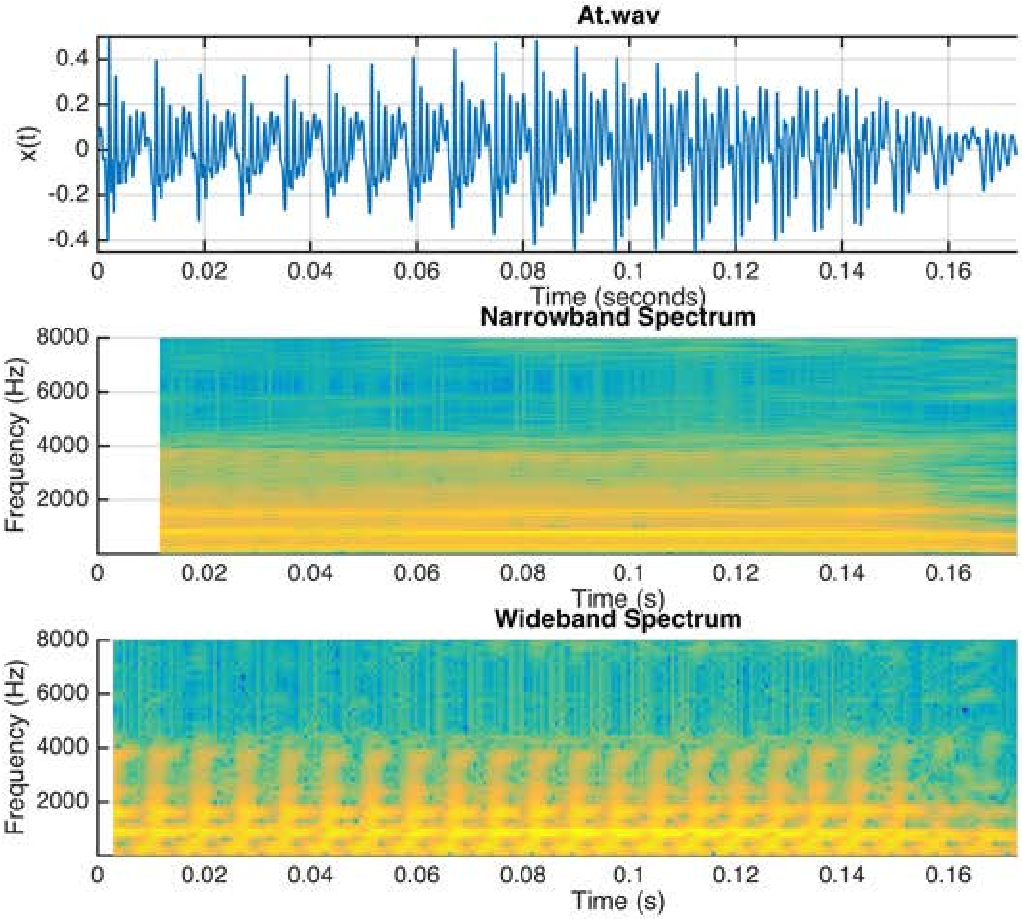

All.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

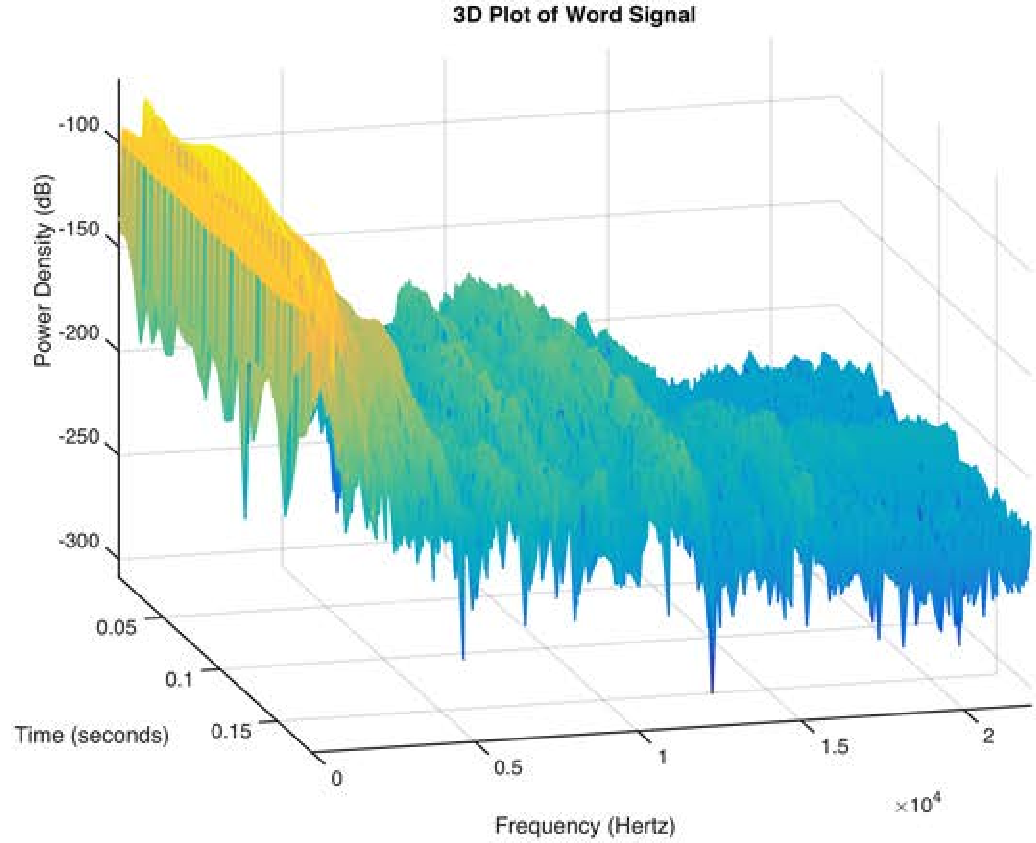

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

/c/ ‘All’

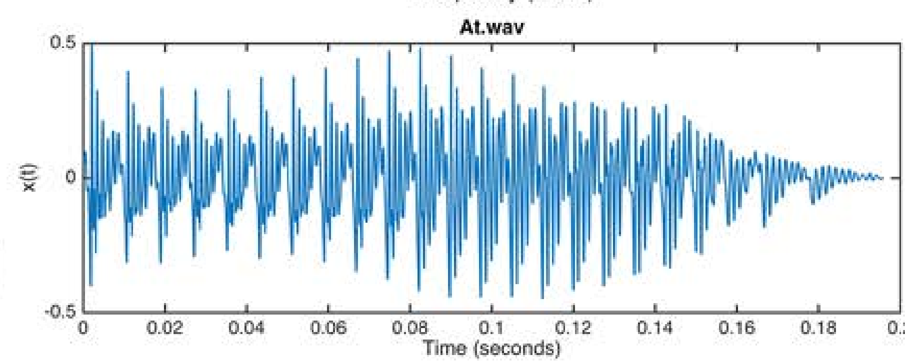

c.wav
x(t)
Time (seconds)

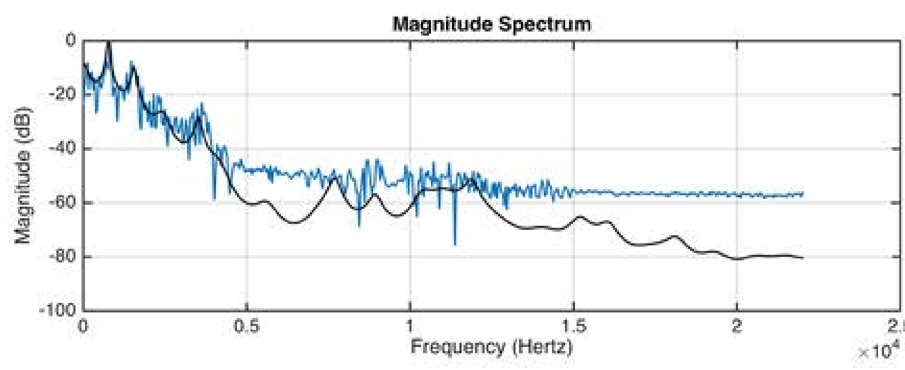

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

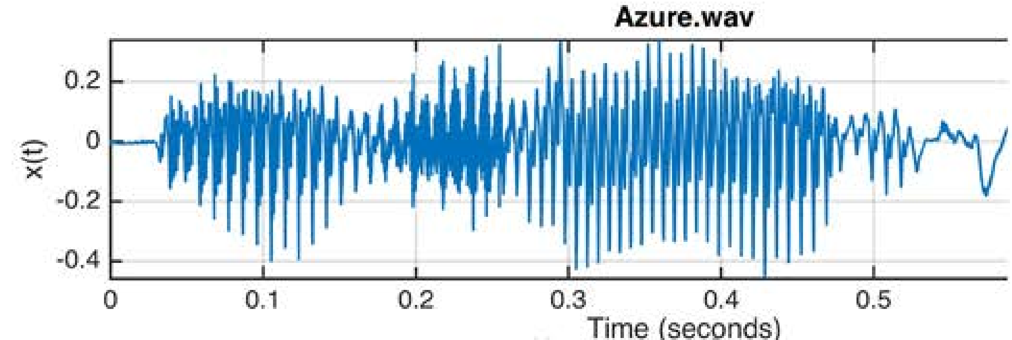

At.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

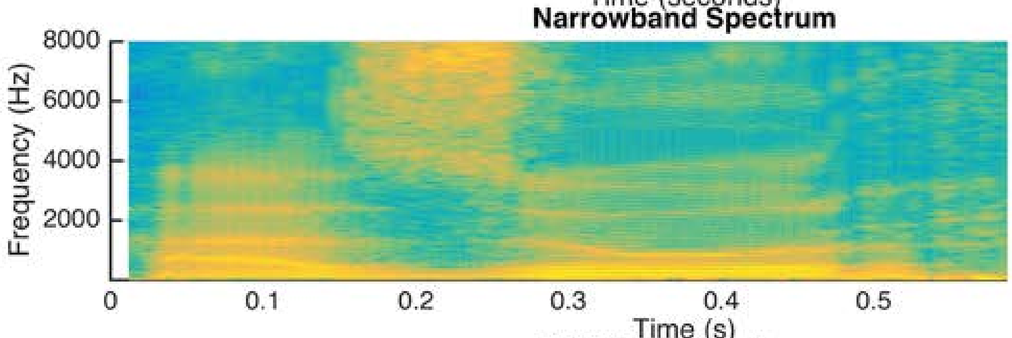

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

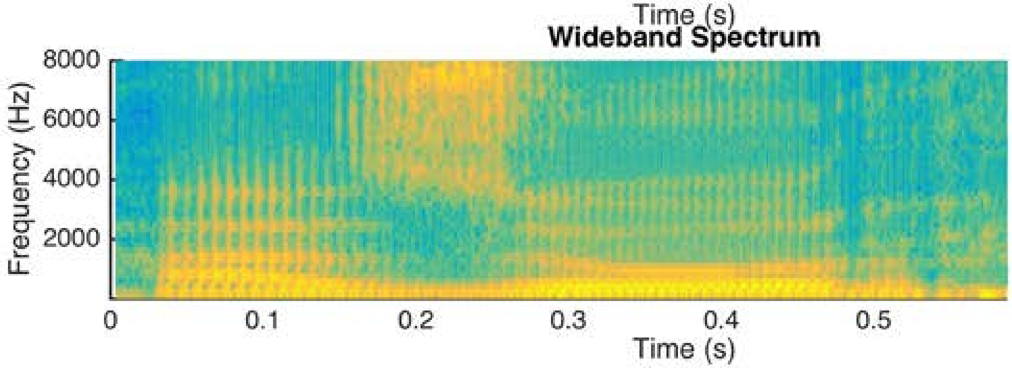
/@/ ‘At’

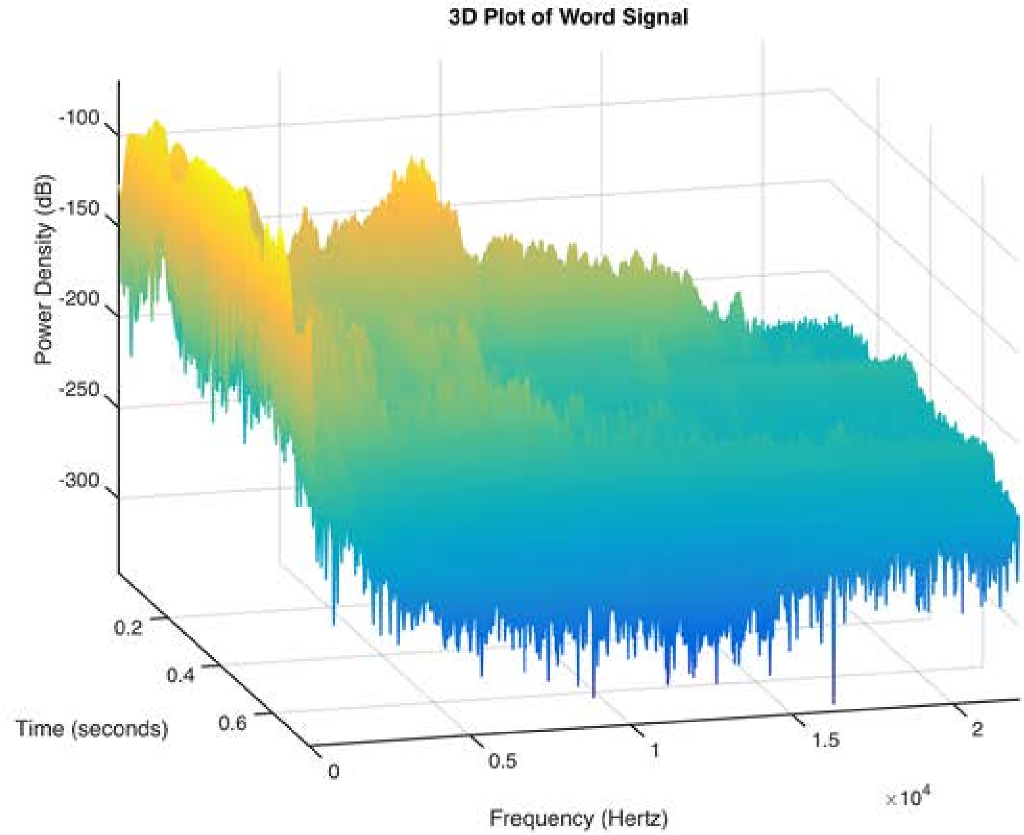

At.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


Azure.wav
x(t)
Time (seconds)

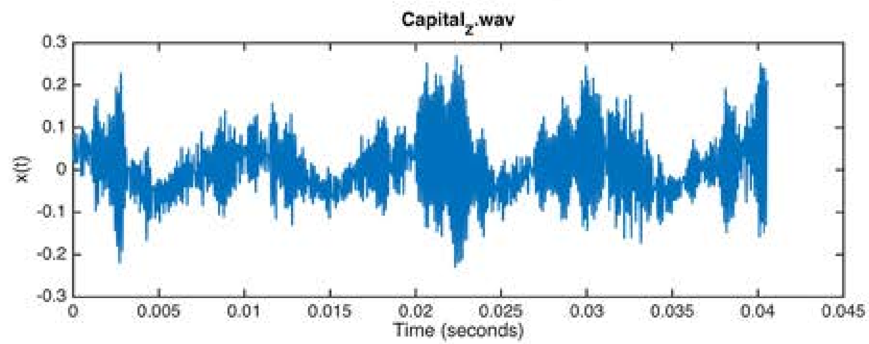

Narrowband Spectrum
Frequency (Hz)
Time (s)

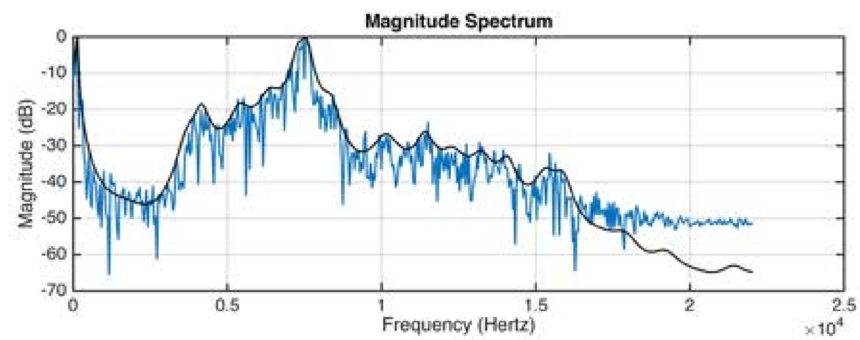

Wideband Spectrum
Frequency (Hz)
Time (s)

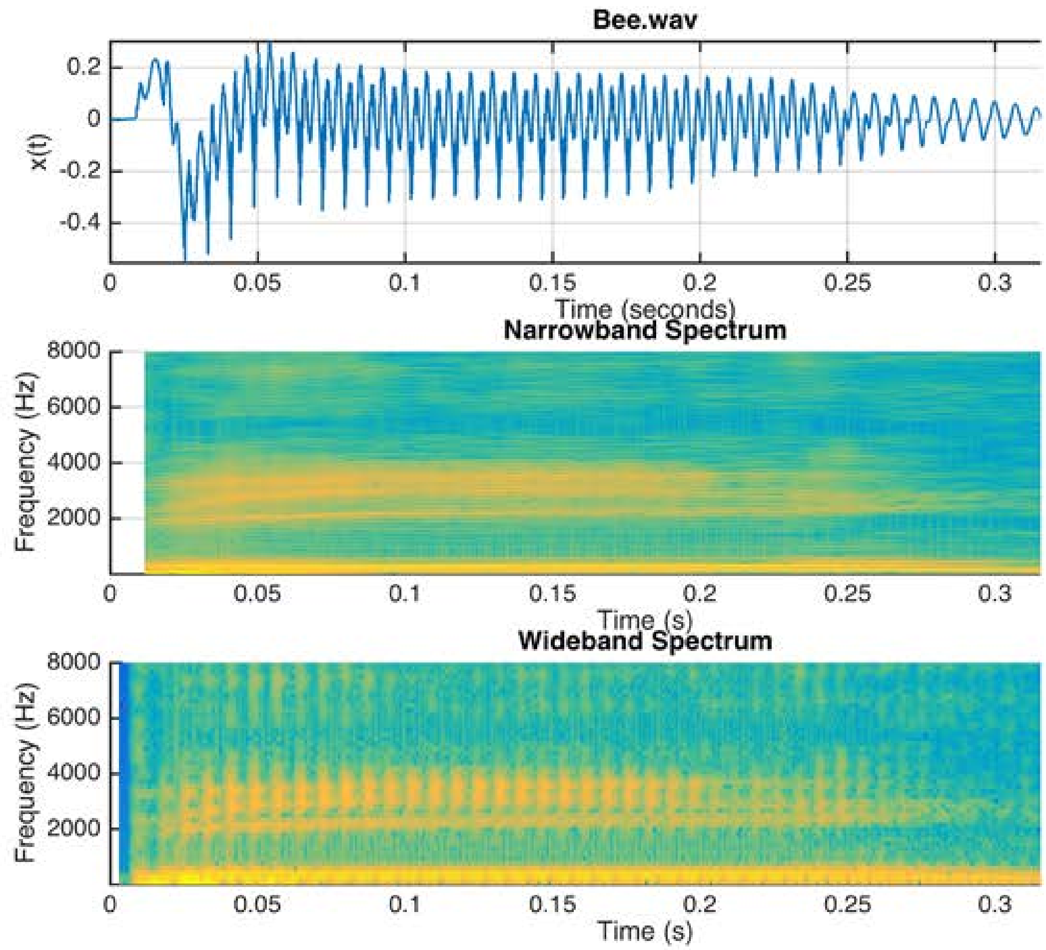

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

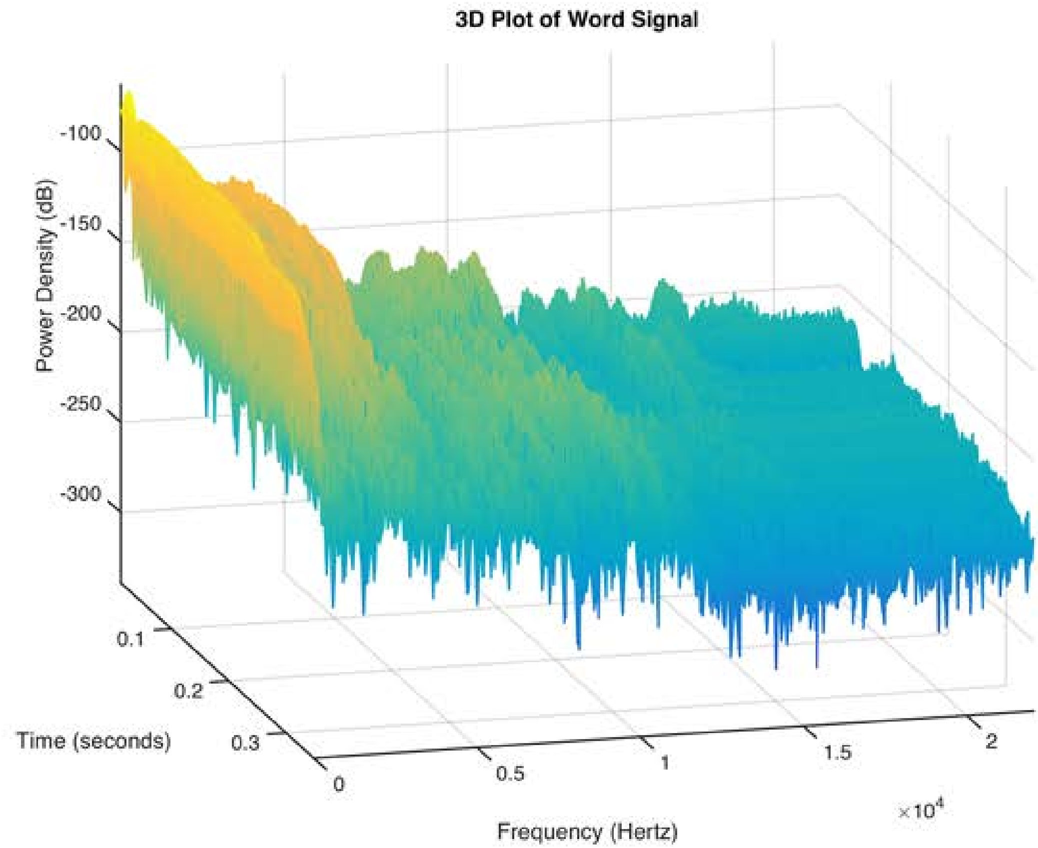
/ Z /

'Azure'

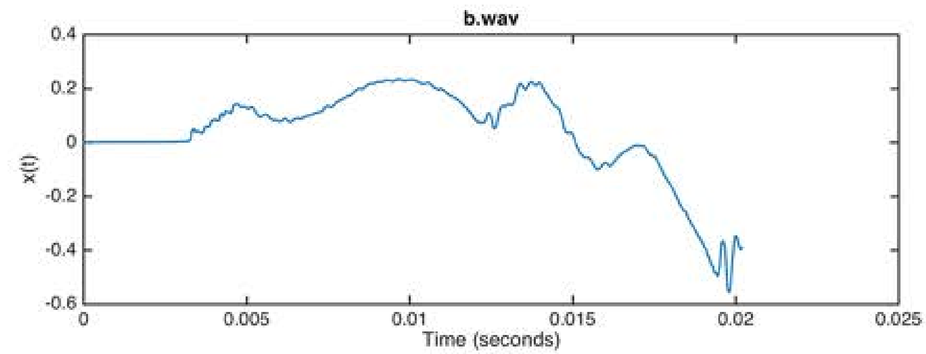

Capital_z.wav
x(t)
Time (seconds)

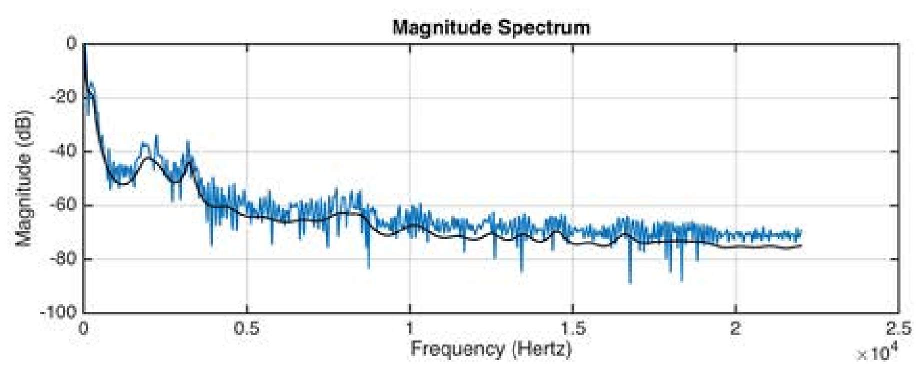

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

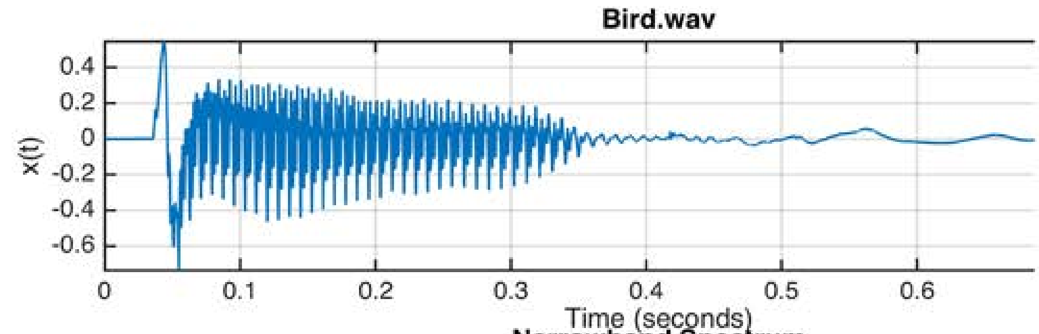

Bee.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

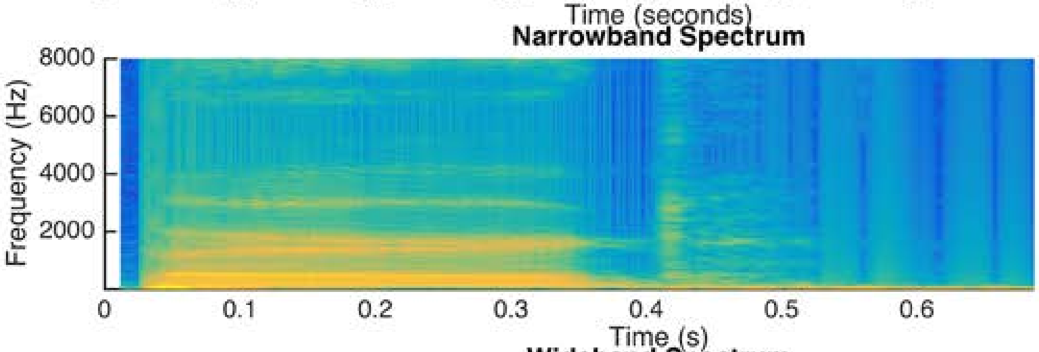

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
$\times 10^4$

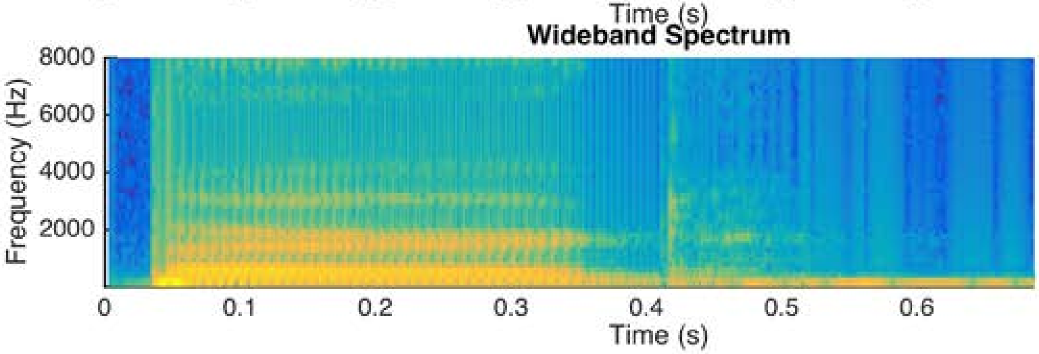
/b/ ‘Be’


b.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
$\times 10^4$

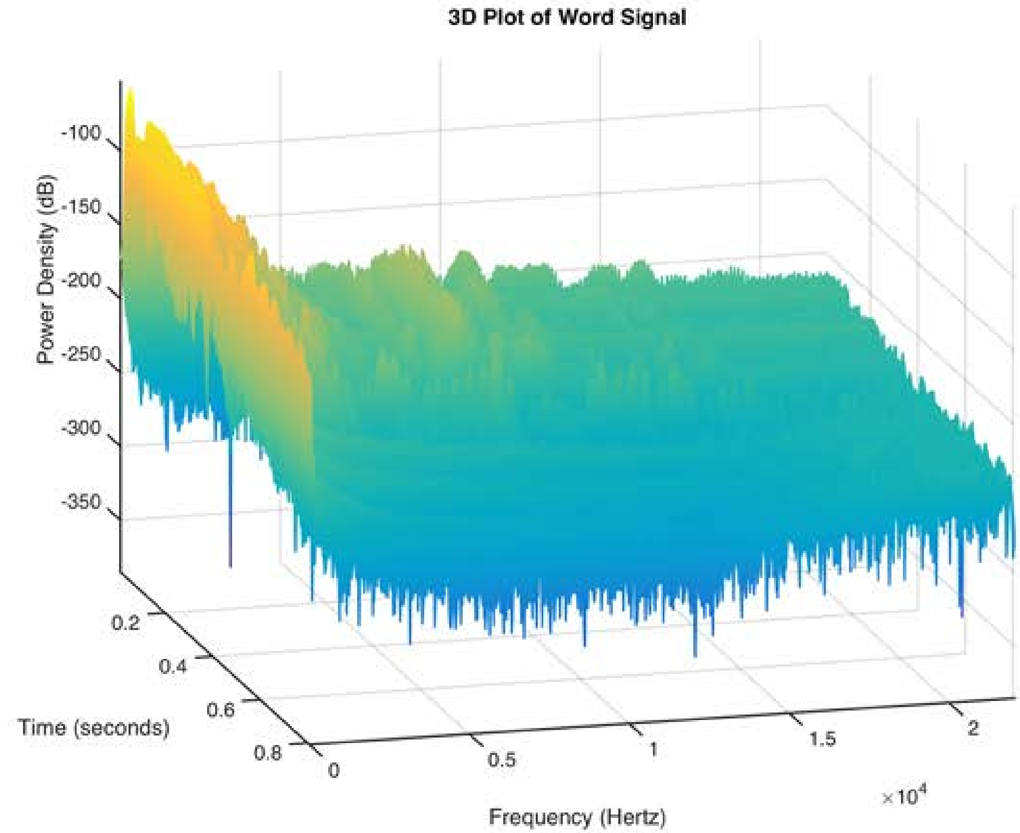

Bird.wav
x(t)
Time (seconds)

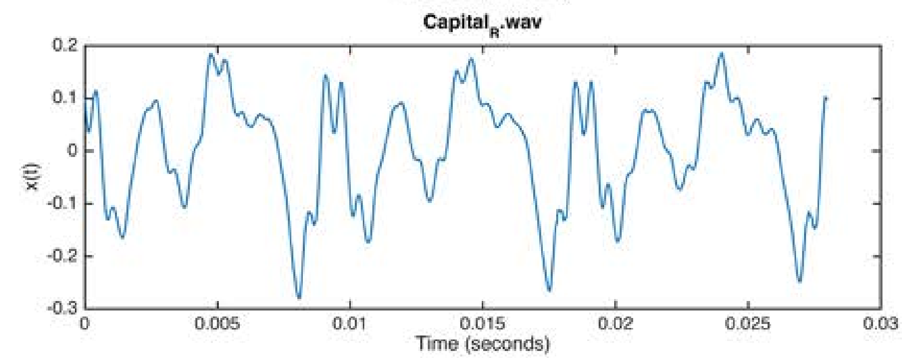

Narrowband Spectrum
Frequency (Hz)
Time (s)

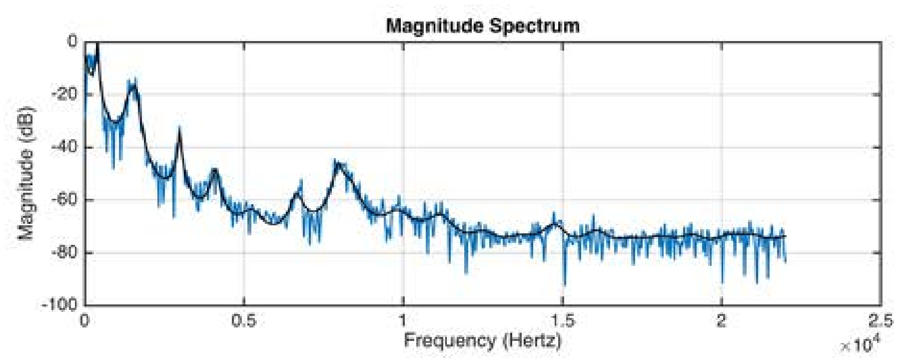

Wideband Spectrum
Frequency (Hz)
Time (s)

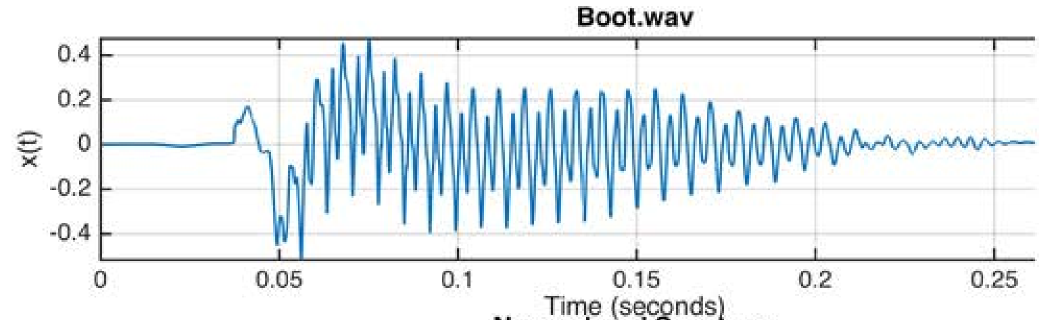
/R/

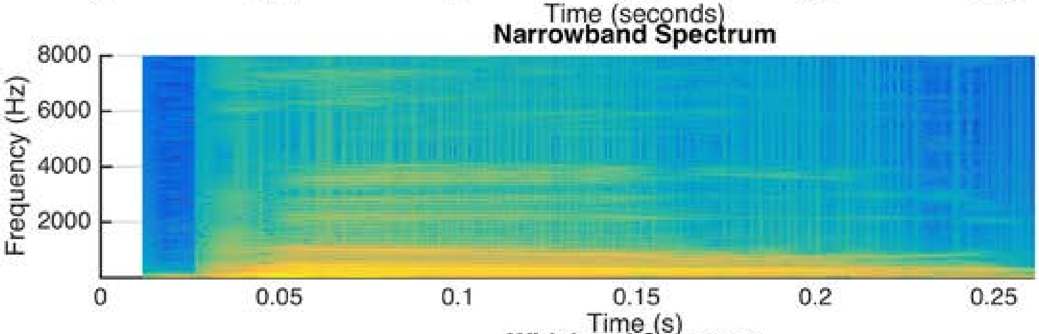
'Bird'

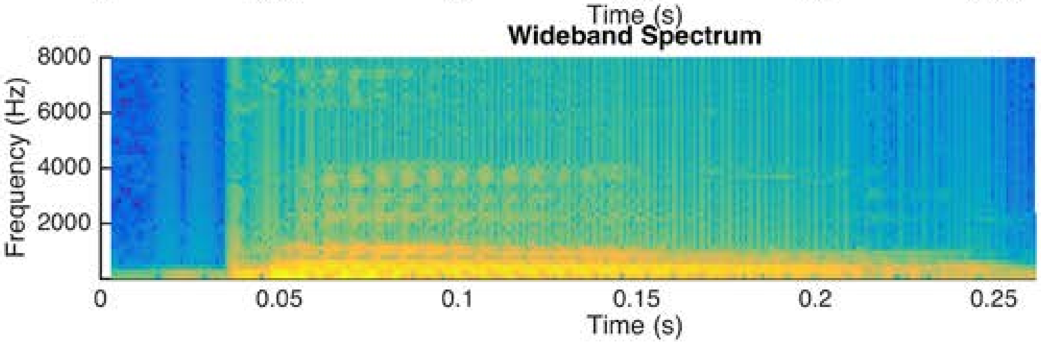

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

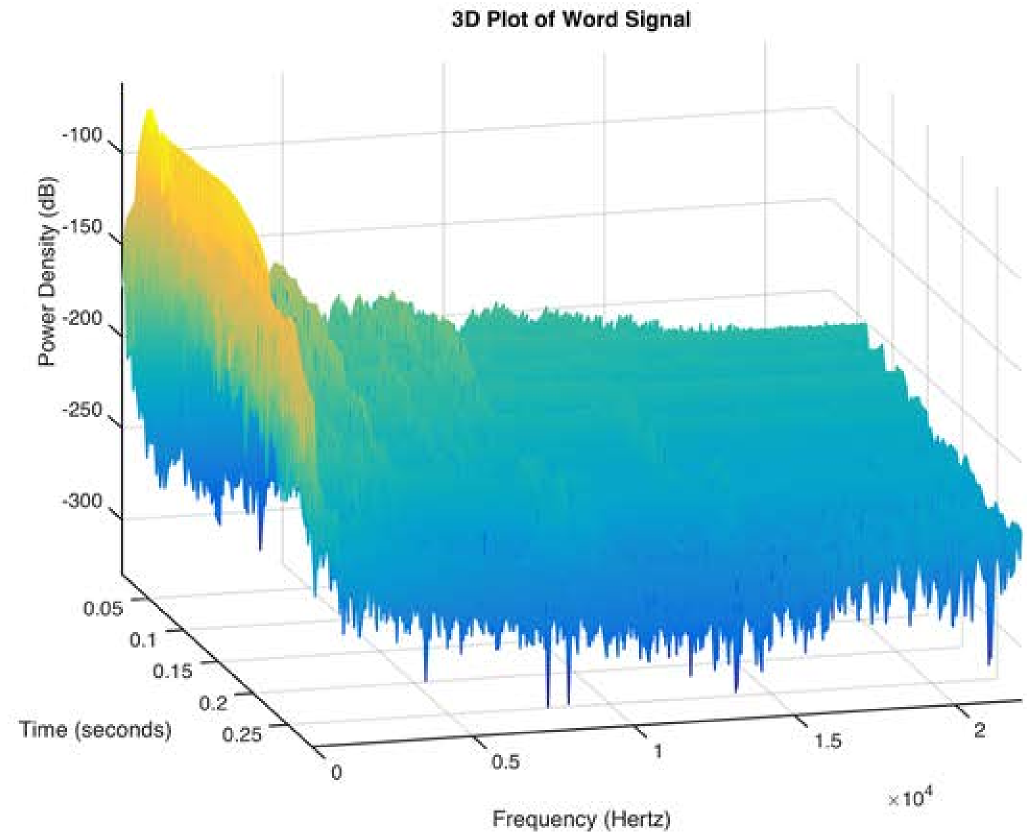

Capital_R.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


Boot.wav
x(t)
Time (seconds)

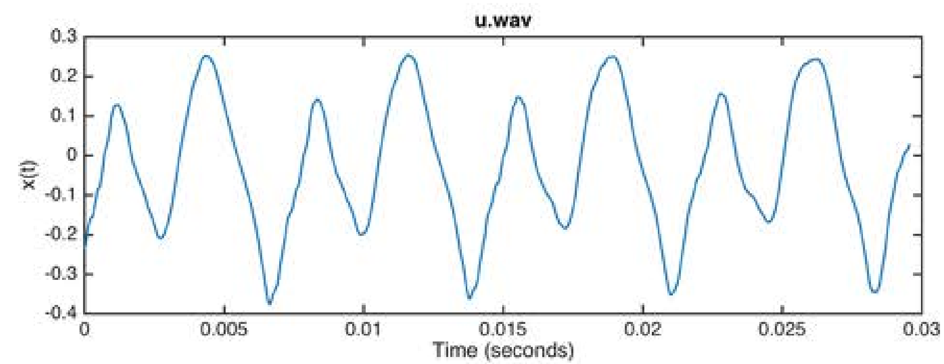

Narrowband Spectrum
Frequency (Hz)
Time (s)

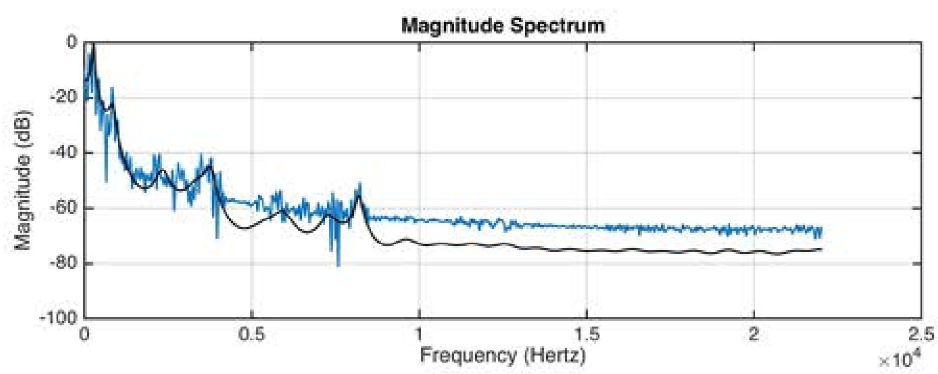

Wideband Spectrum
Frequency (Hz)
Time (s)

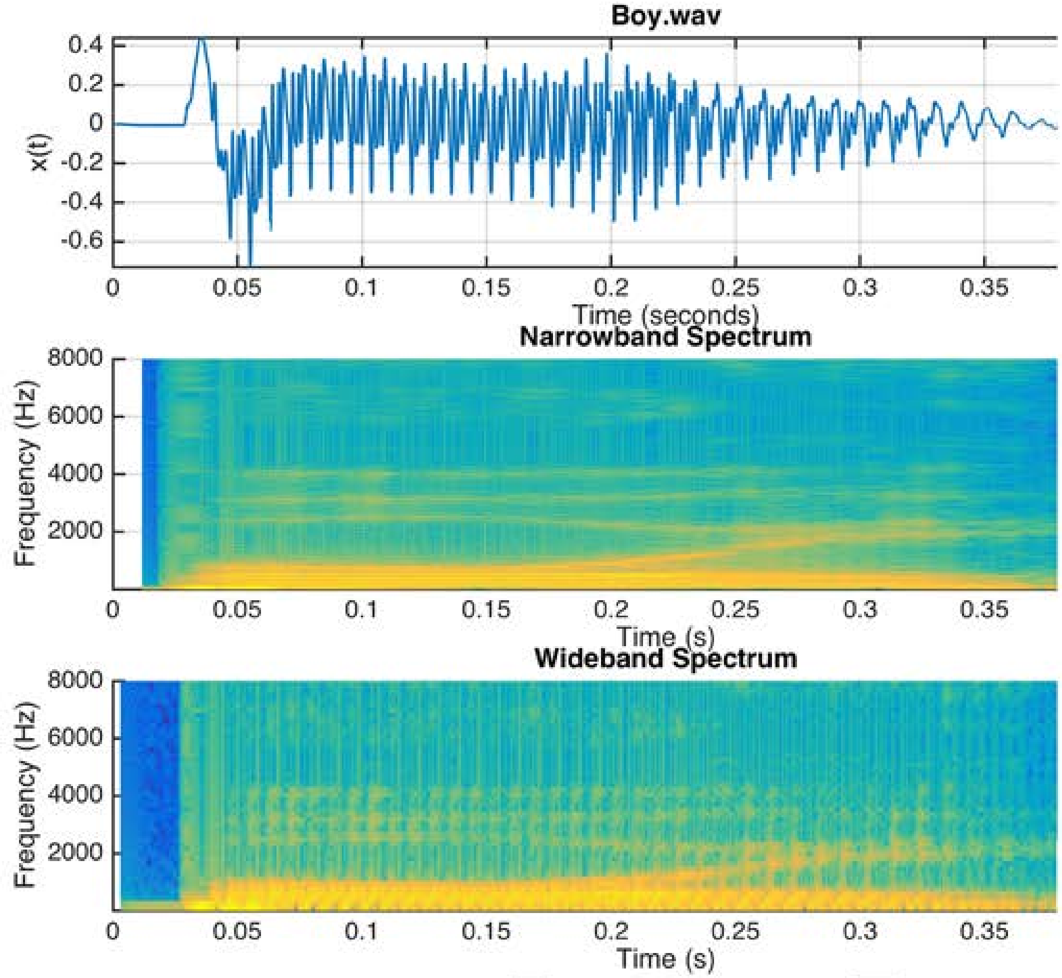

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

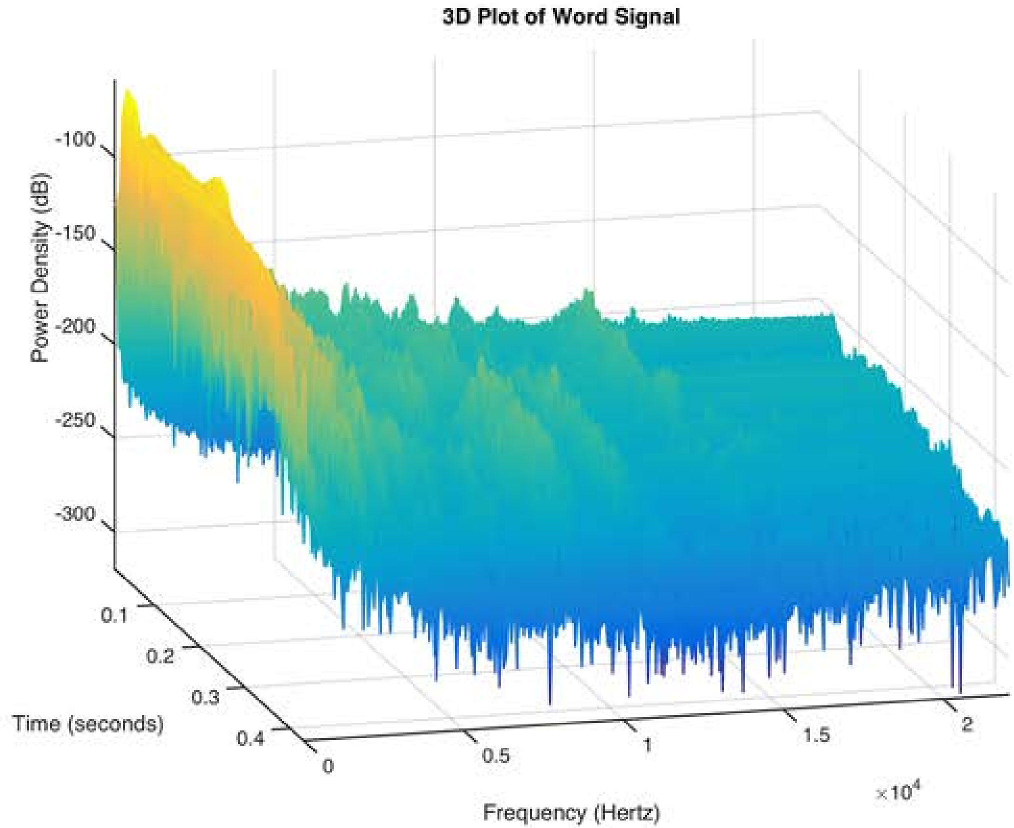
/u/

'Boot'


u.wav
x(t)
Time (seconds)

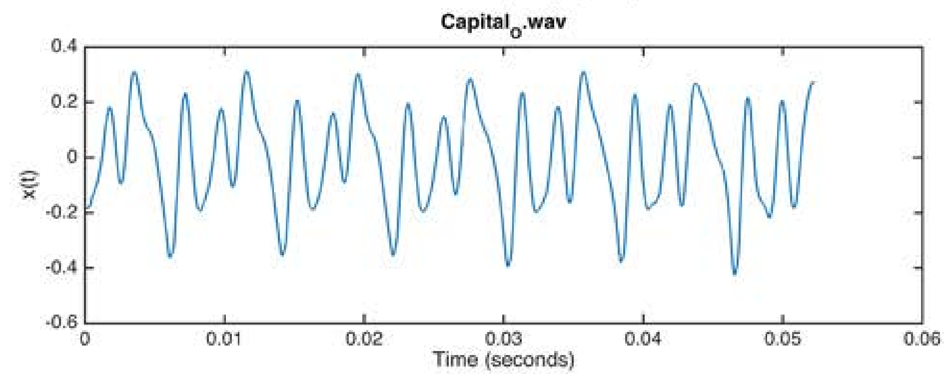

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

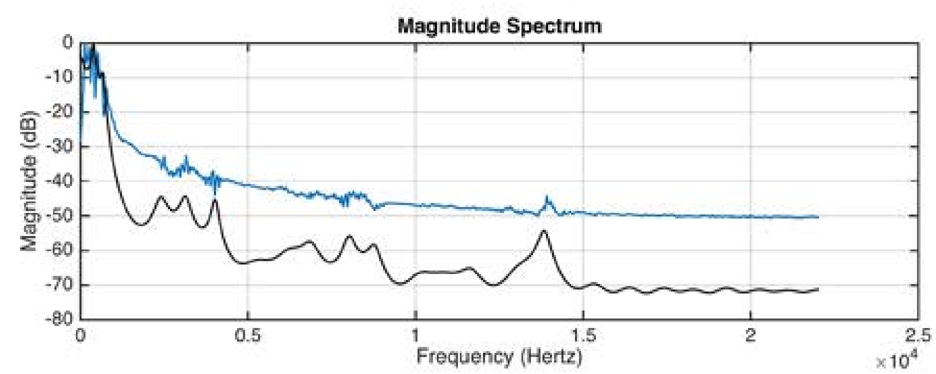

Boy.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

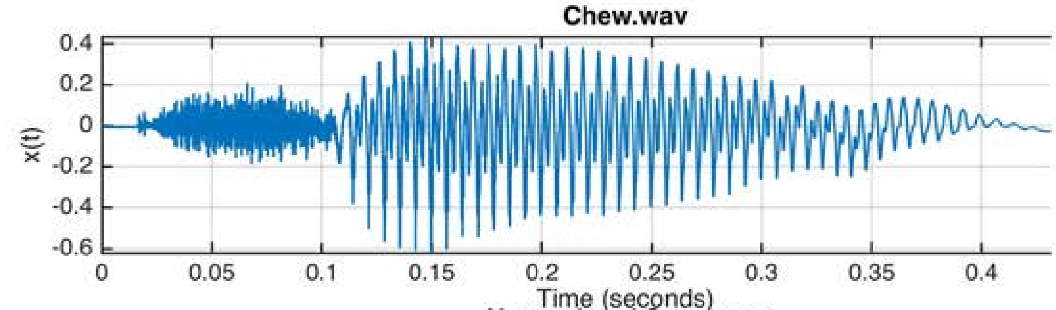

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

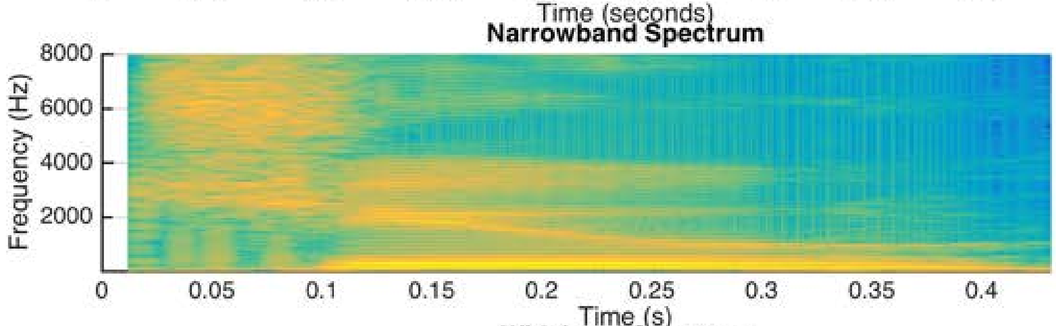
/O/

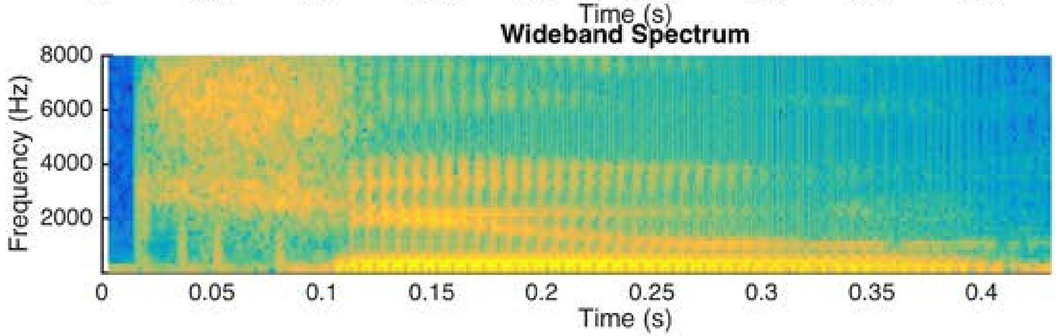
‘Boy’

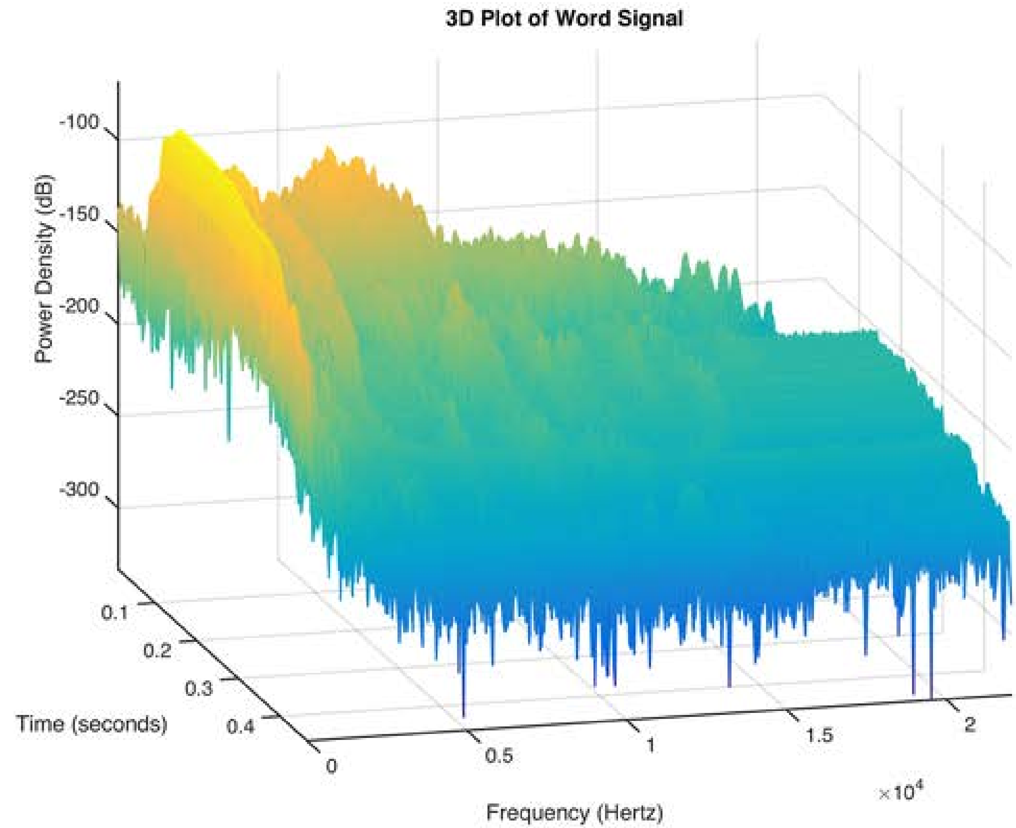

Capital_o.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

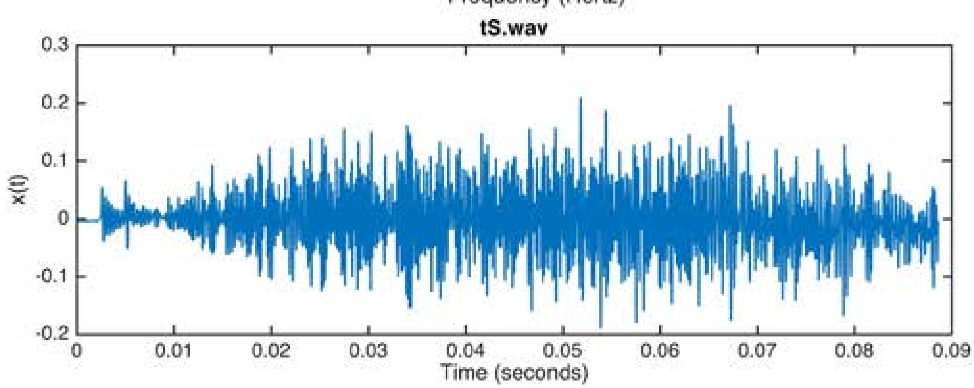

Chew.wav
x(t)
Time (seconds)

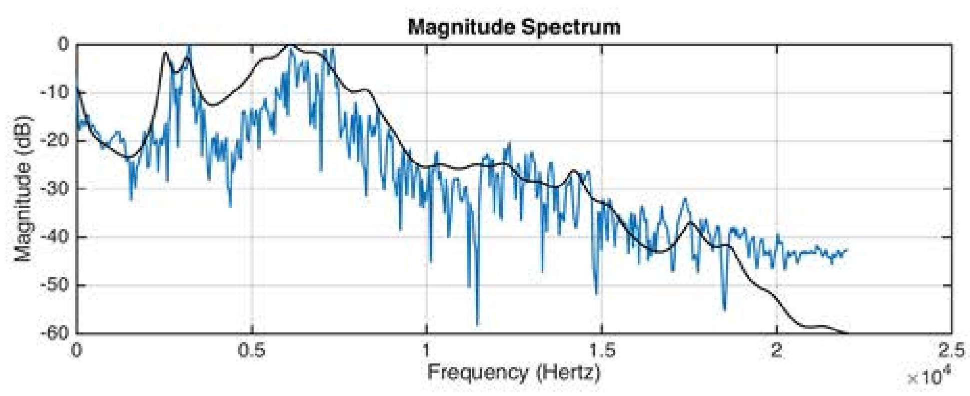

Narrowband Spectrum
Frequency (Hz)
Time (s)

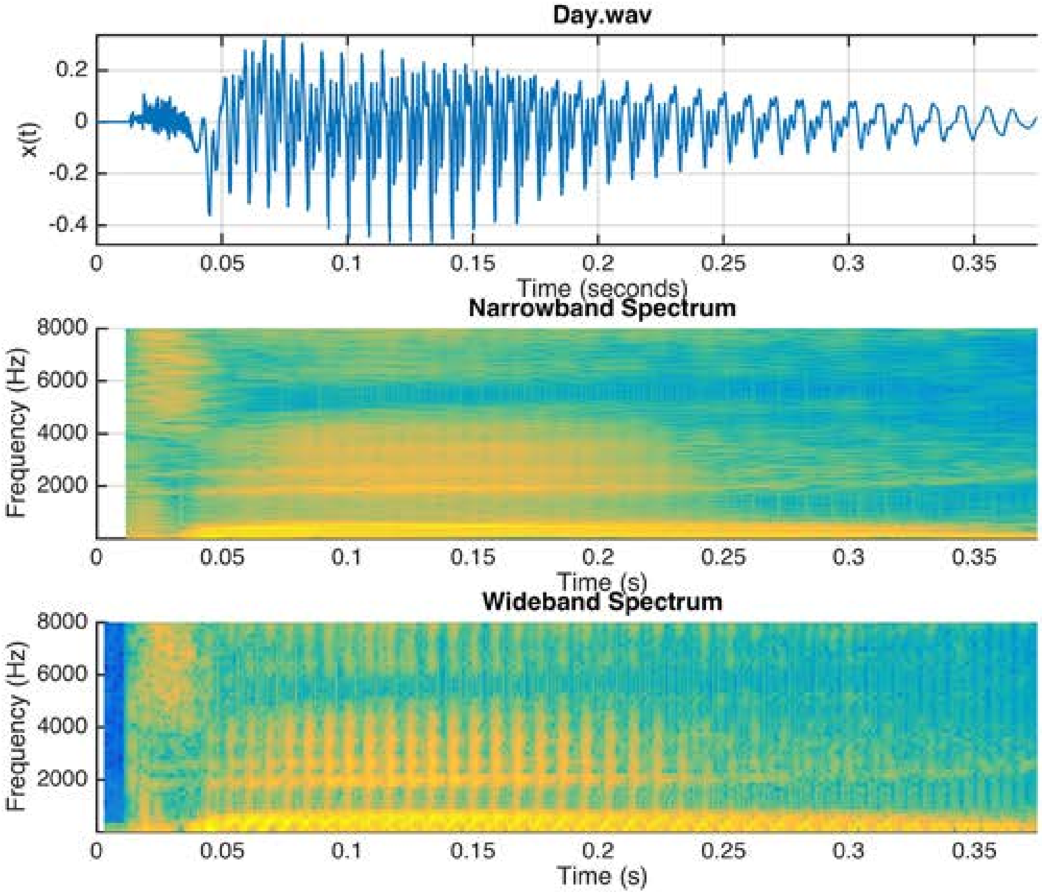

Wideband Spectrum
Frequency (Hz)
Time (s)

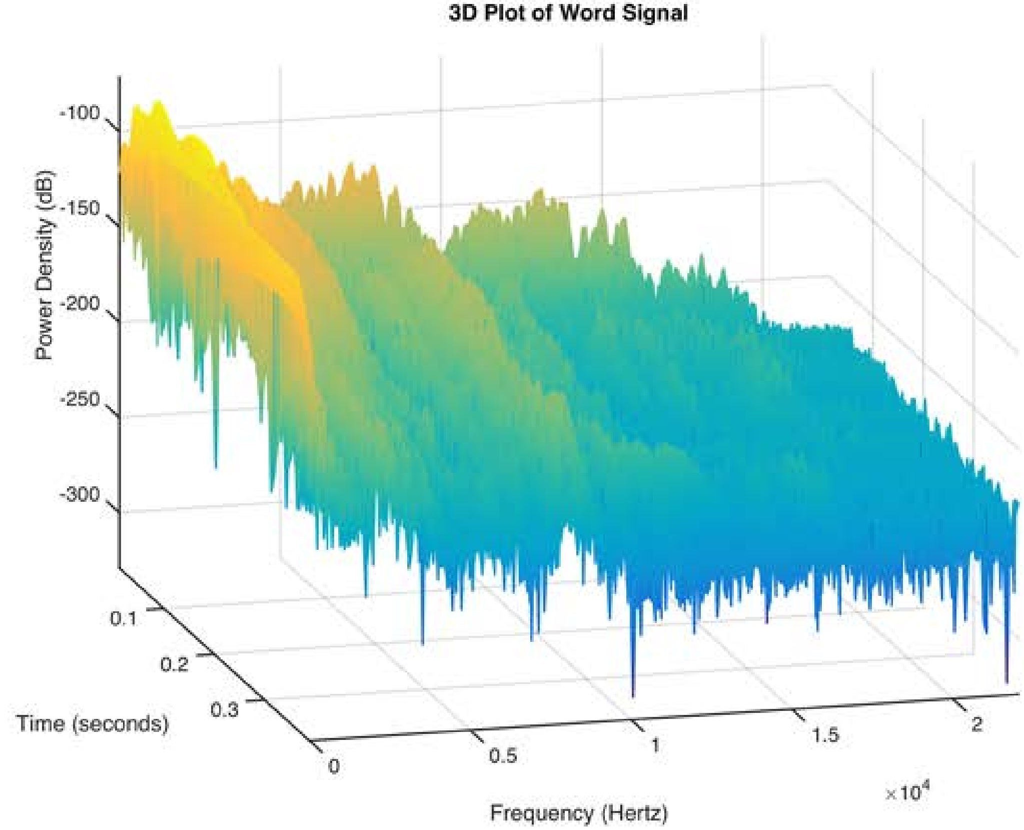

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

/tS/
‘Chew’


tS.wav
x(t)
Time (seconds)

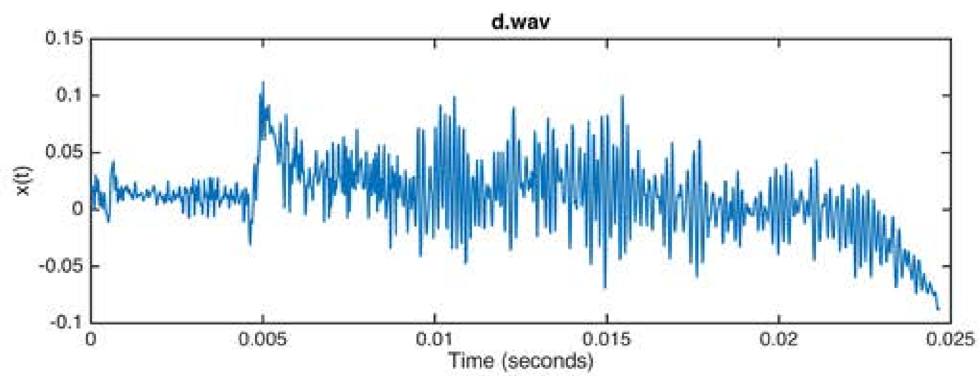

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

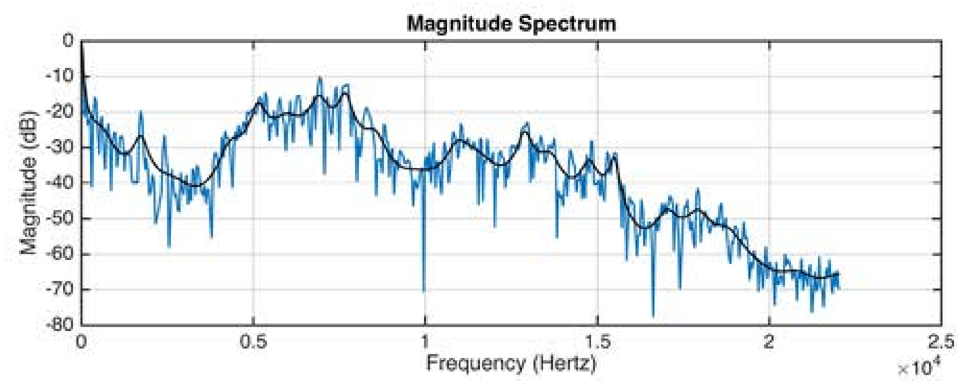

Day.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

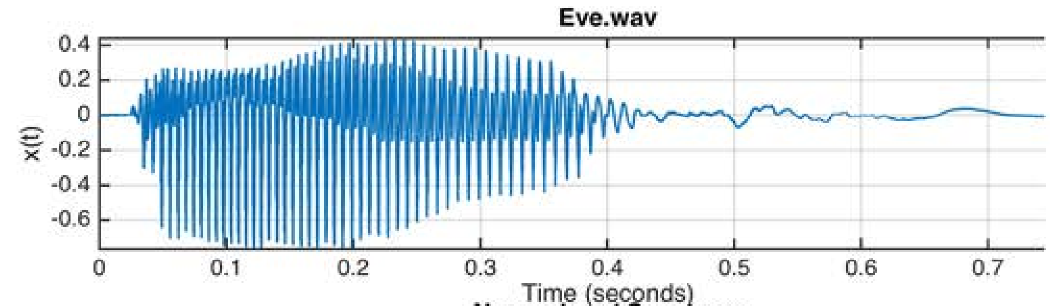

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

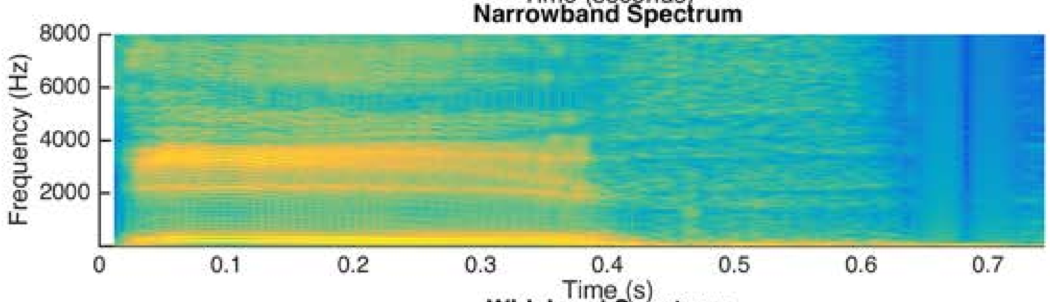
/d/

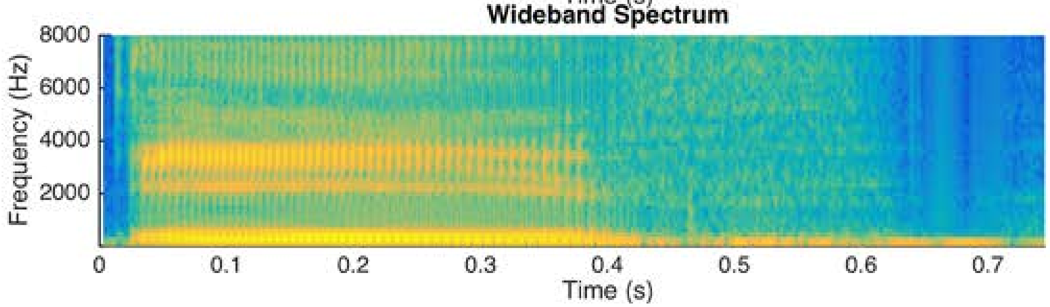
‘Day’

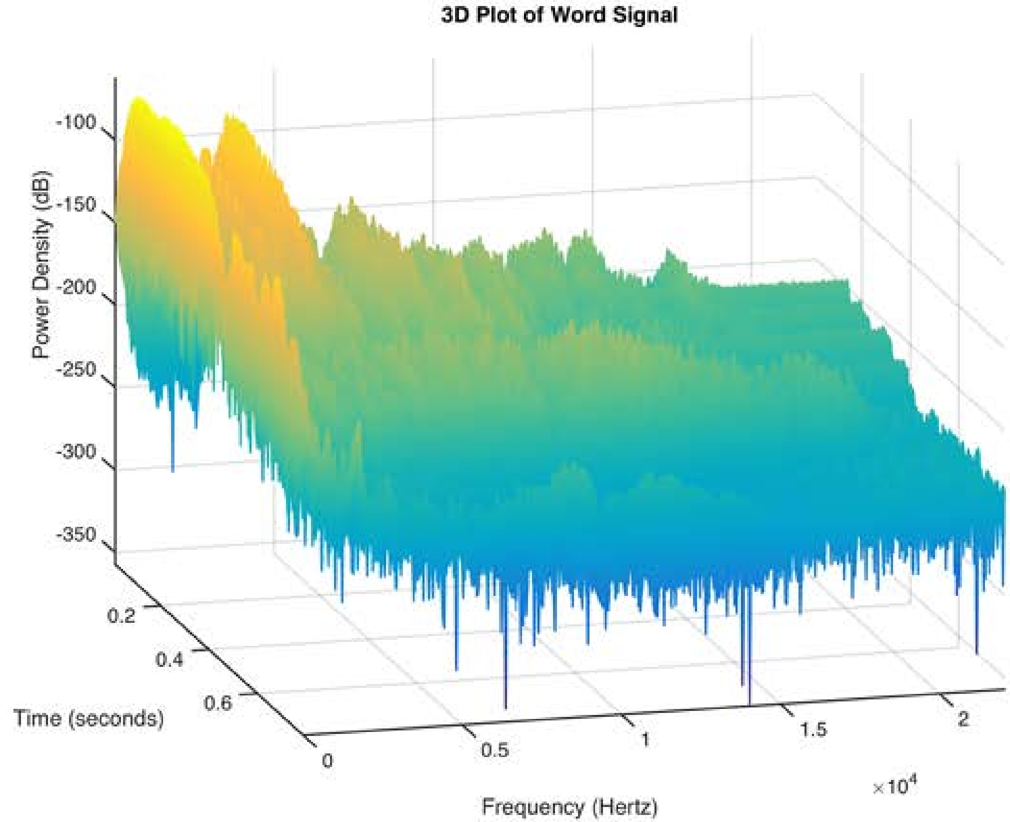

d.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


Eve.wav
x(t)
Time (seconds)

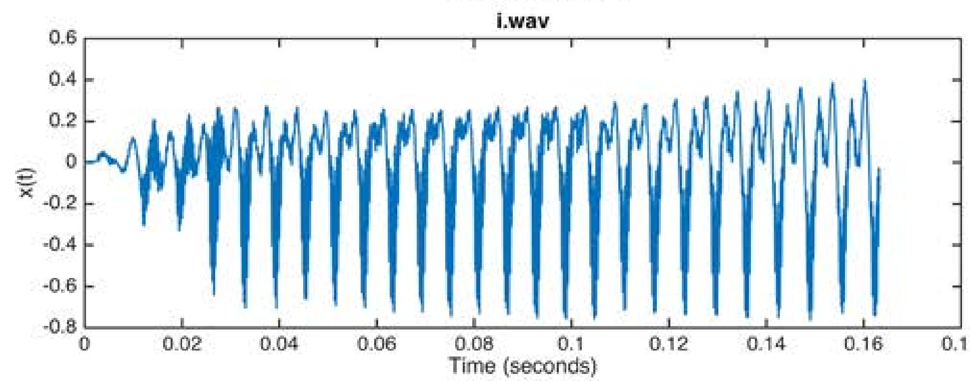

Narrowband Spectrum
Frequency (Hz)
Time (s)

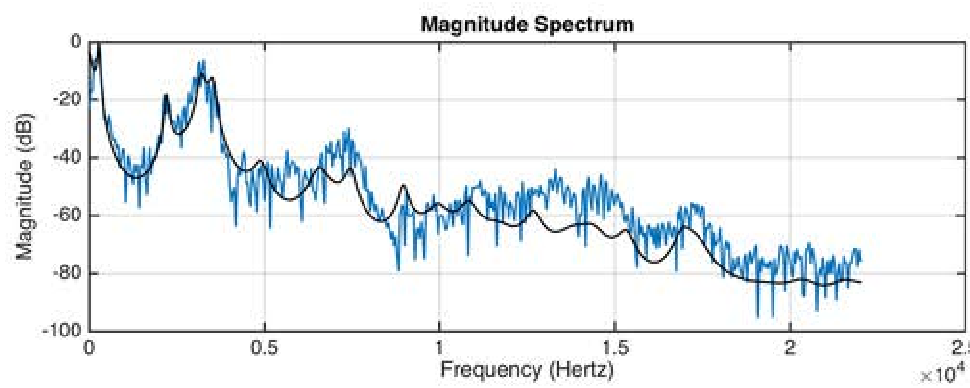

Wideband Spectrum
Frequency (Hz)
Time (s)

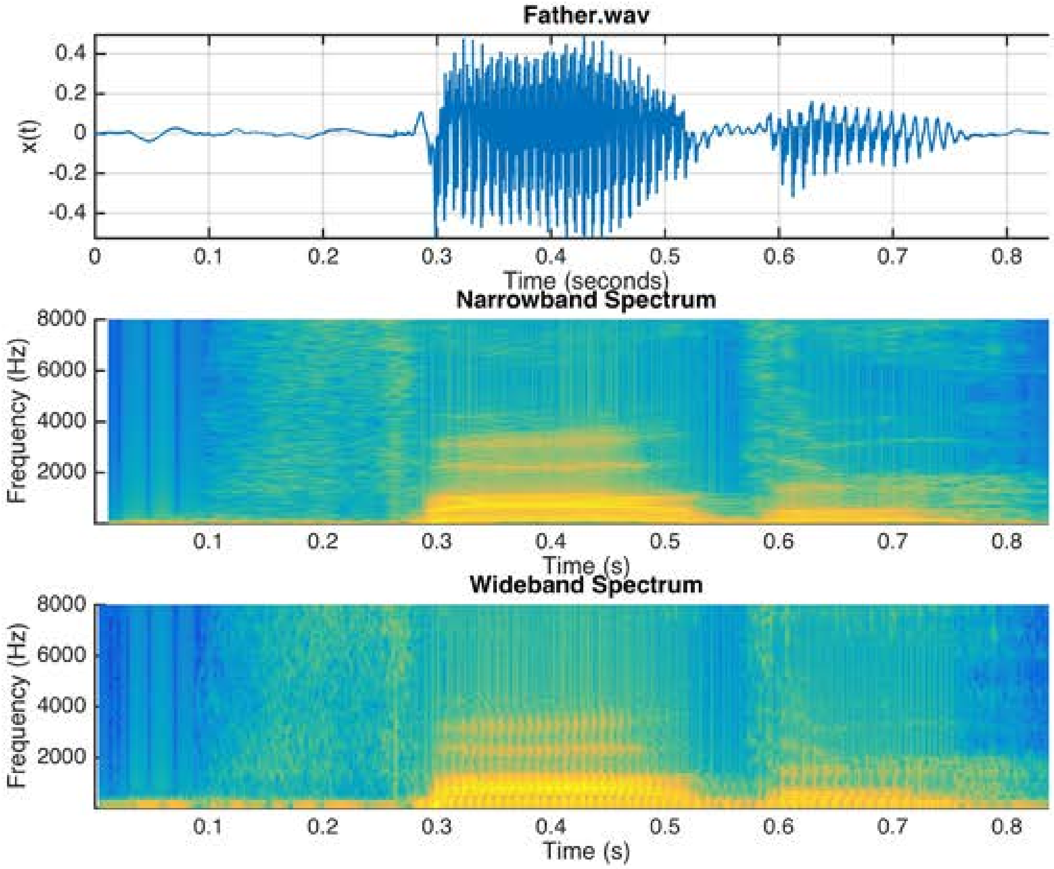

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

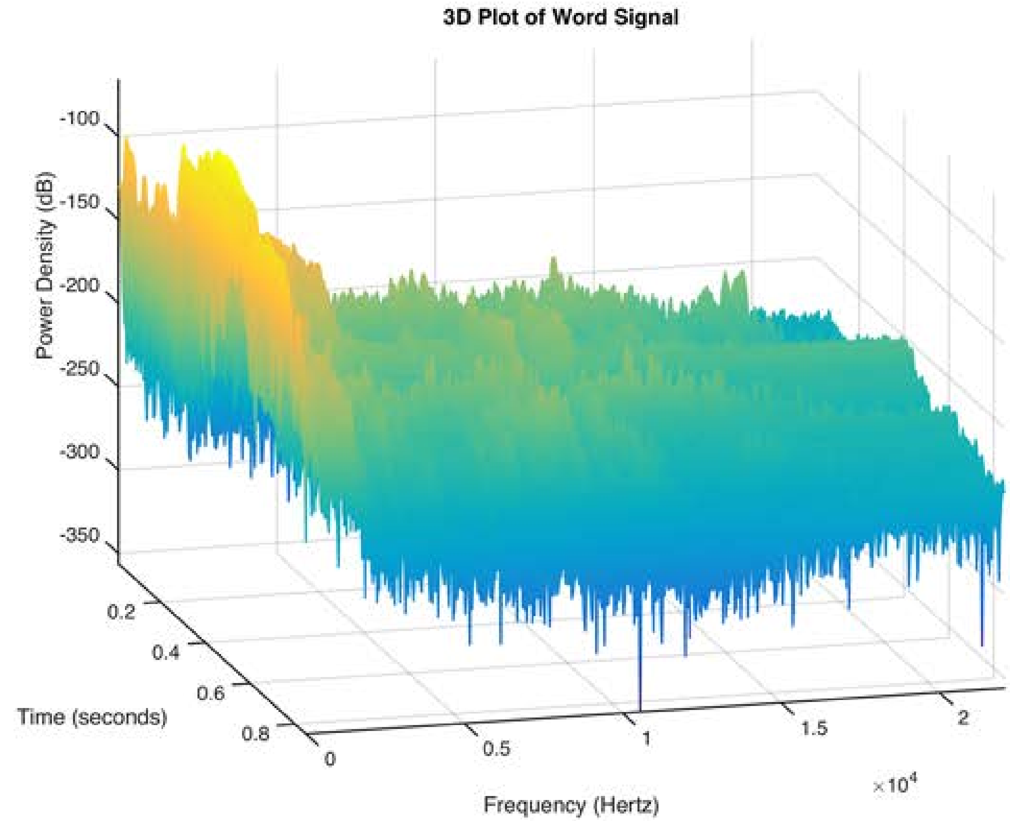
/i/

'Eve'


i.wav
x(t)
Time (seconds)

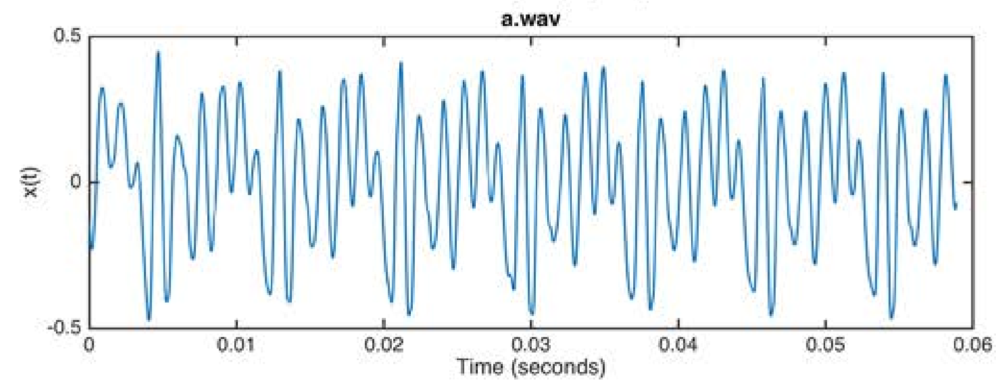

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

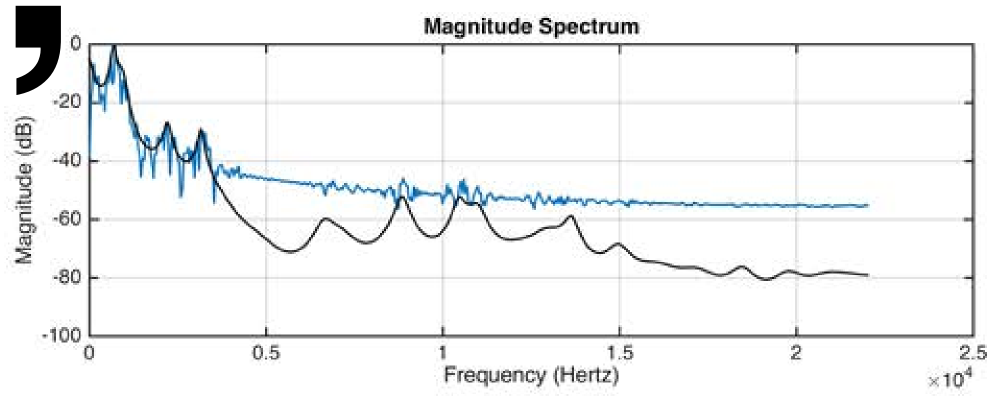

Father.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

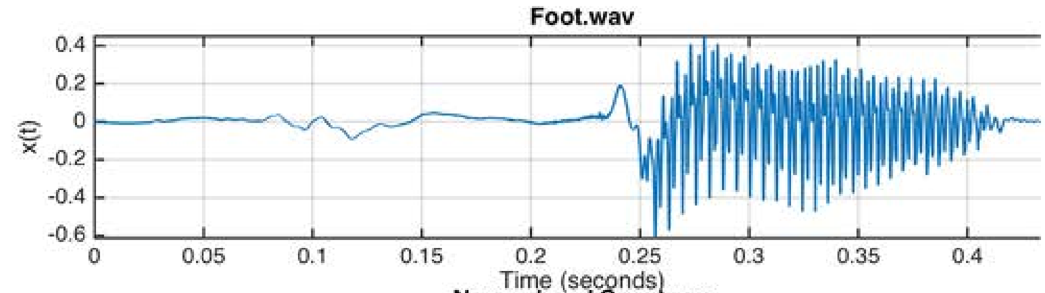

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

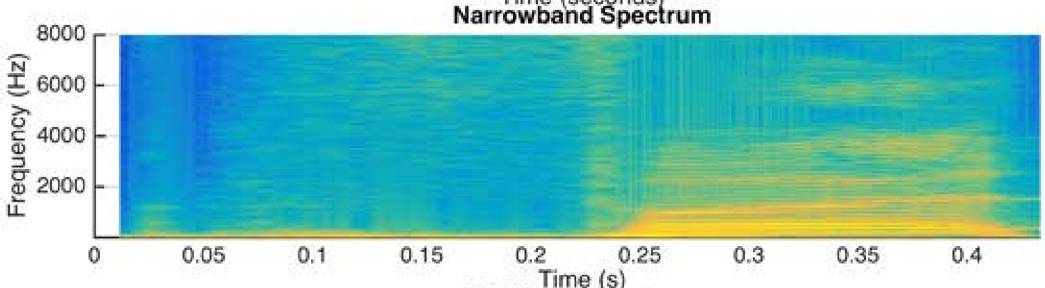
/a/

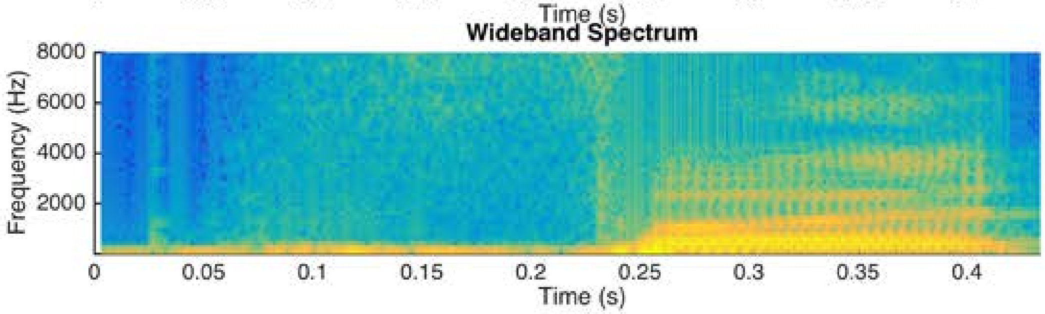
‘Father’

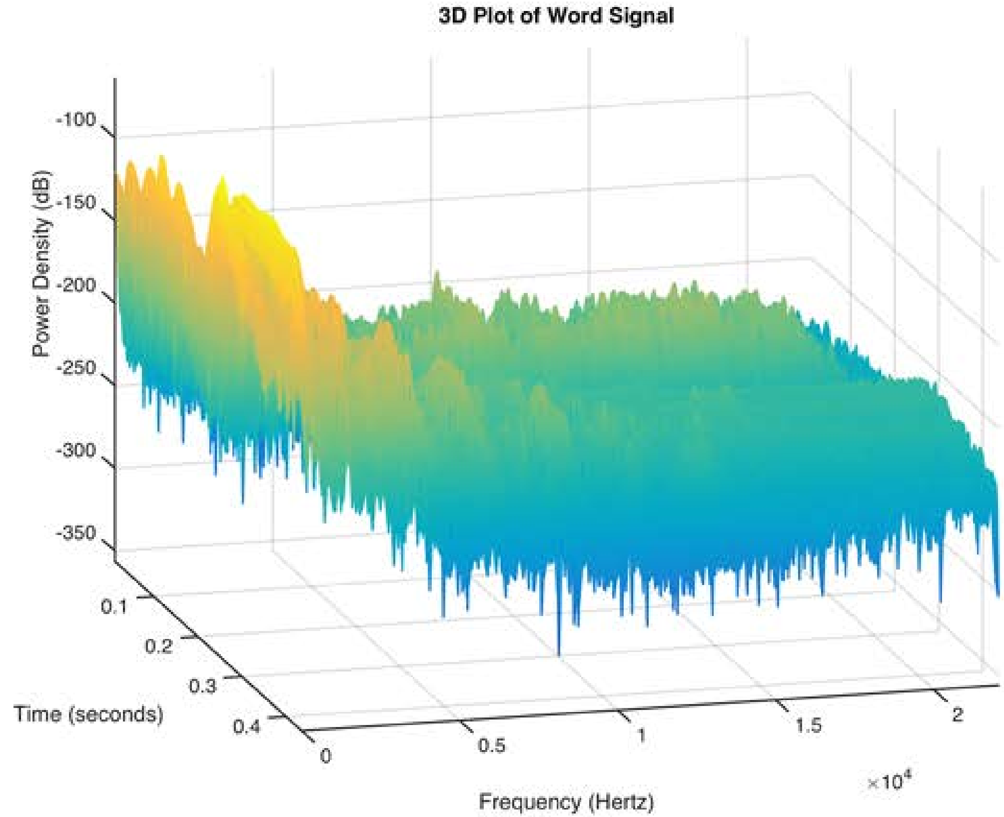

a.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


Foot.wav
x(t)
Time (seconds)

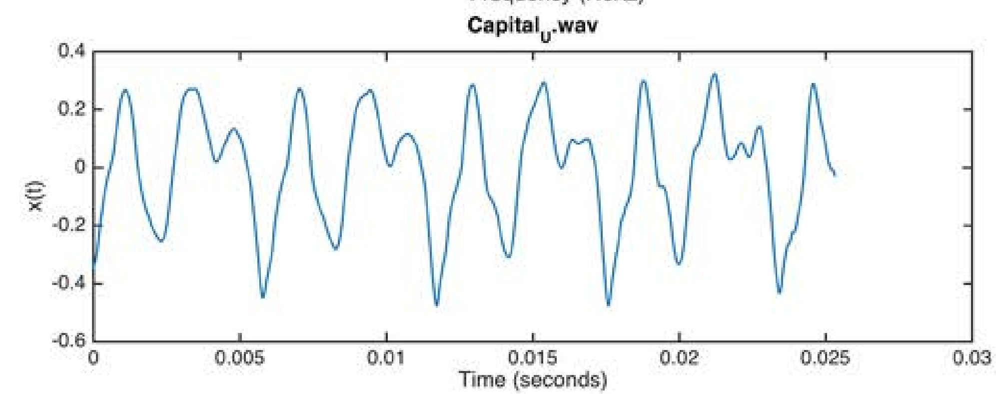

Narrowband Spectrum
Frequency (Hz)
Time (s)

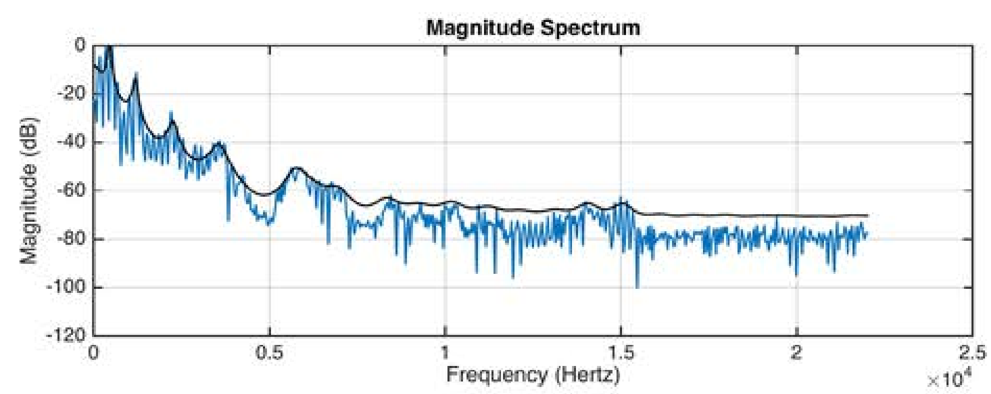

Wideband Spectrum
Frequency (Hz)
Time (s)

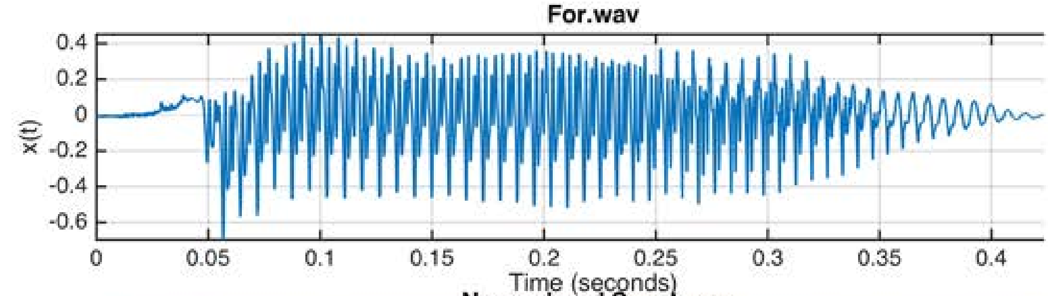

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

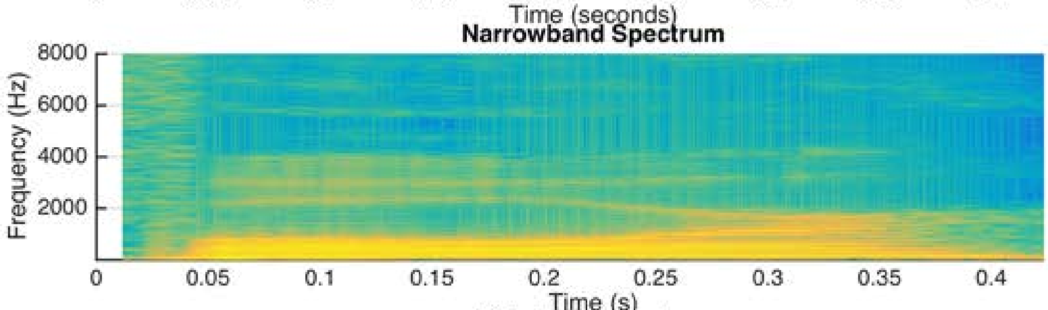
/U/

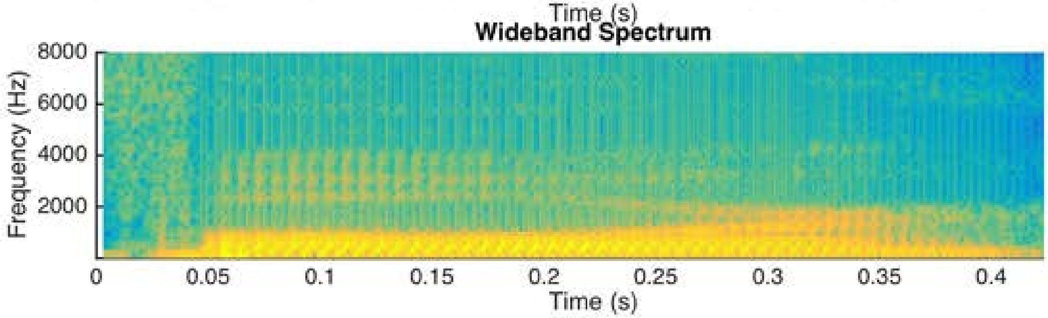
'Foot'

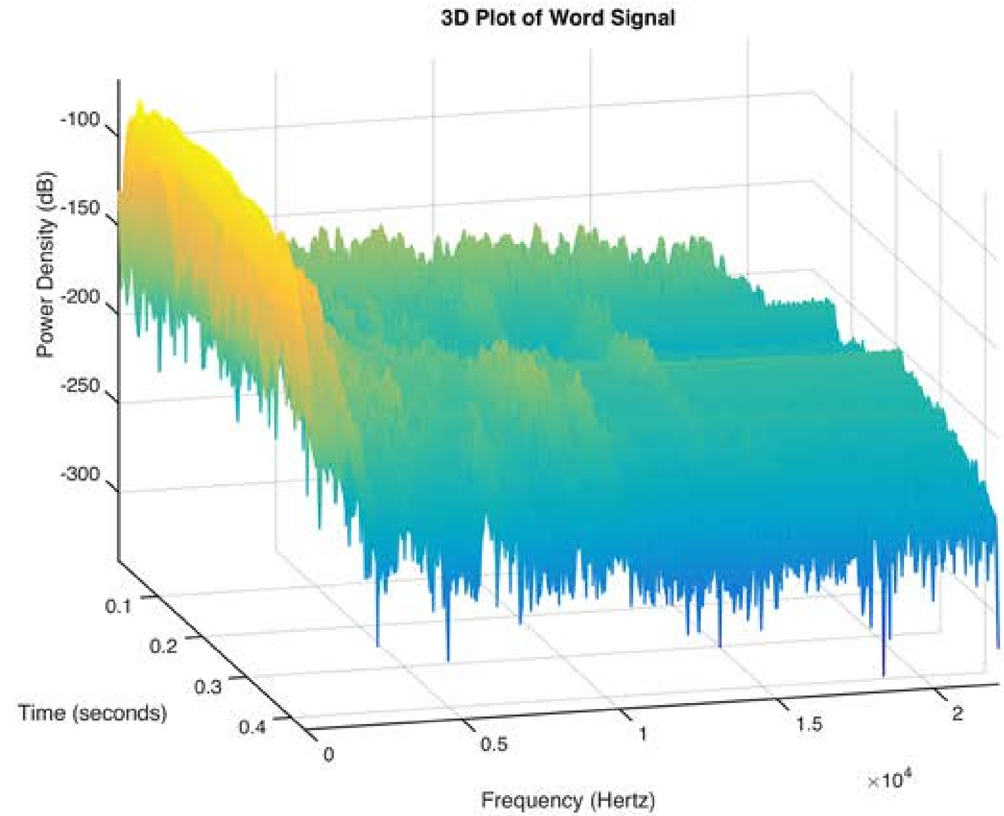

Capital_U.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


For.wav
x(t)
Time (seconds)

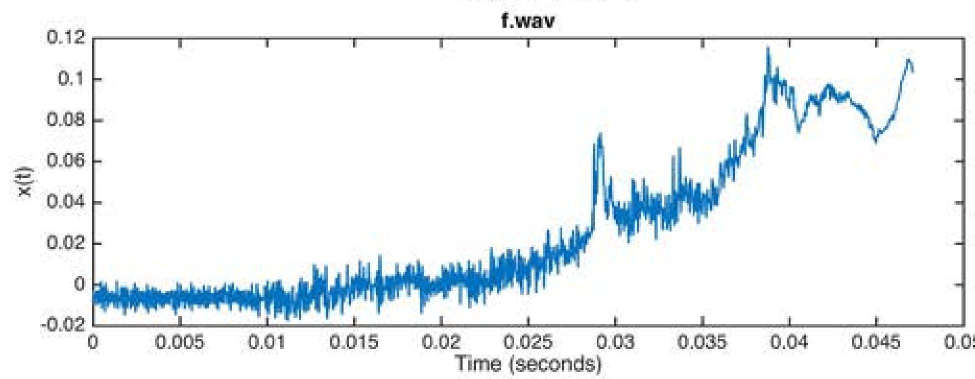

Narrowband Spectrum
Frequency (Hz)
Time (s)

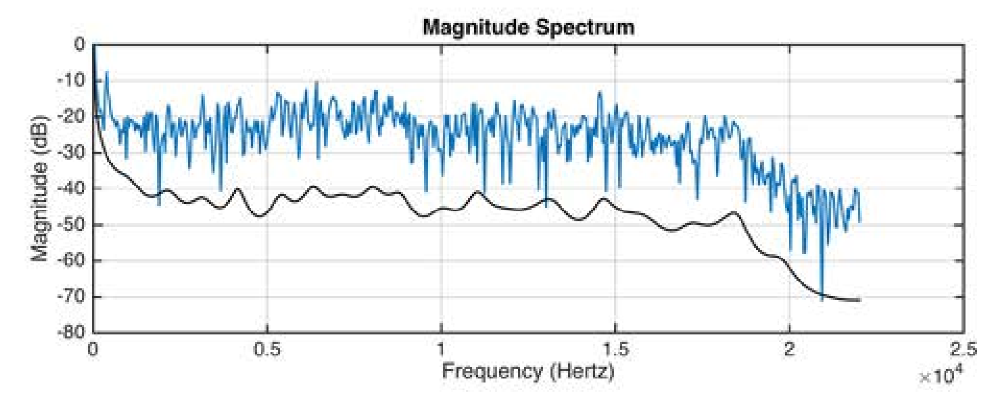

Wideband Spectrum
Frequency (Hz)
Time (s)

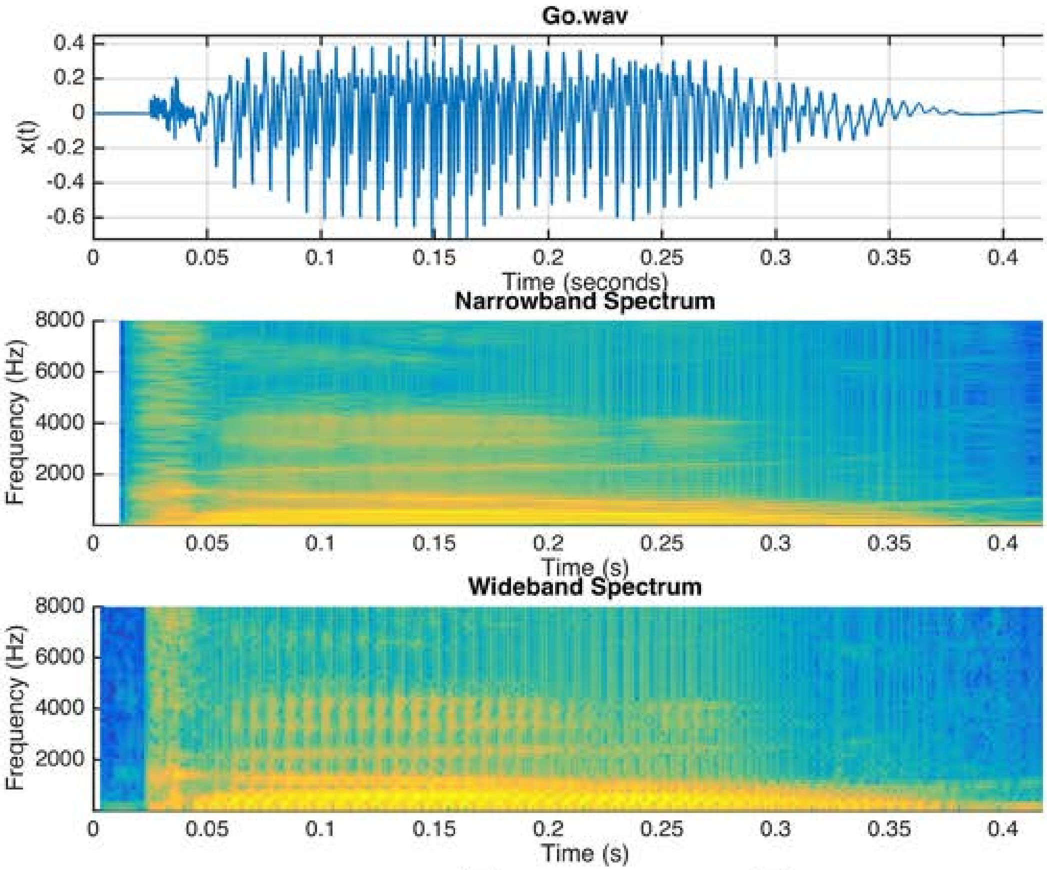

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

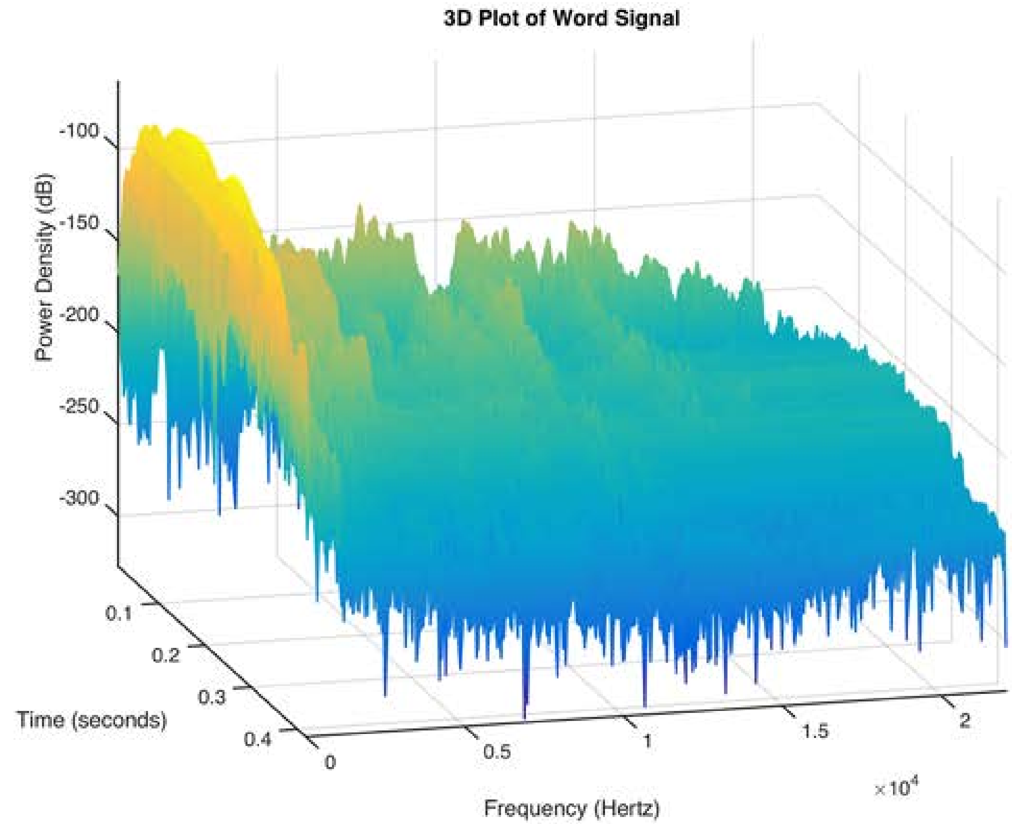
/f/

'For'


f.wav
x(t)
Time (seconds)

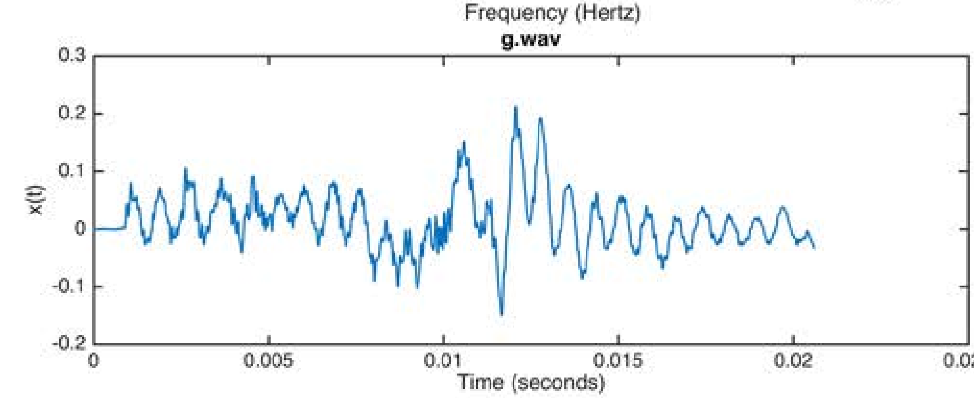

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

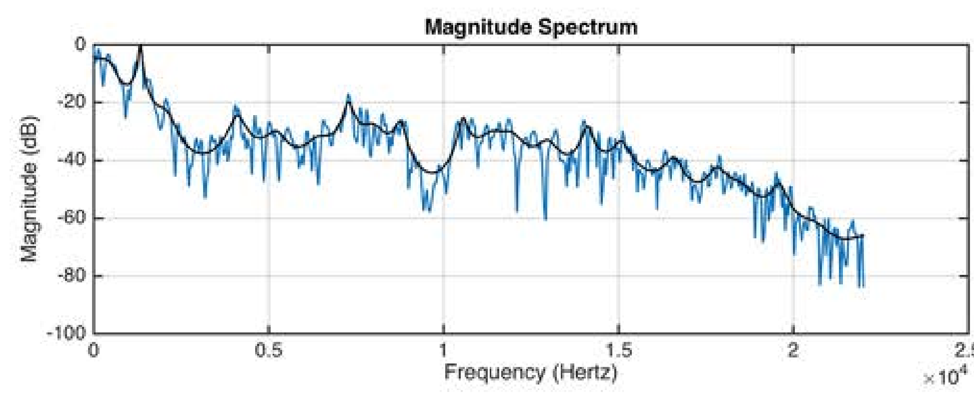

Go.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

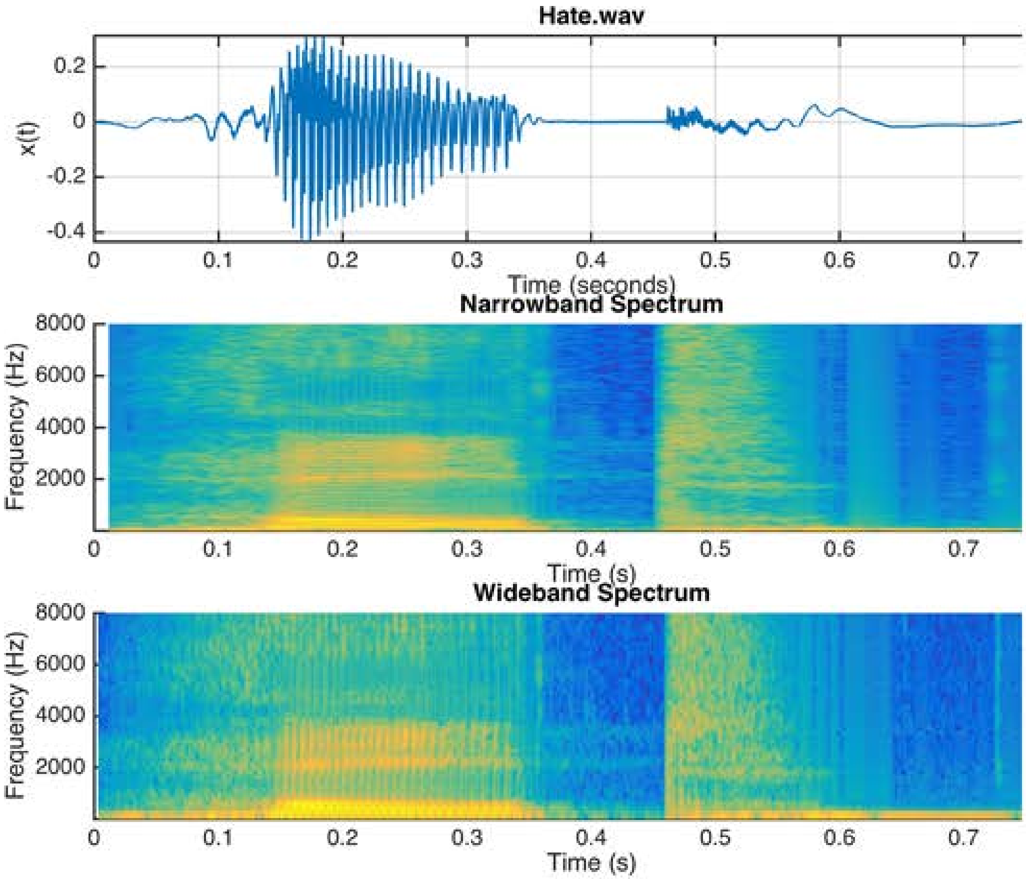

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

/g/

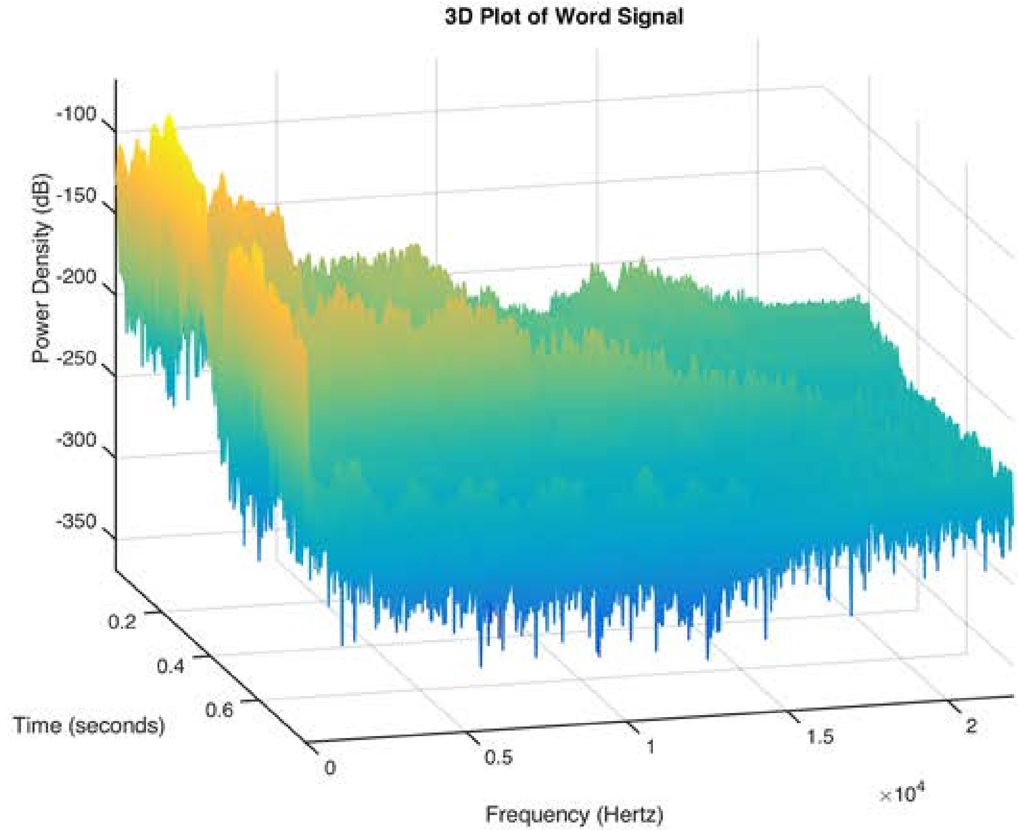
'Go'


g.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4



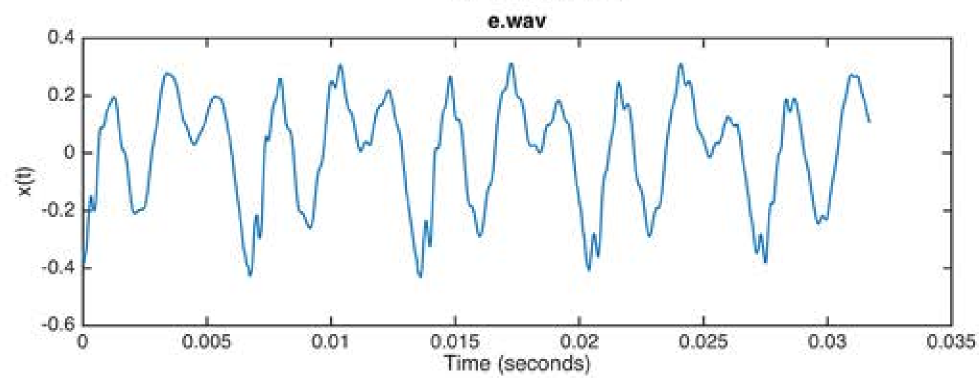
Hate.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)



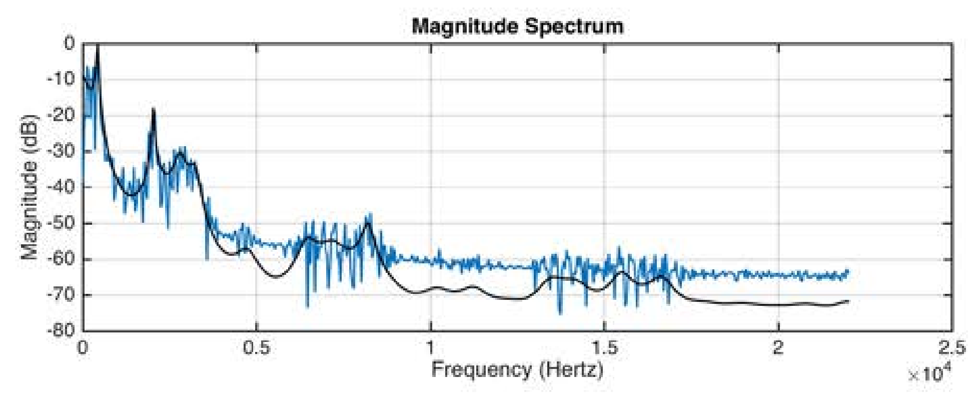
3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

/e/

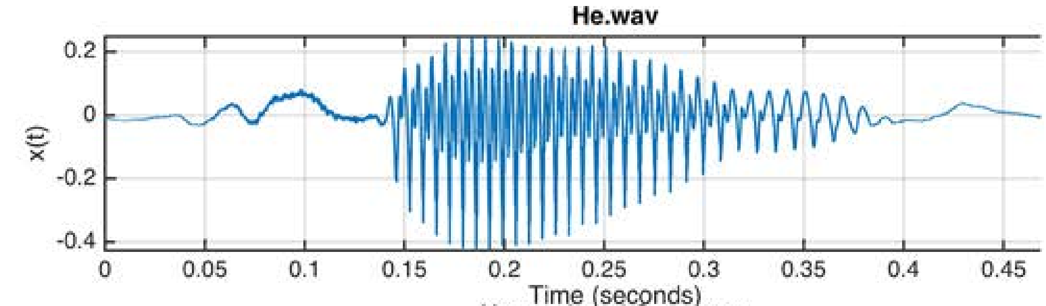
‘Hate’

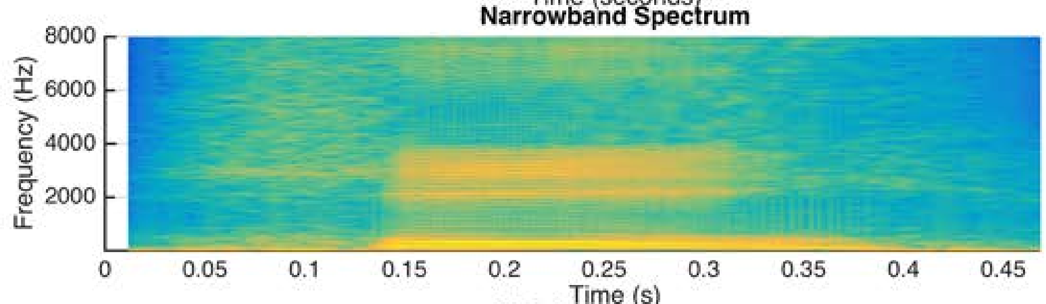


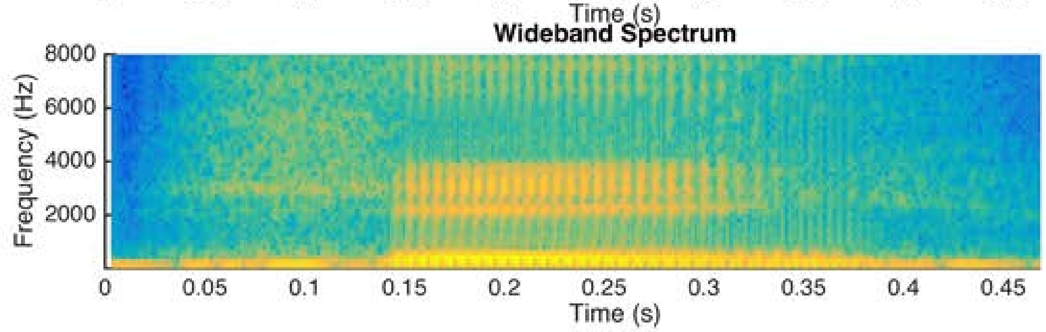
e.wav
x(t)
Time (seconds)

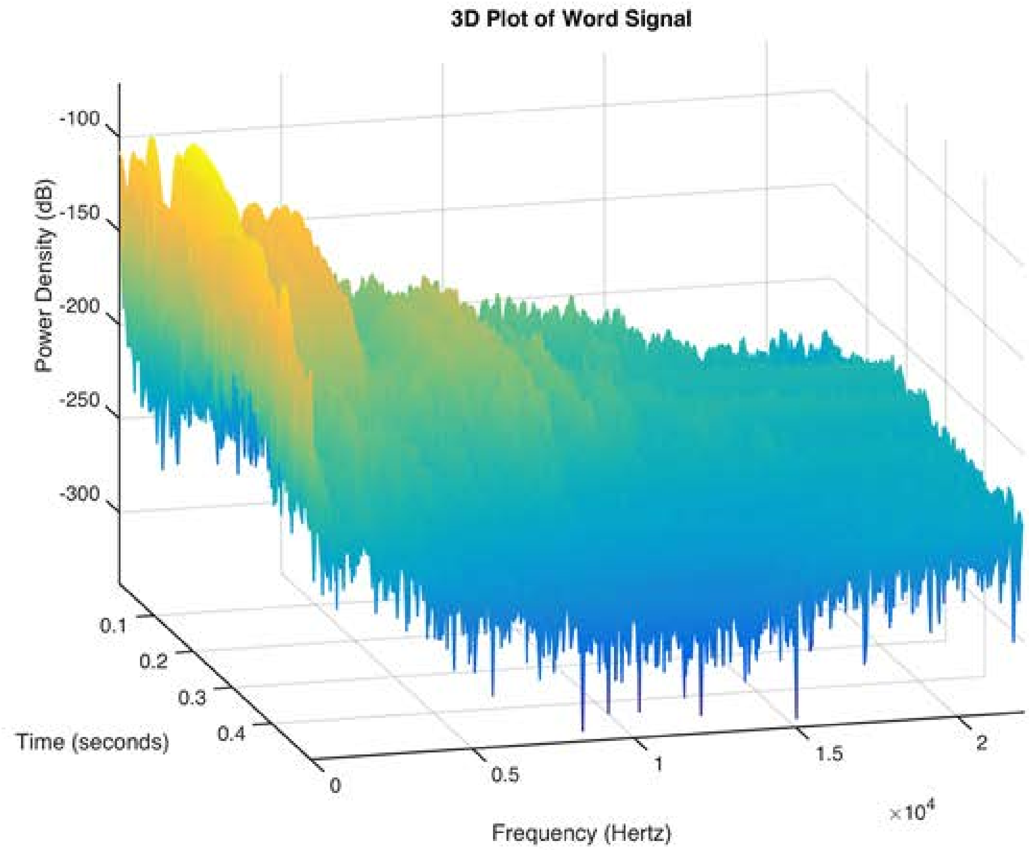


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


He.wav
x(t)
Time (seconds)

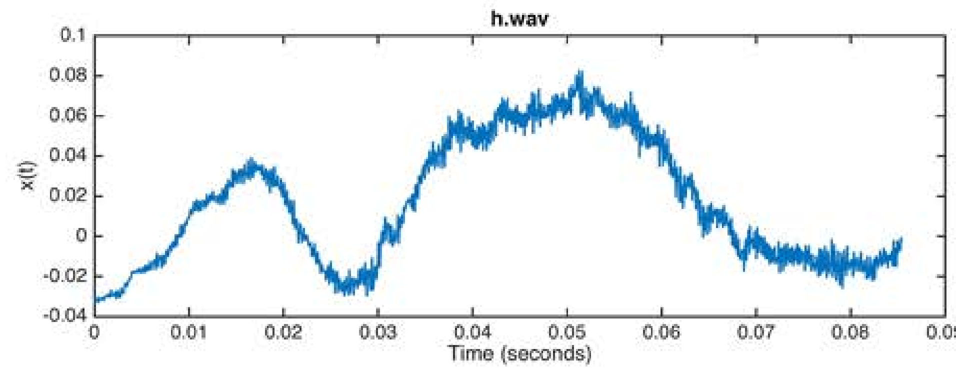

Narrowband Spectrum
Frequency (Hz)
Time (s)

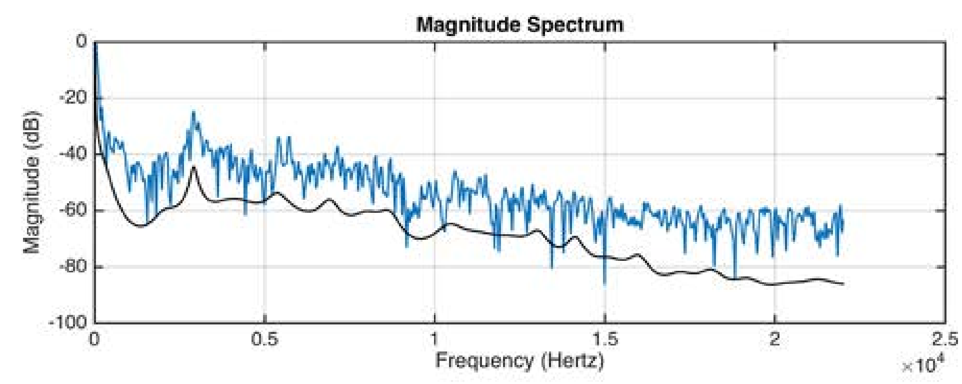

Wideband Spectrum
Frequency (Hz)
Time (s)

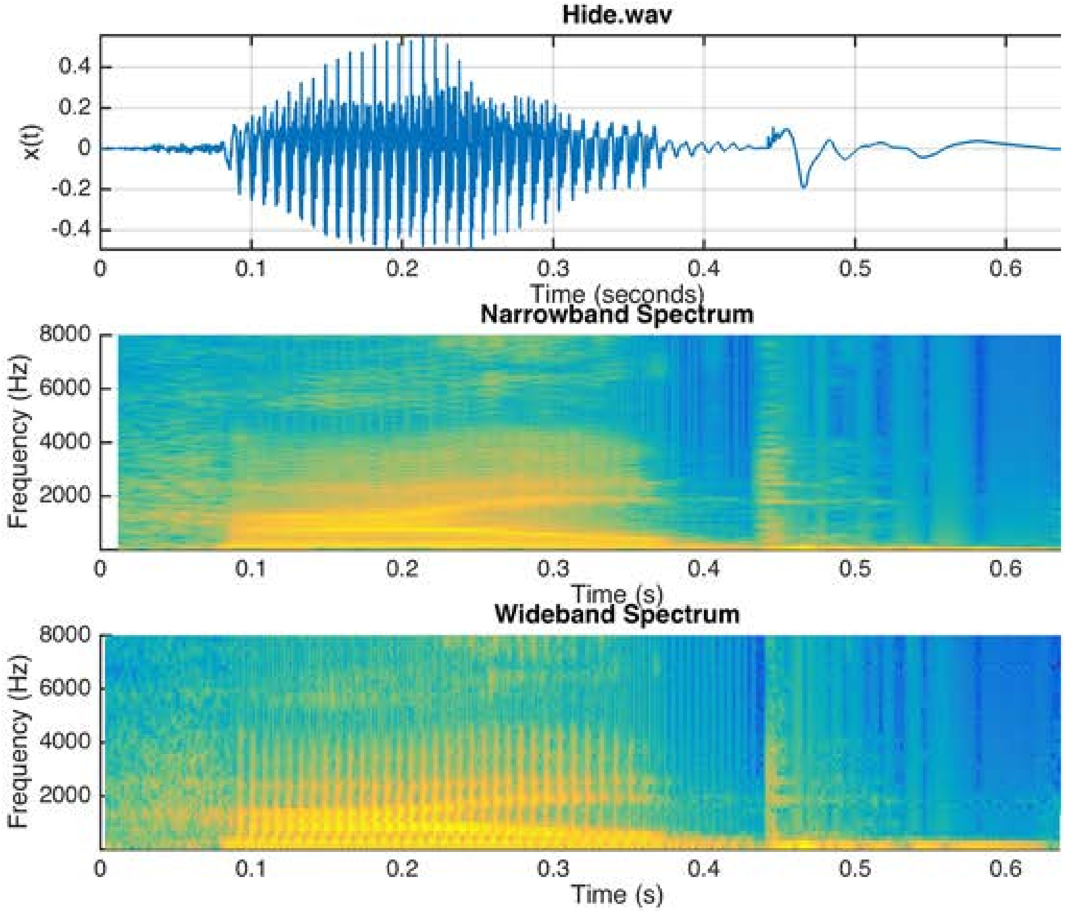

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

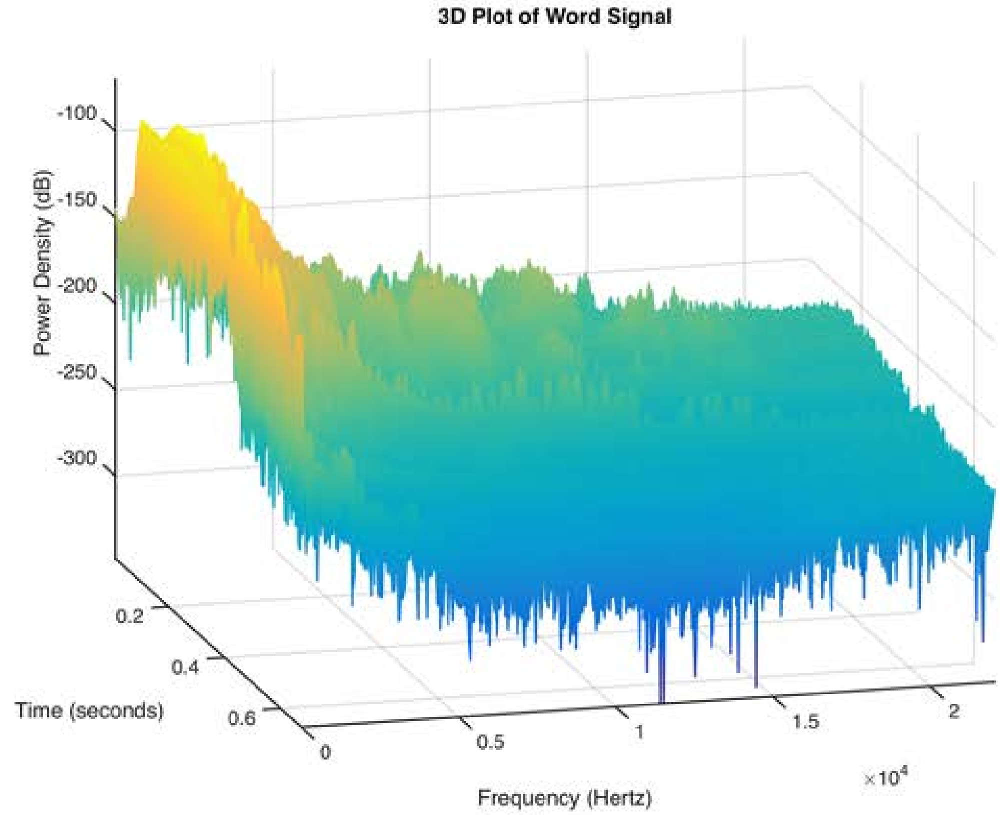
/h/

‘He’


h.wav
x(t)
Time (seconds)

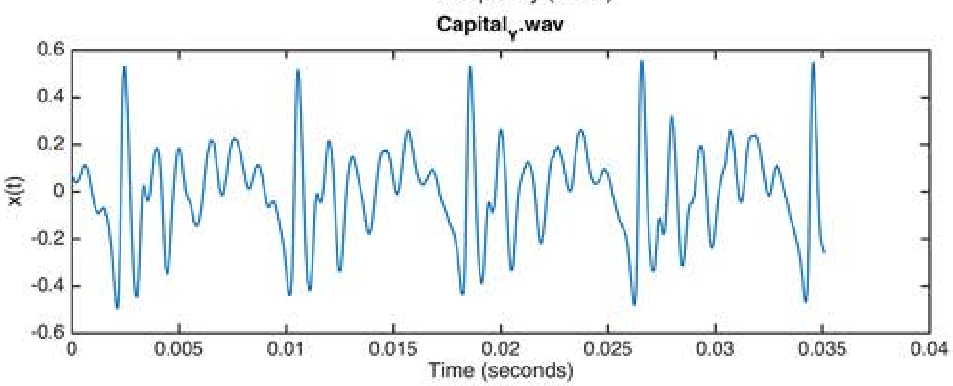

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

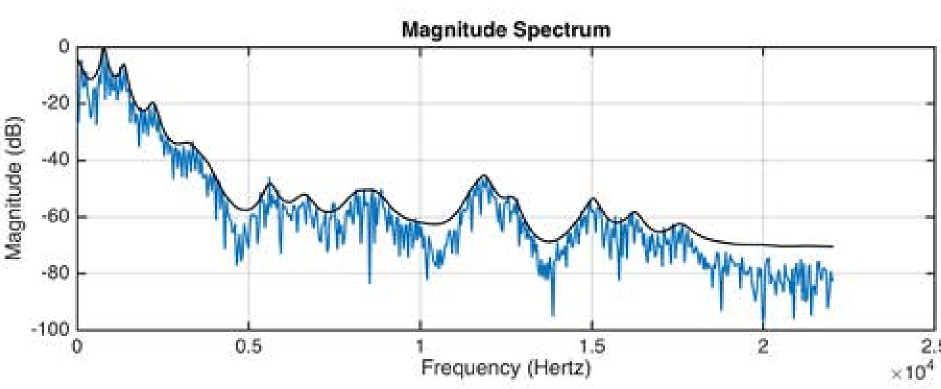

Hide.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

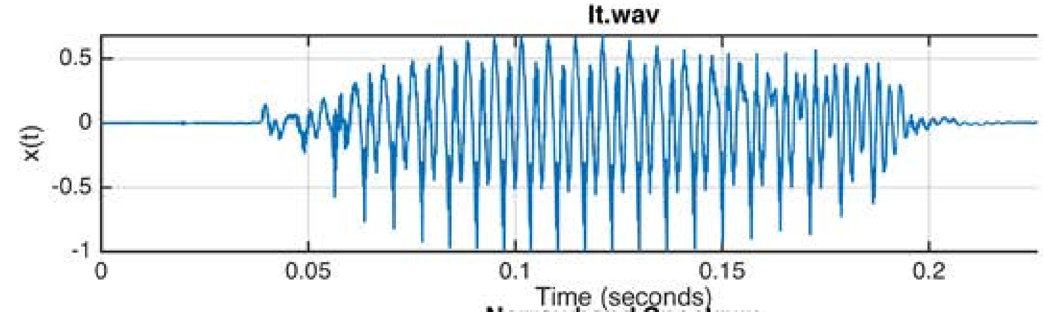

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

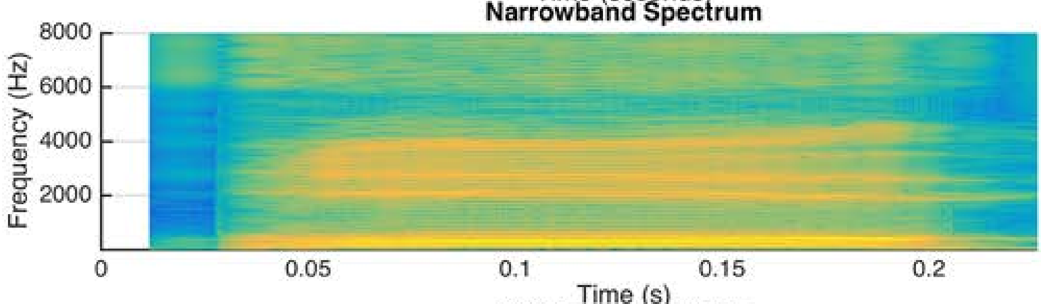
/Y/

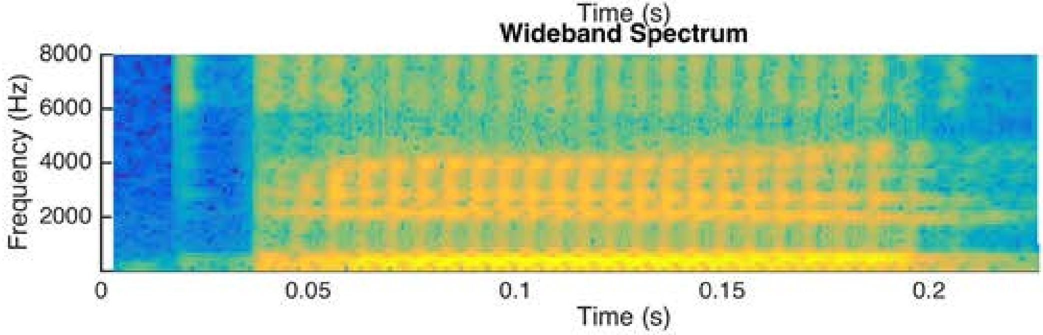
'Hide'

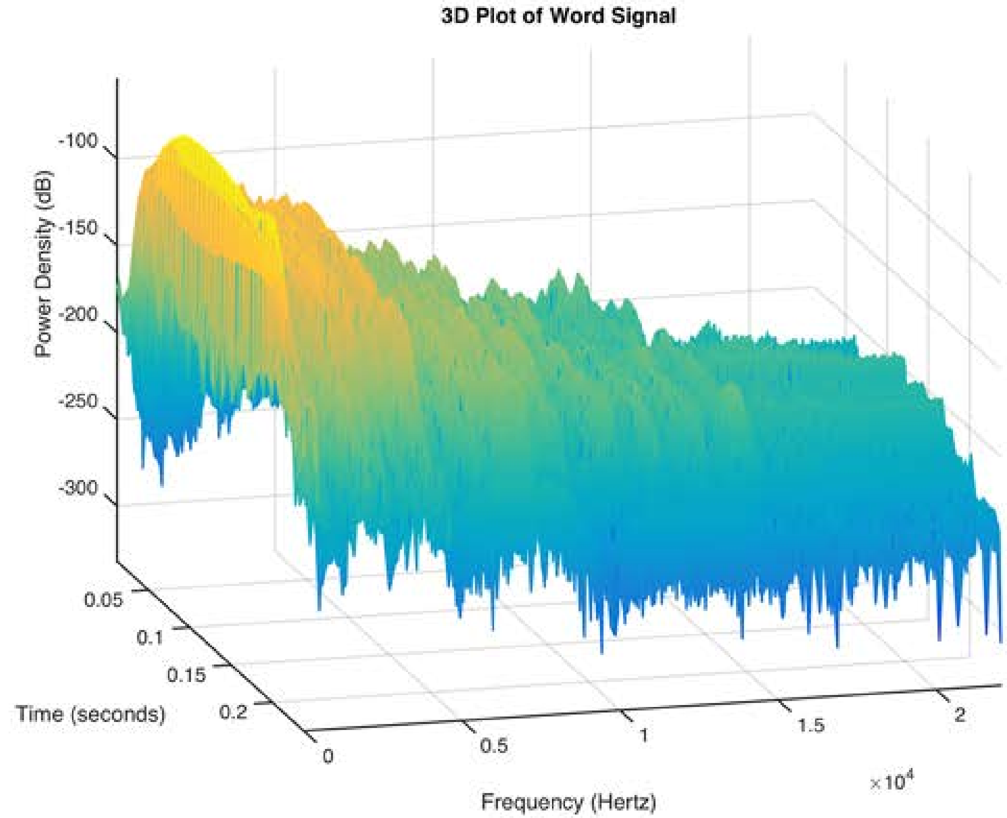

Capital_Y.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


It.wav
x(t)
Time (seconds)

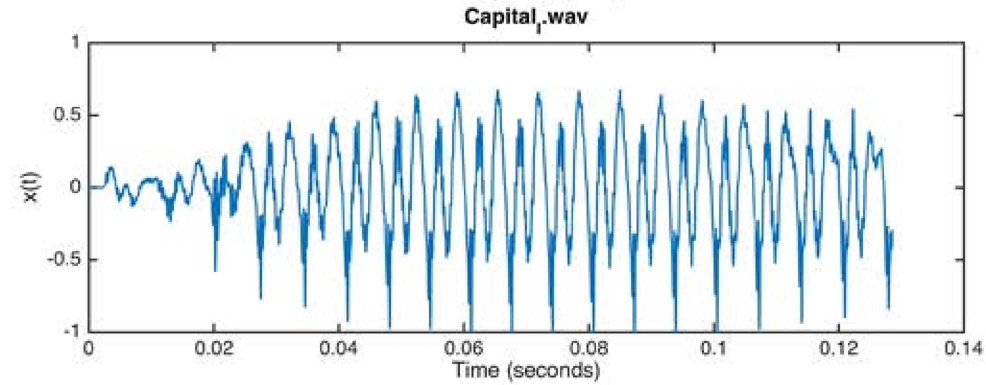

Narrowband Spectrum
Frequency (Hz)
Time (s)

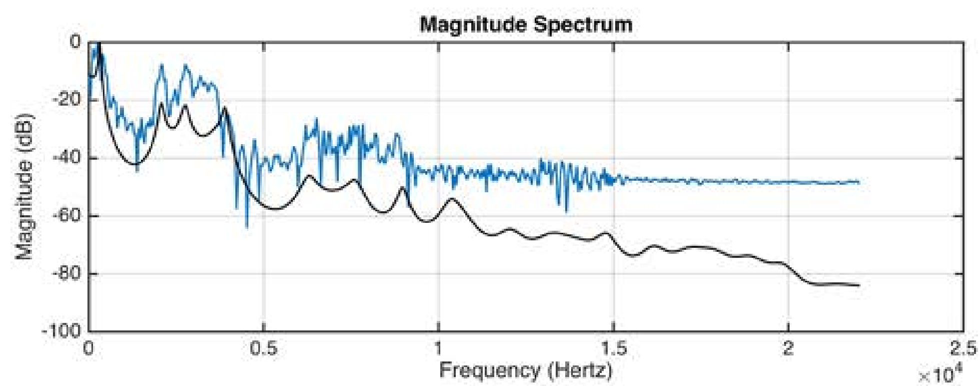

Wideband Spectrum
Frequency (Hz)
Time (s)

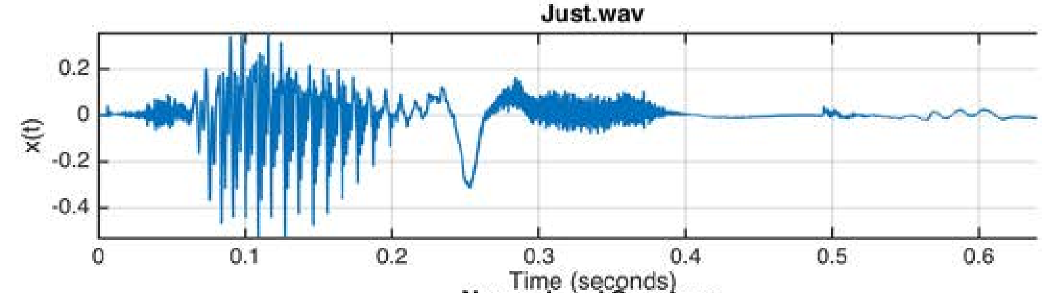

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

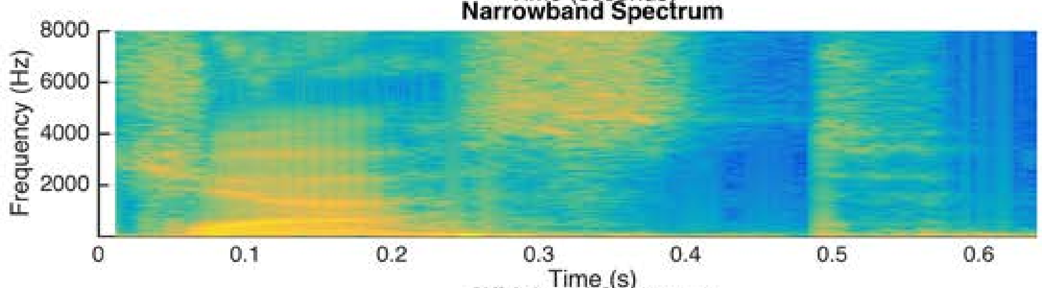
/I/

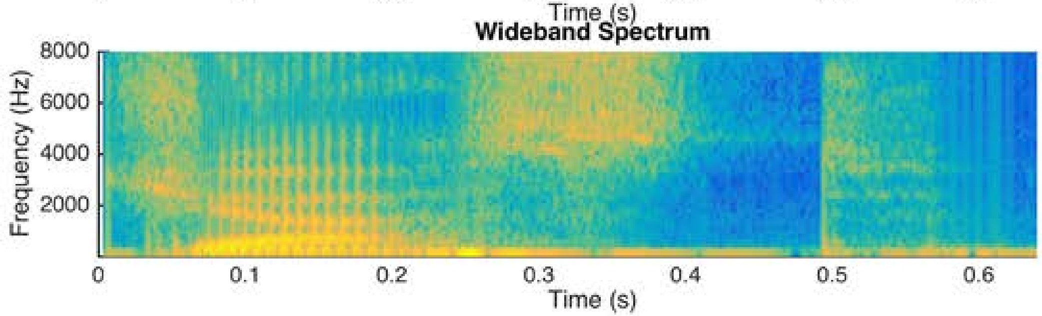
'It'

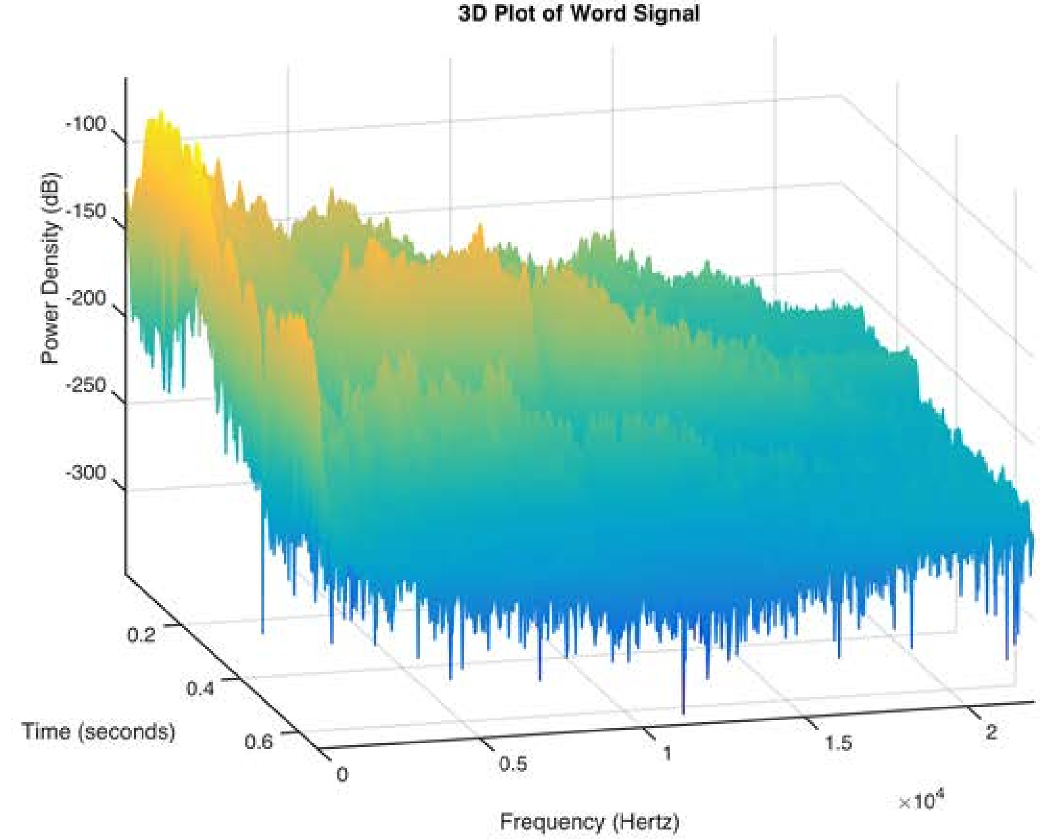

Capital_I.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


Just.wav
x(t)
Time (seconds)

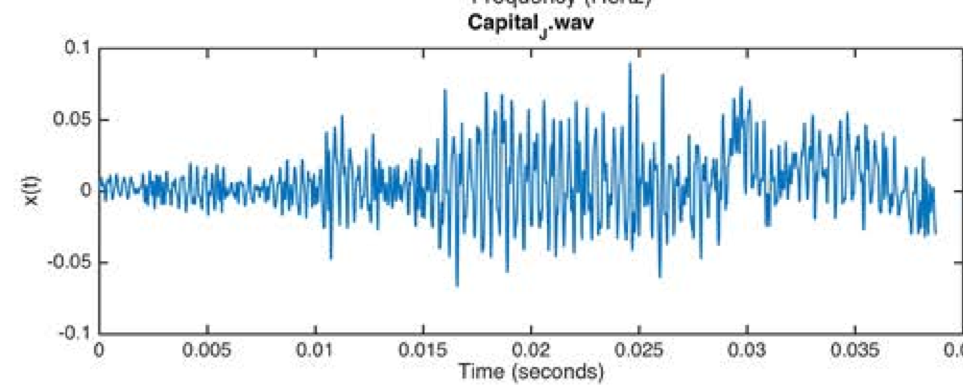

Narrowband Spectrum
Frequency (Hz)
Time (s)

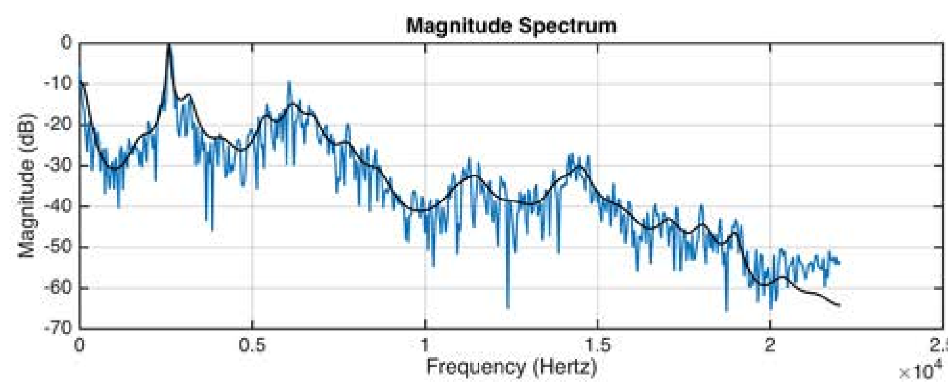

Wideband Spectrum
Frequency (Hz)
Time (s)

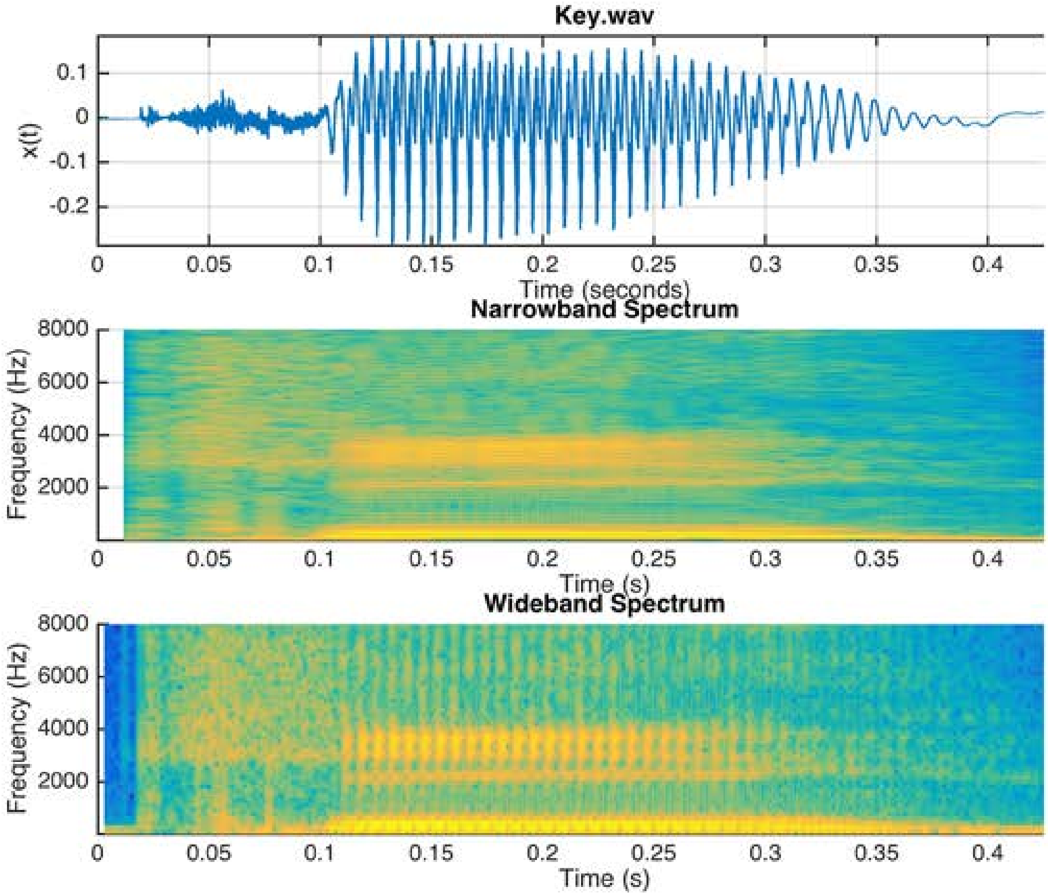

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

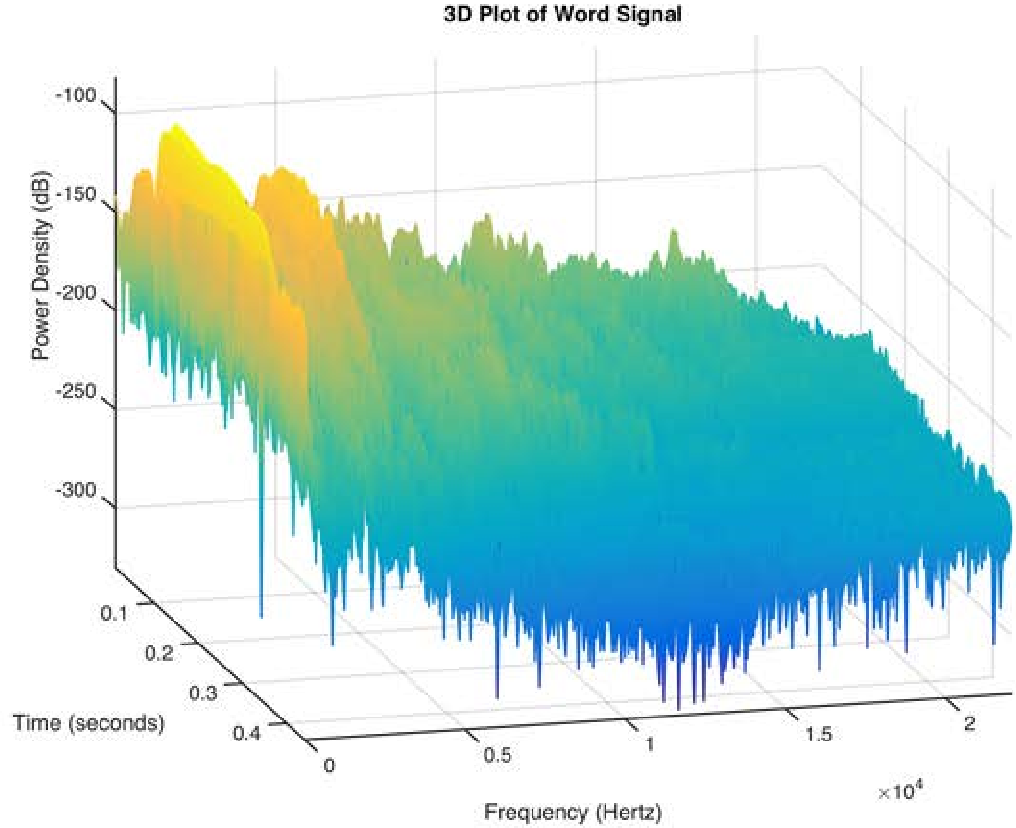
/J/

‘Just’


Capital_J.wav
x(t)
Time (seconds)

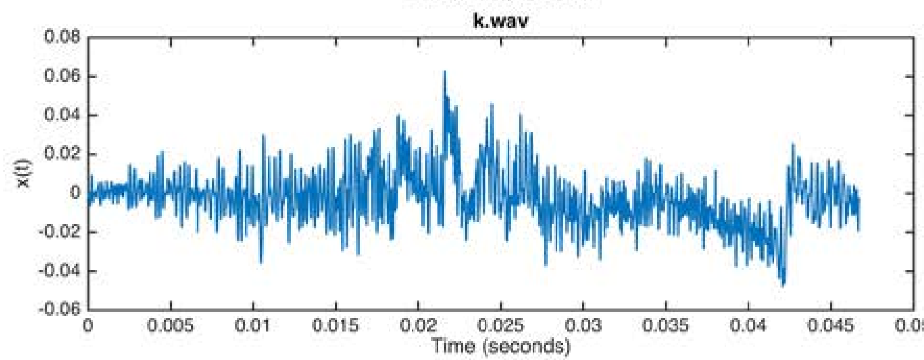

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

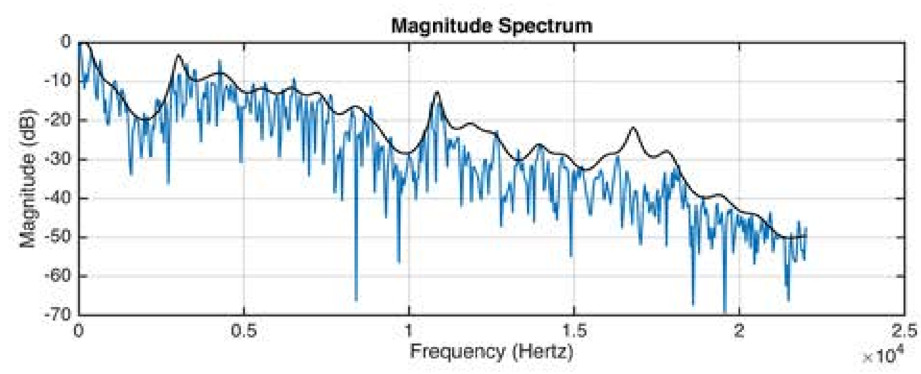

Key.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

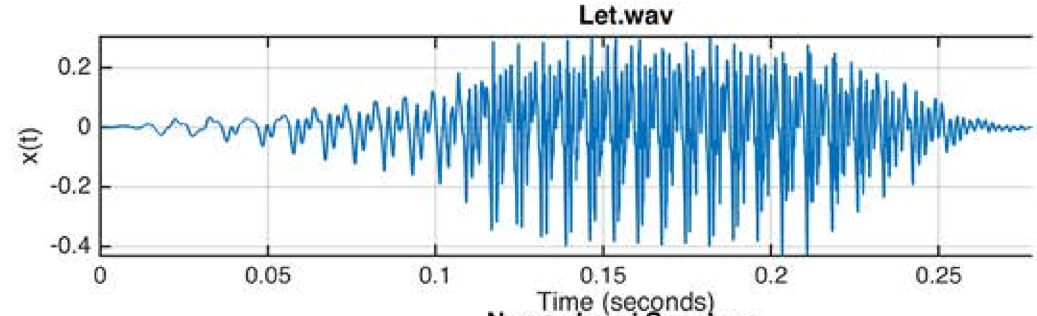

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

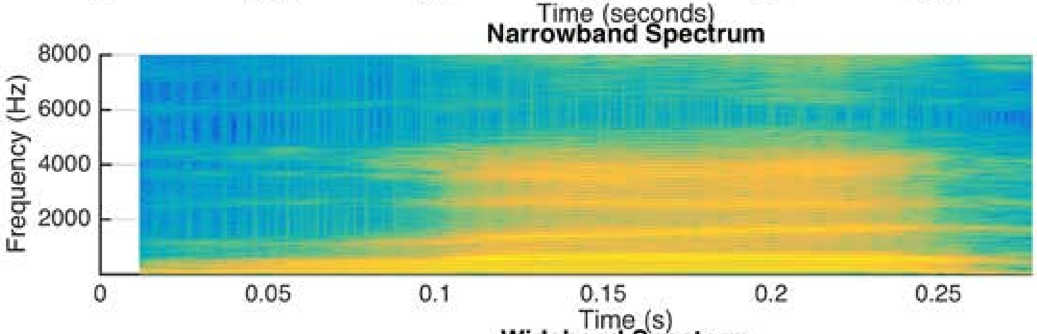
/k/

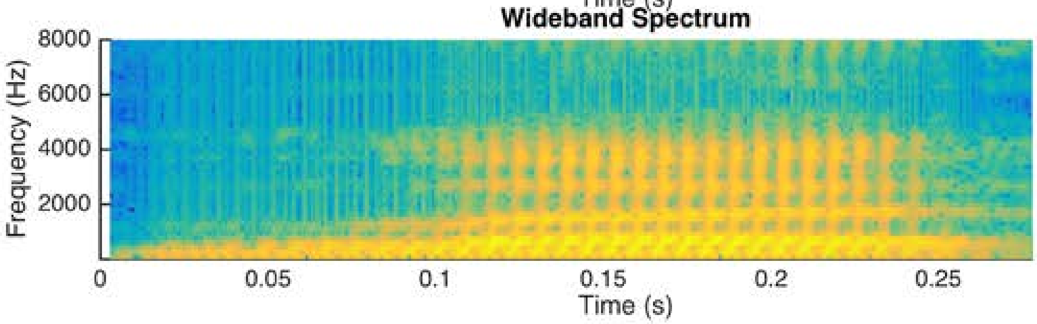
‘Key’

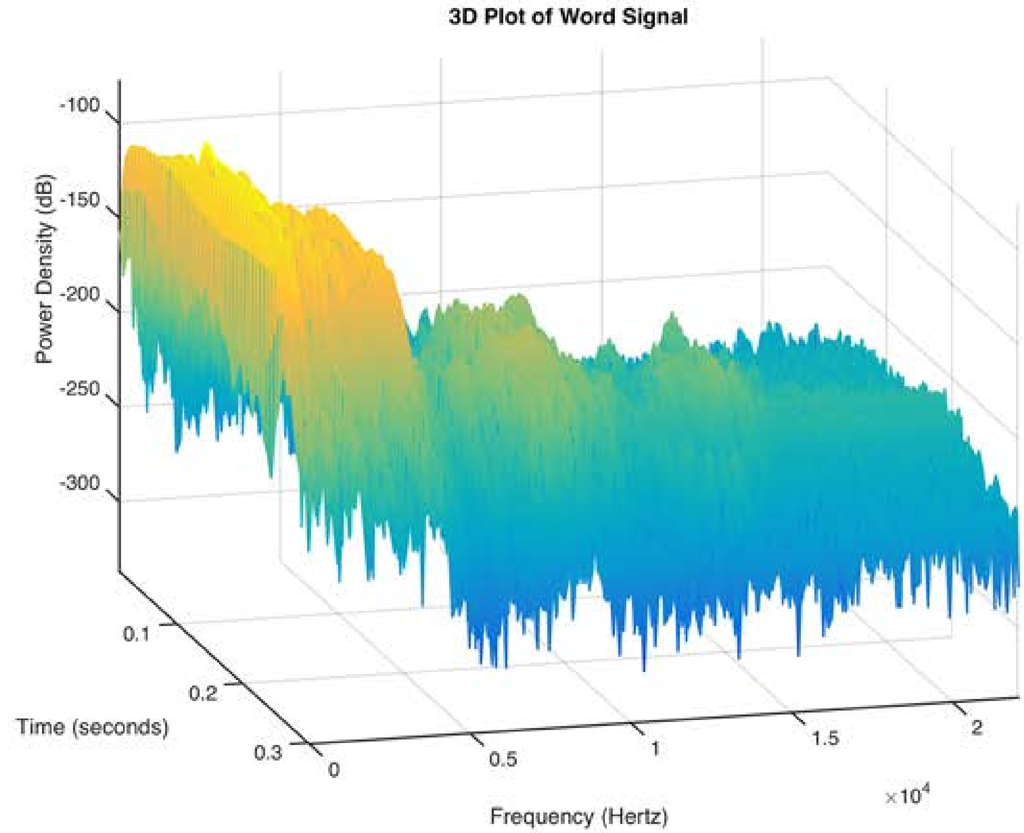

k.wav
x(t)
Time (seconds)

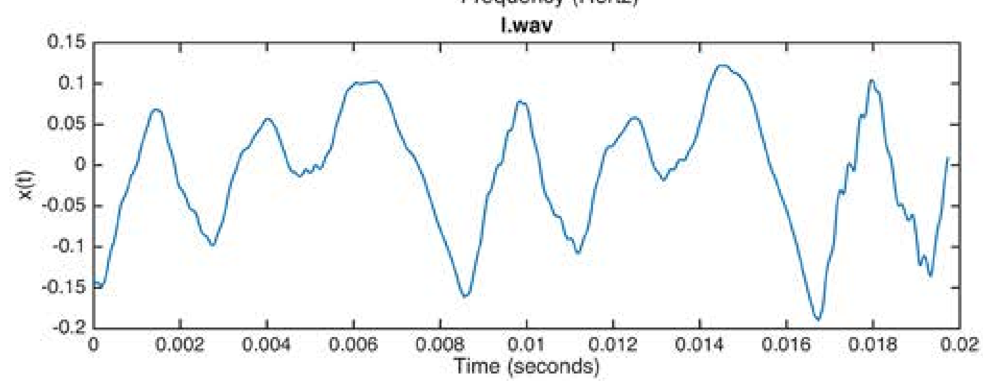

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

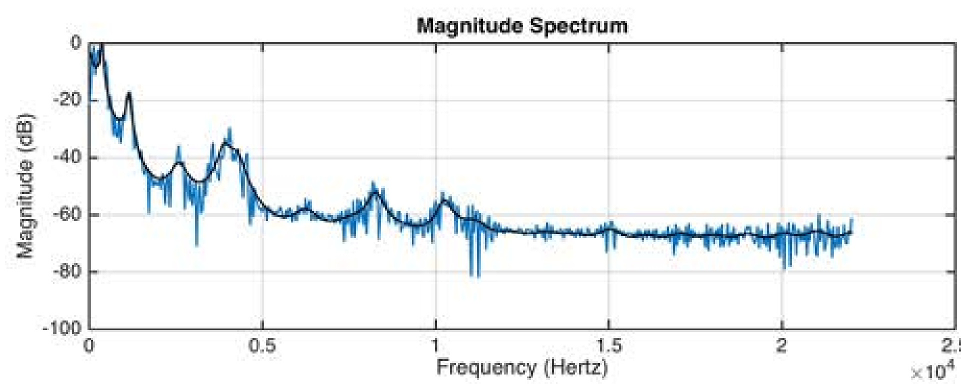

Let.wav
x(t)
Time (seconds)


Narrowband Spectrum
Frequency (Hz)
Time (s)


Wideband Spectrum
Frequency (Hz)
Time (s)

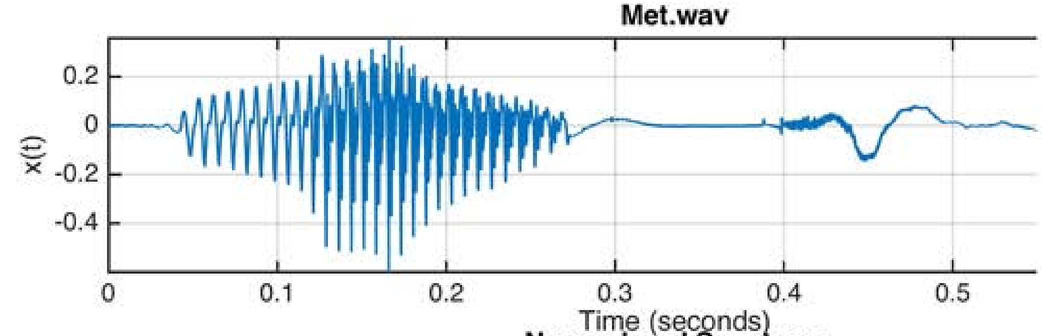

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
$\times 10^4$

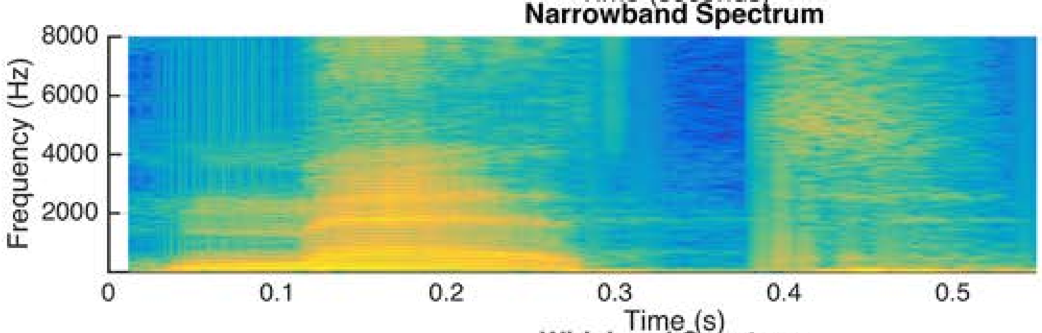

l.wav
x(t)
Time (seconds)

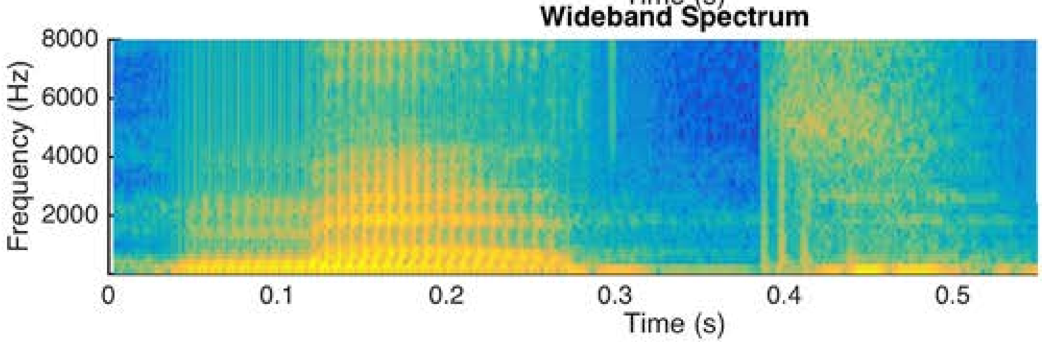

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
$\times 10^4$

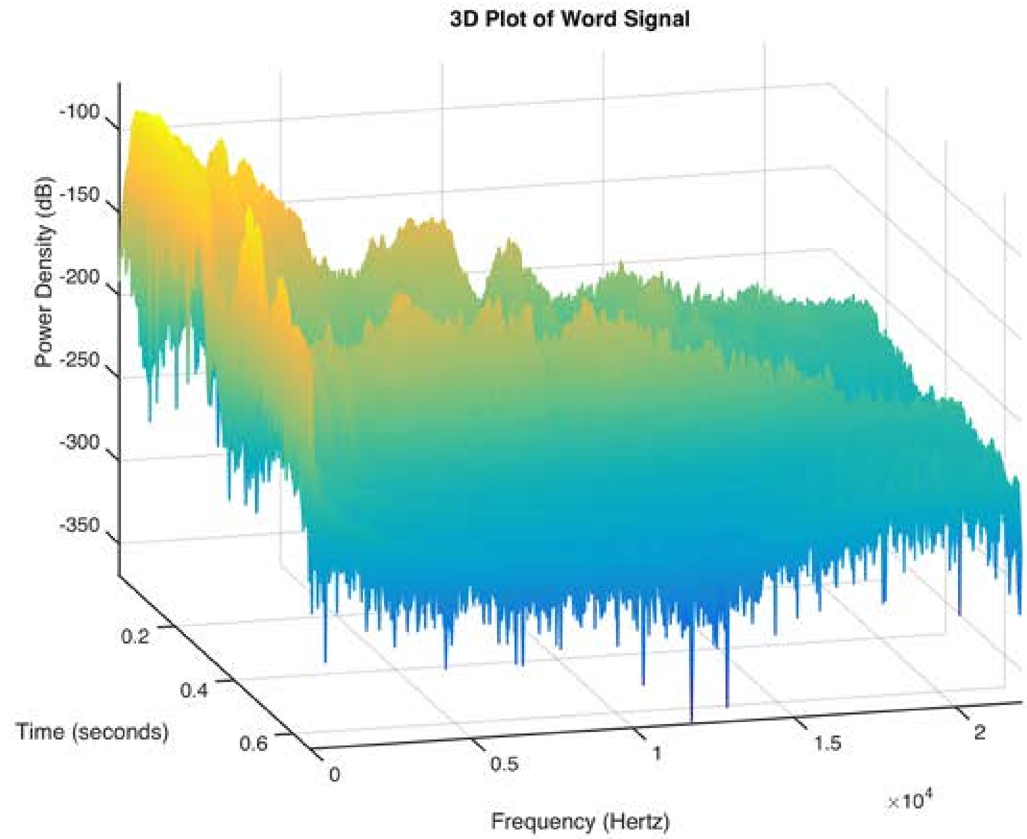
/l/

'Let'


Met.wav
x(t)
Time (seconds)

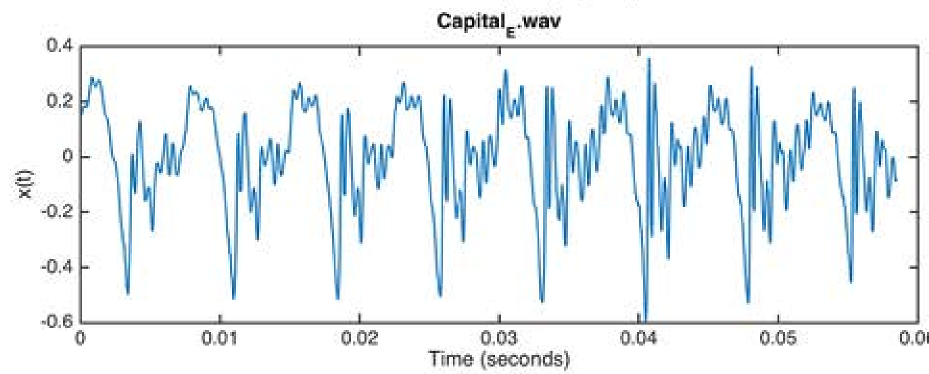

Narrowband Spectrum
Frequency (Hz)
Time (s)

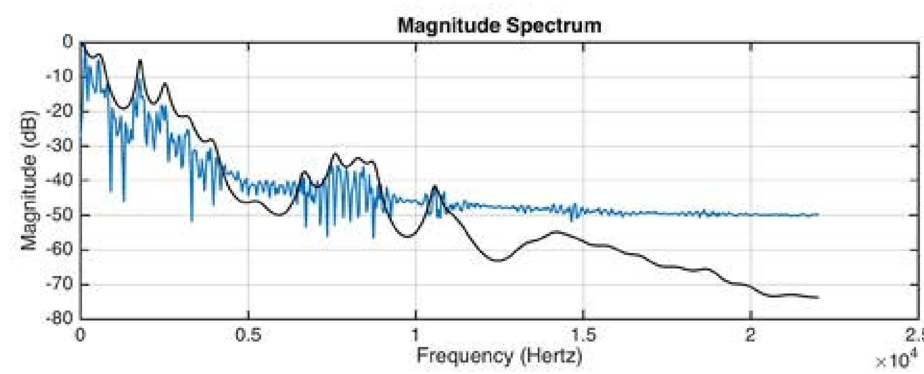

Wideband Spectrum
Frequency (Hz)
Time (s)

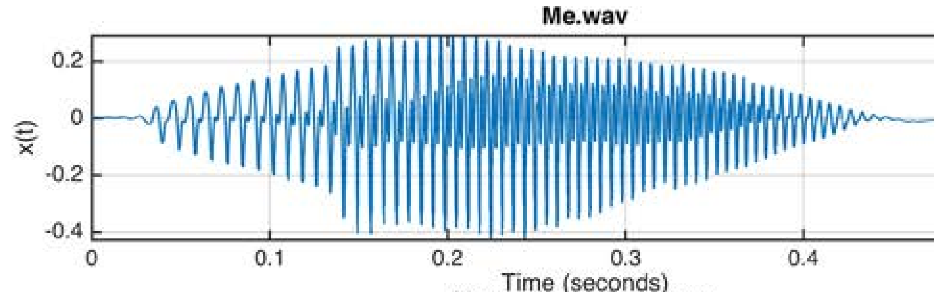

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

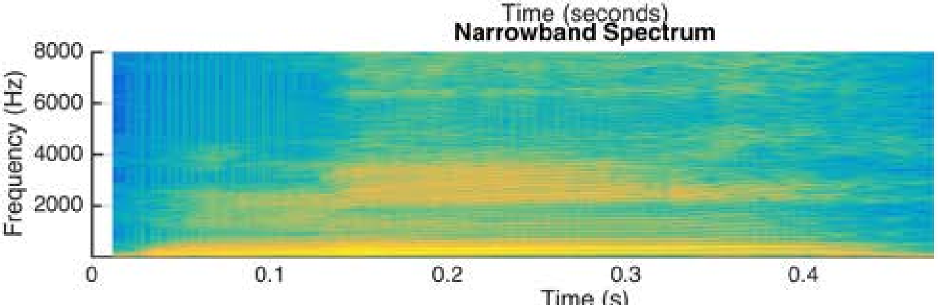
/E/

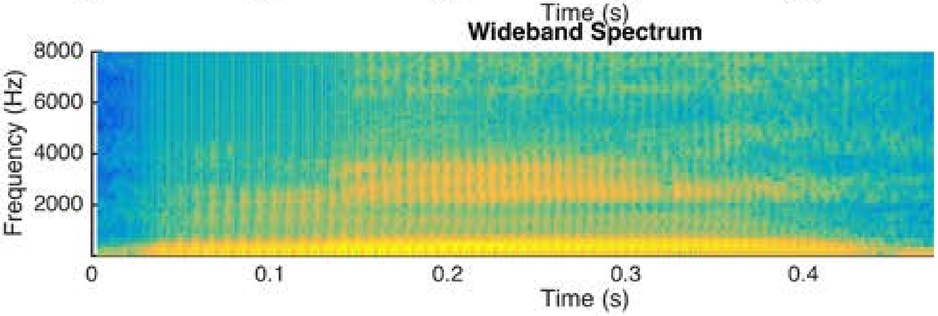
'Met'

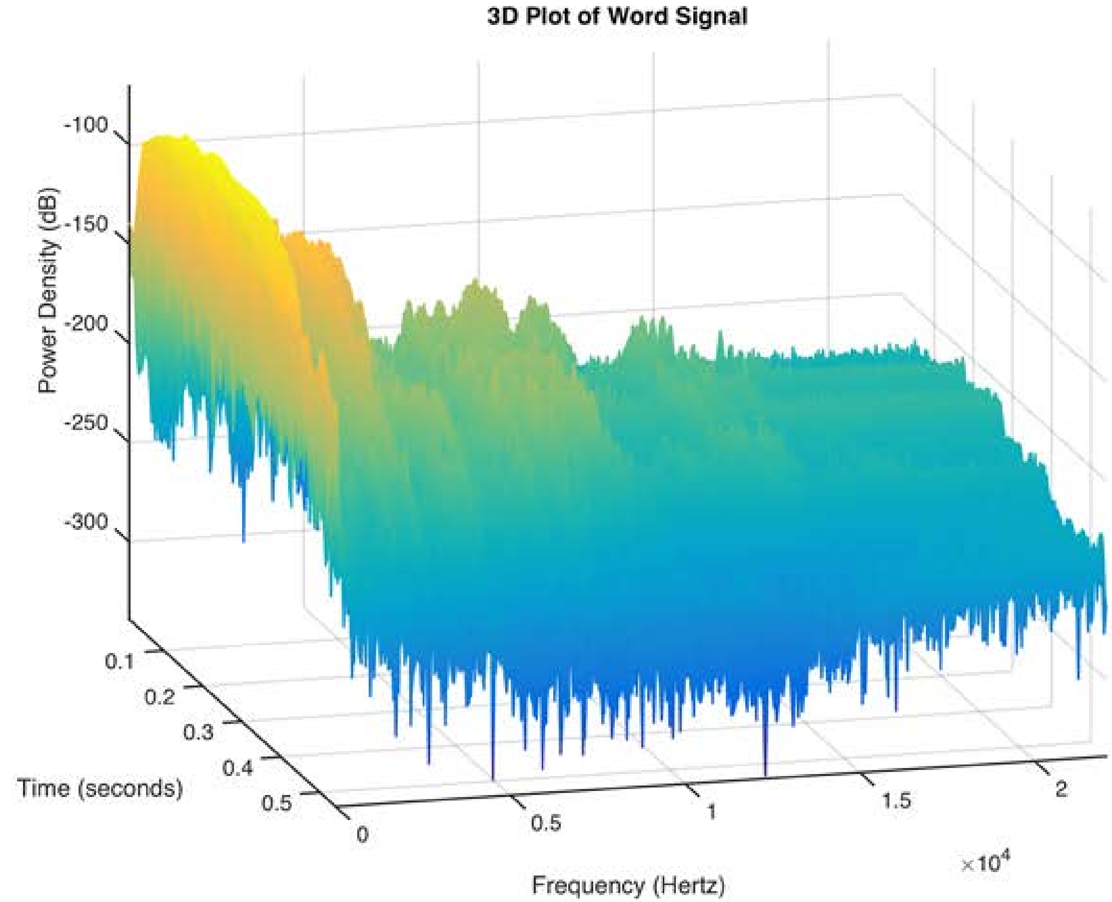

Capital_E.wav
x(t)
Time (seconds)

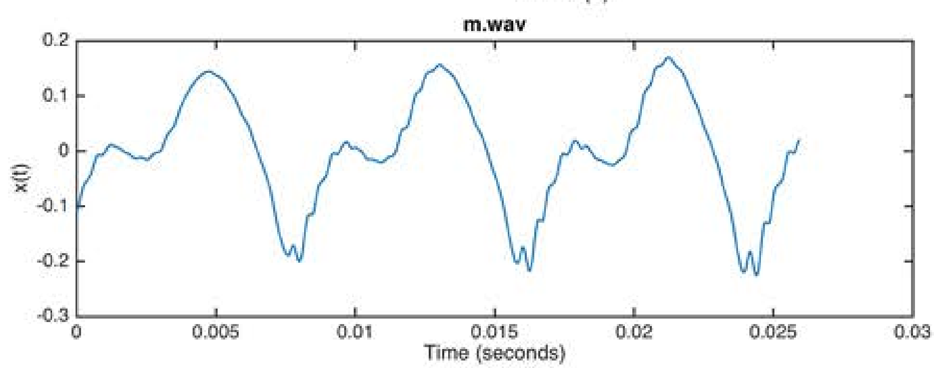

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

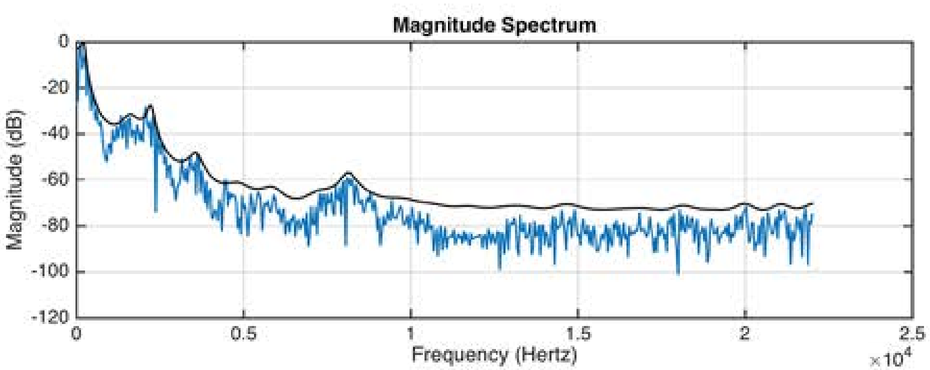

Me.wav
x(t)
Time (seconds)


Narrowband Spectrum
Frequency (Hz)
Time (s)


Wideband Spectrum
Frequency (Hz)
Time (s)

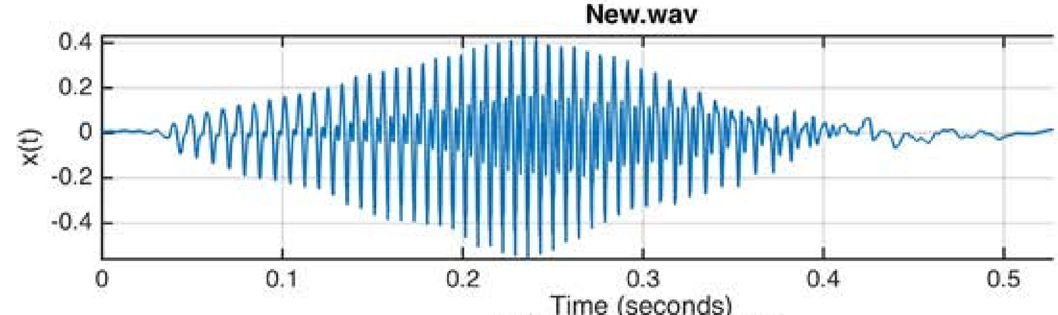

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

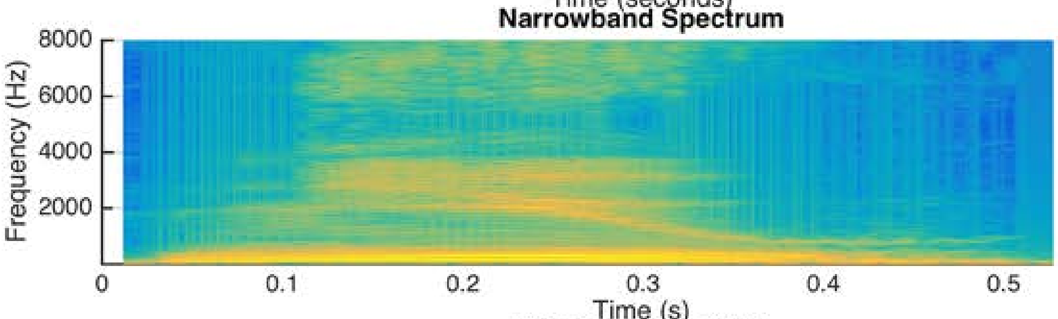

m.wav
x(t)
Time (seconds)

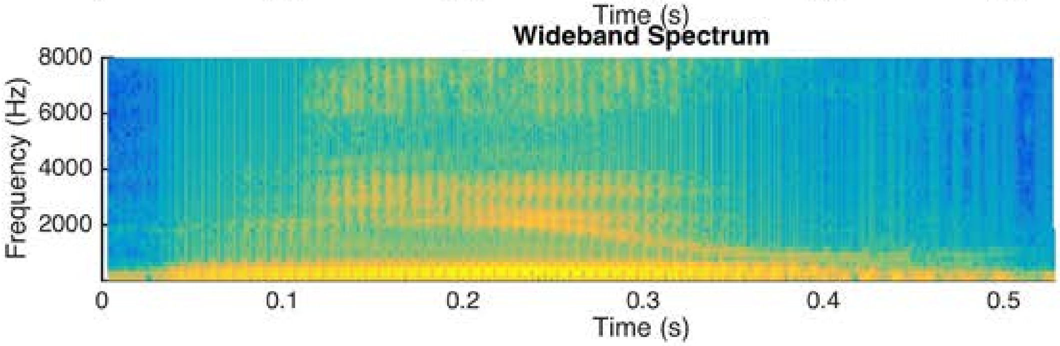

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

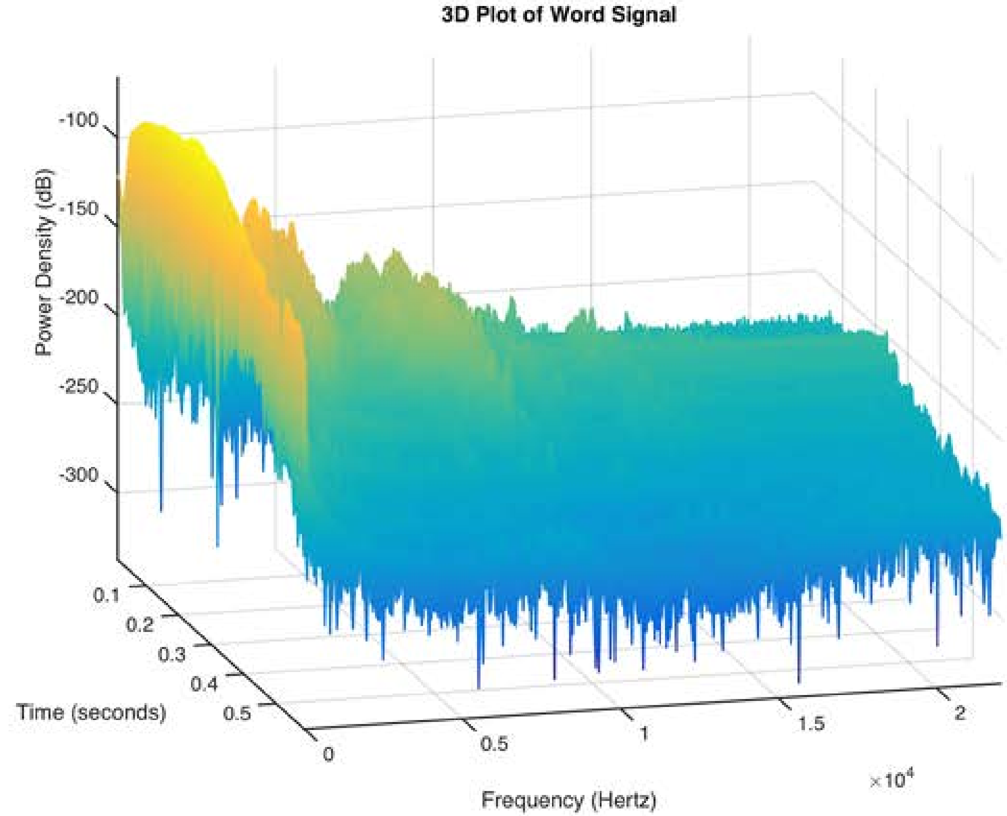
/m/

‘Me’


New.wav
x(t)
Time (seconds)

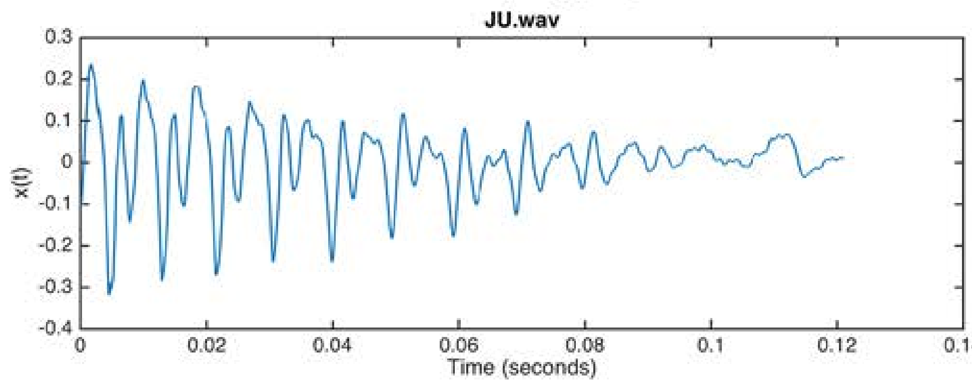

Narrowband Spectrum
Frequency (Hz)
Time (s)

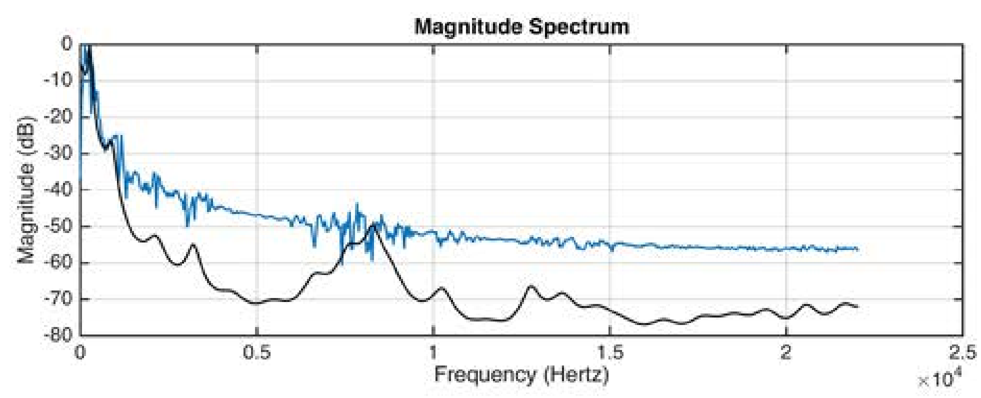

Wideband Spectrum
Frequency (Hz)
Time (s)

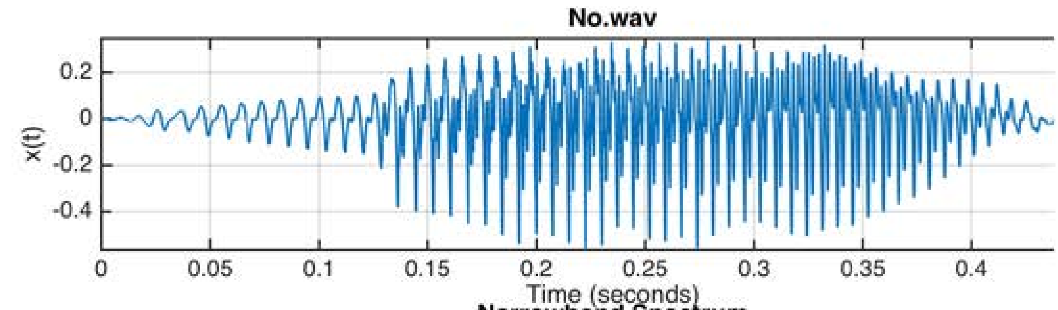

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

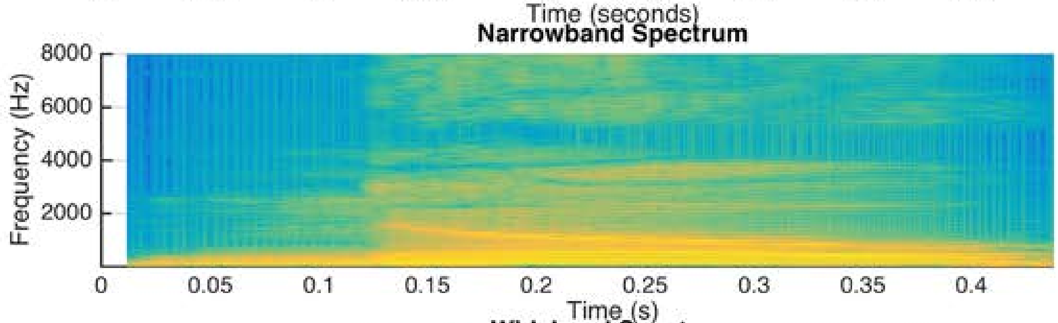
/JU/

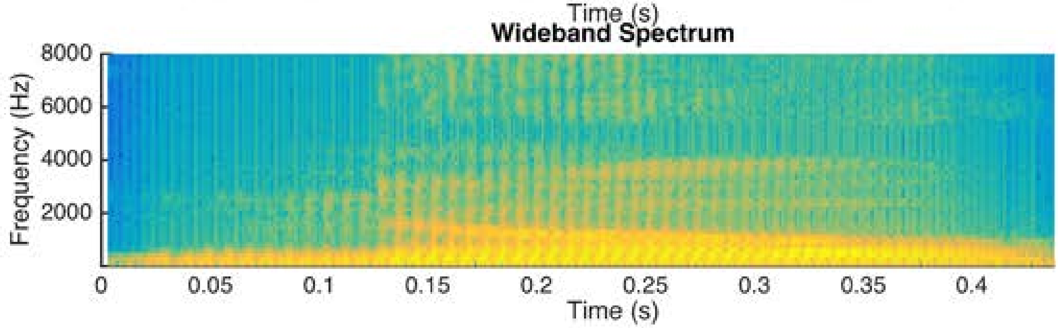
‘New’


JU.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

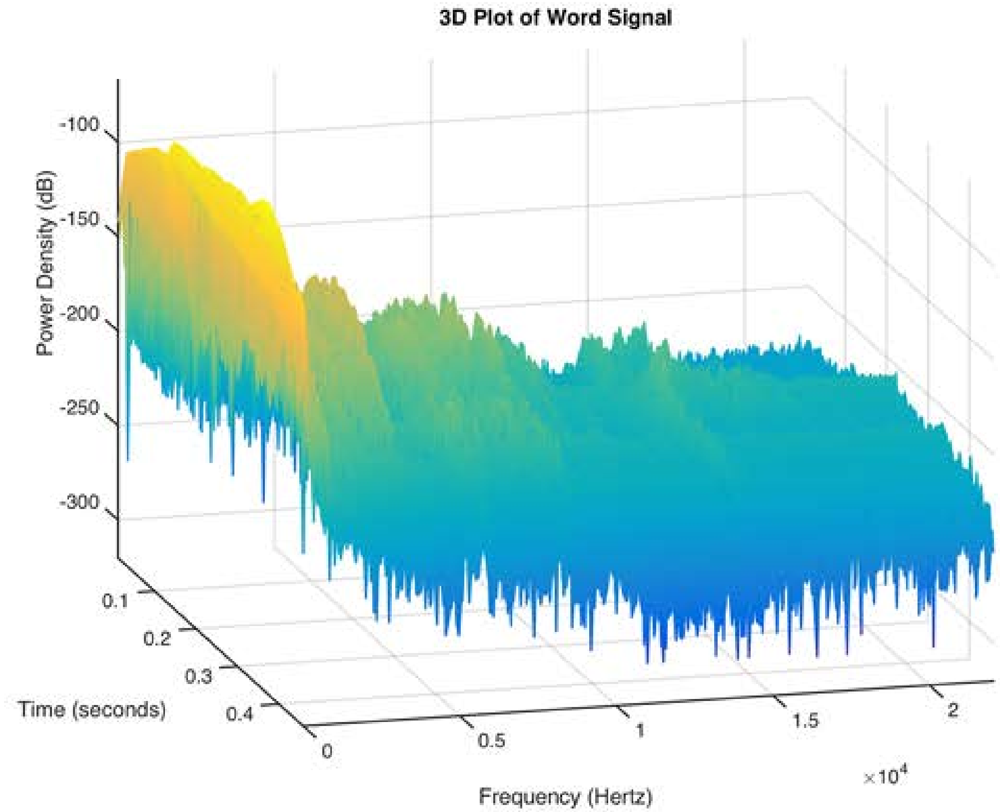

No.wav
x(t)
Time (seconds)

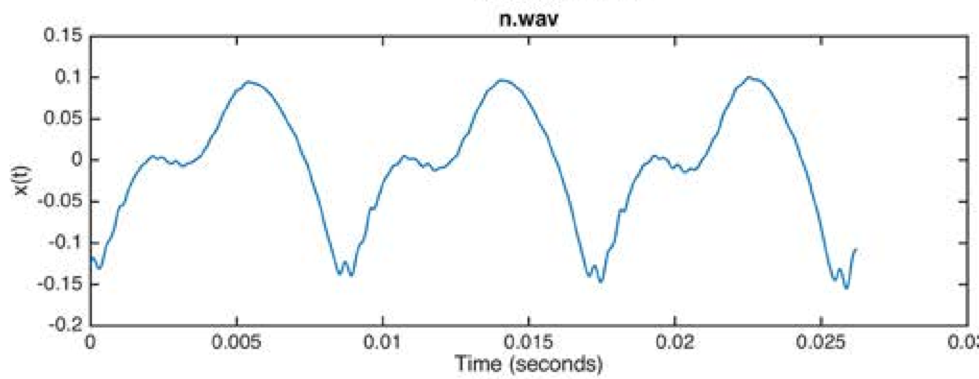

Narrowband Spectrum
Frequency (Hz)
Time (s)

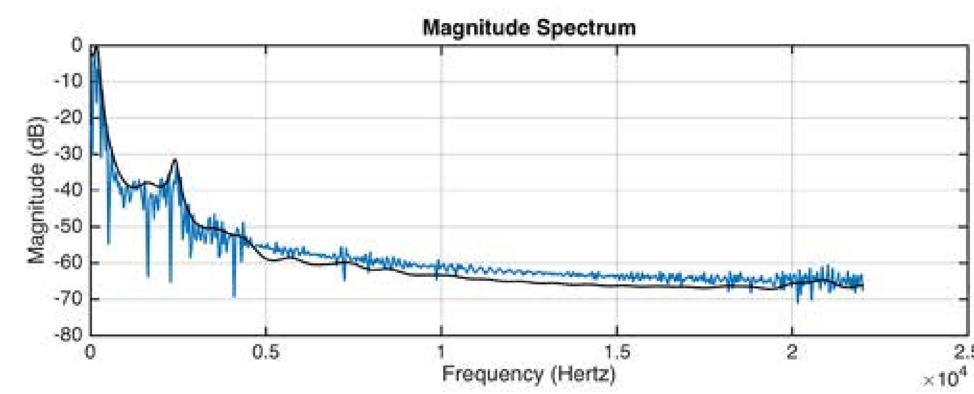

Wideband Spectrum
Frequency (Hz)
Time (s)

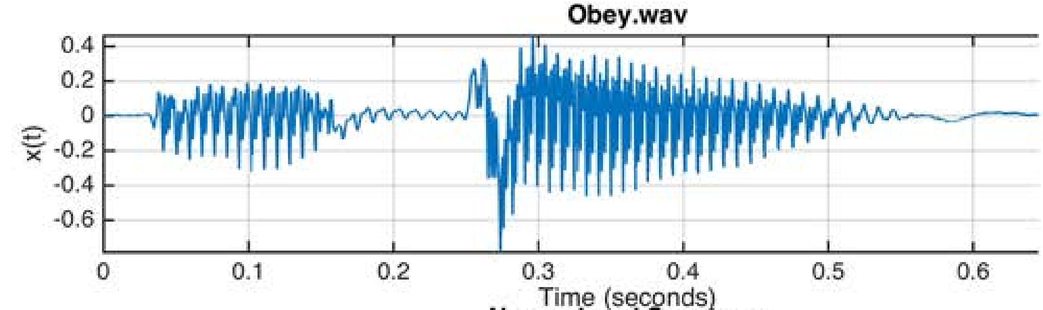
/n/

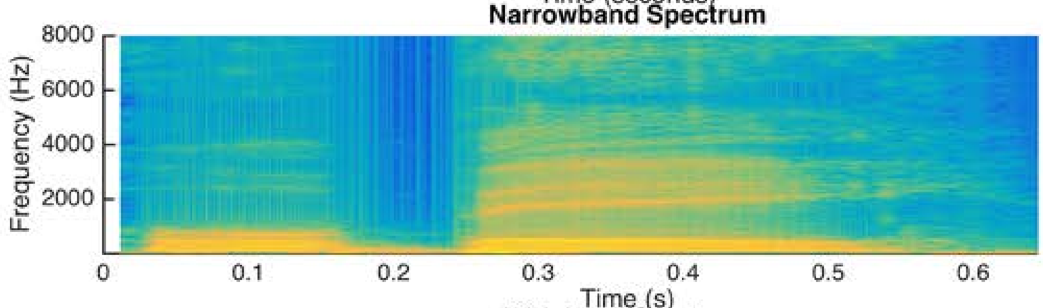
‘No’

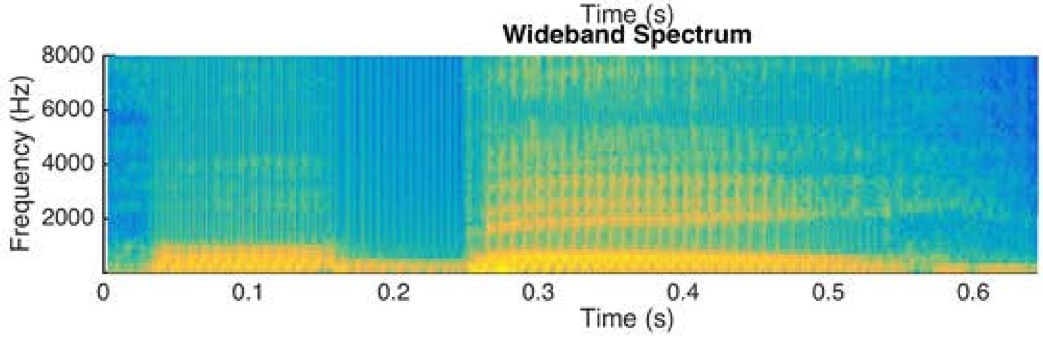

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

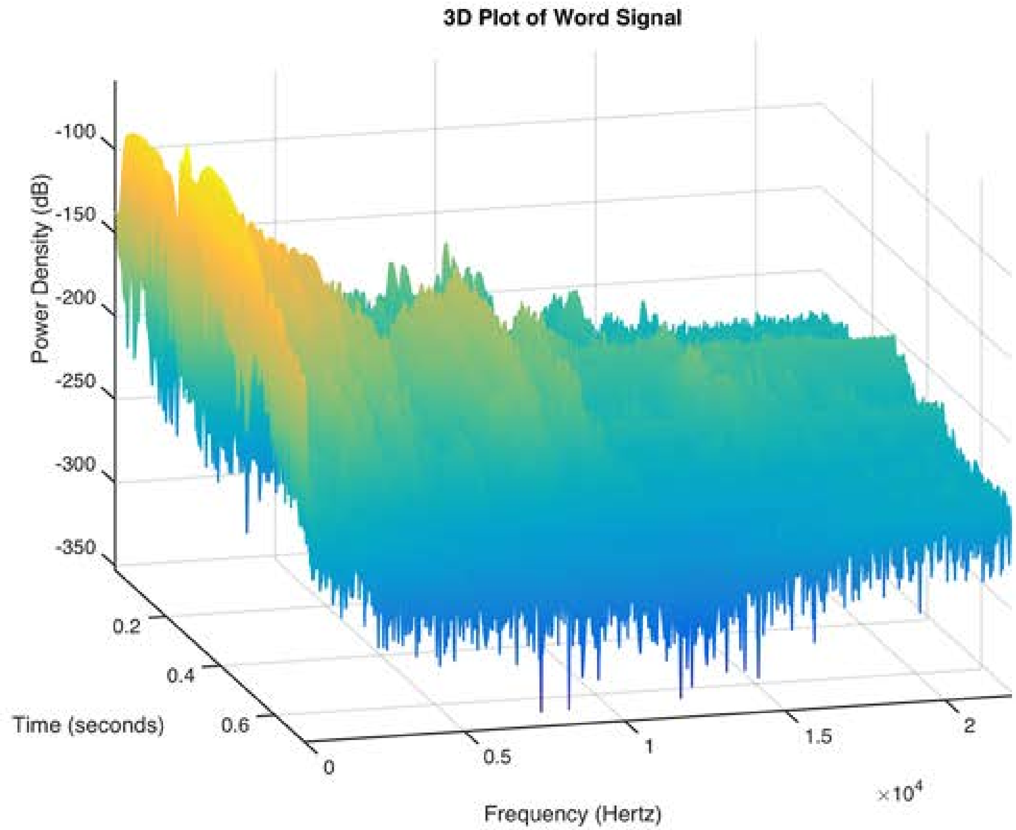

n.wav
x(t)
Time (seconds)

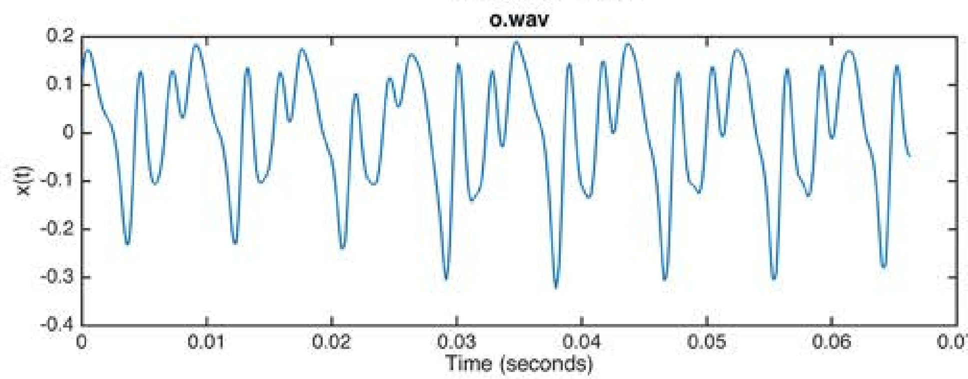

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

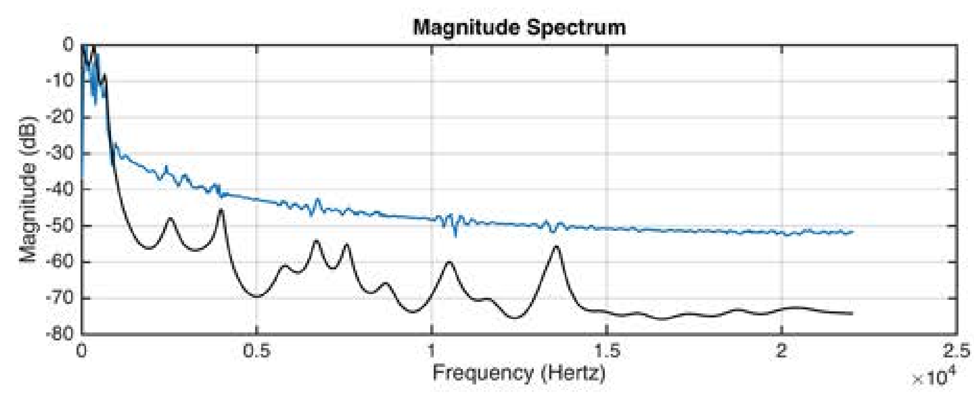

Obey.wav
x(t)
Time (seconds)


Narrowband Spectrum
Frequency (Hz)
Time (s)


Wideband Spectrum
Frequency (Hz)
Time (s)

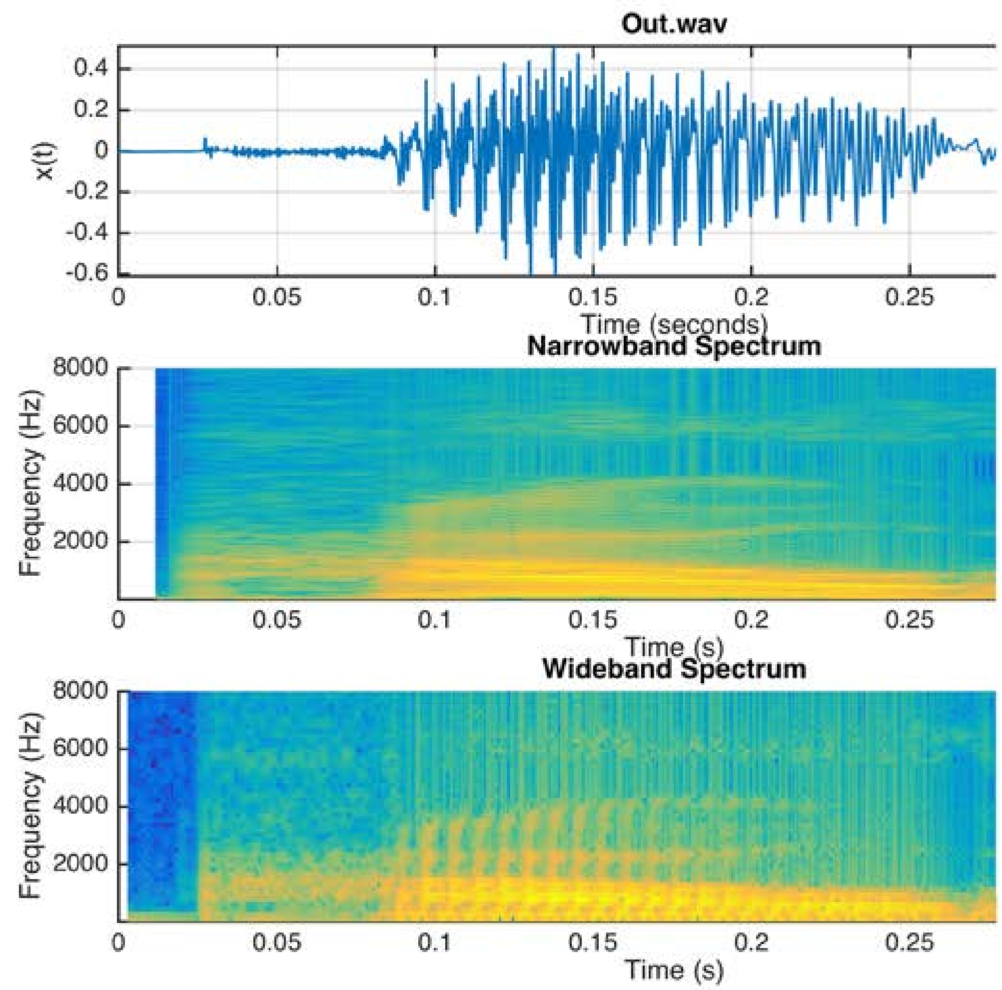

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)


o.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

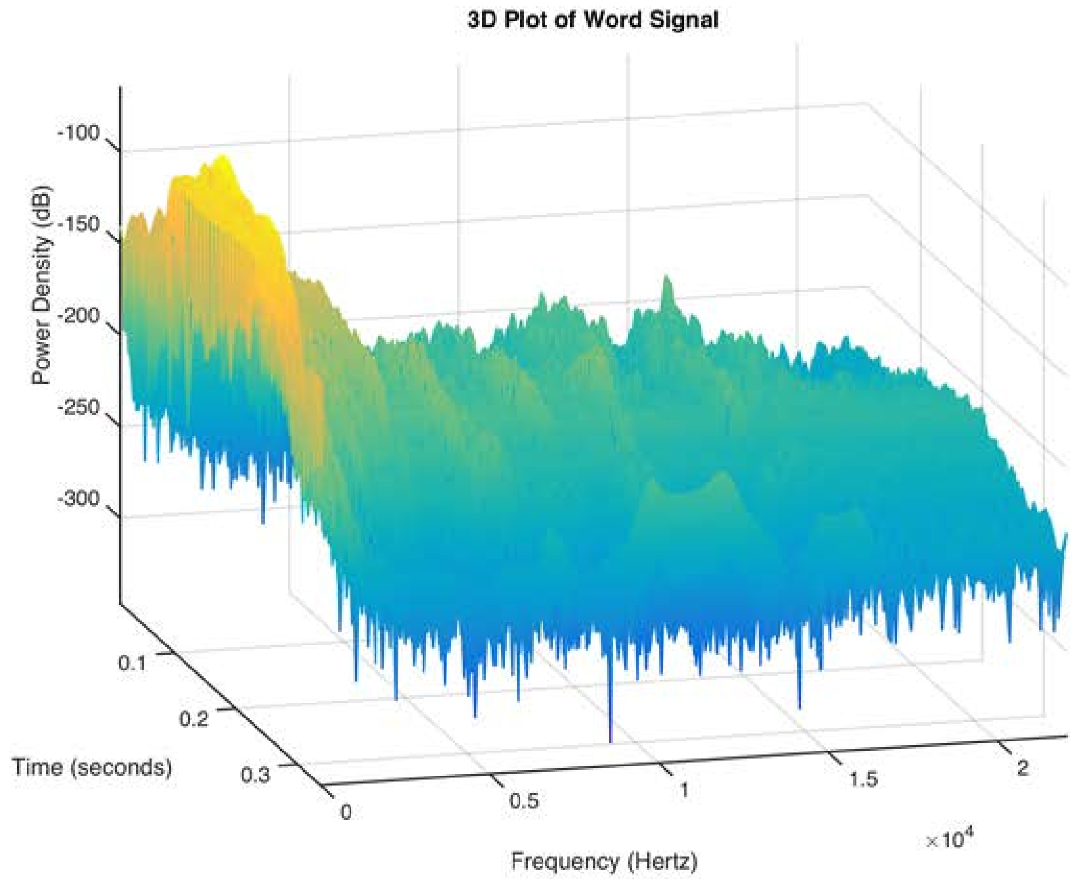
/o/

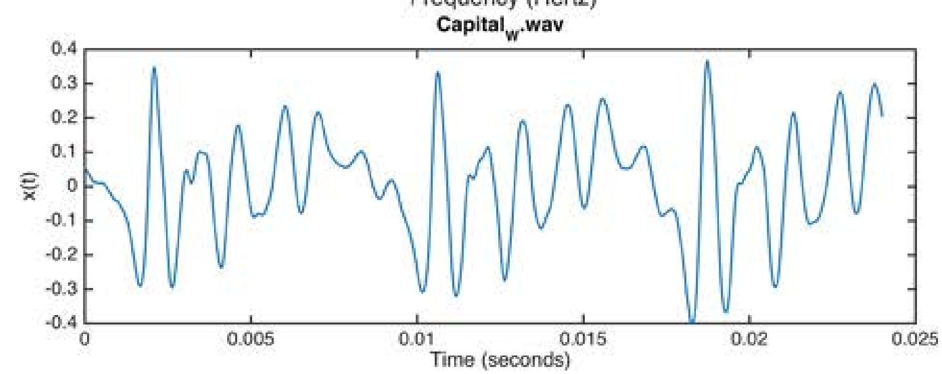
‘Obey’

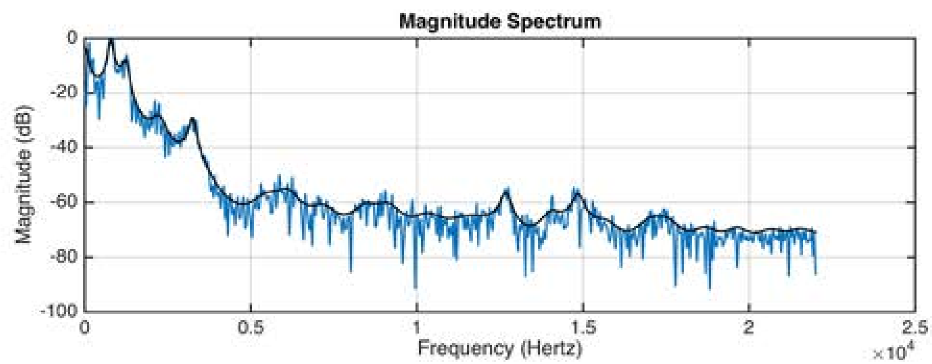

Out.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

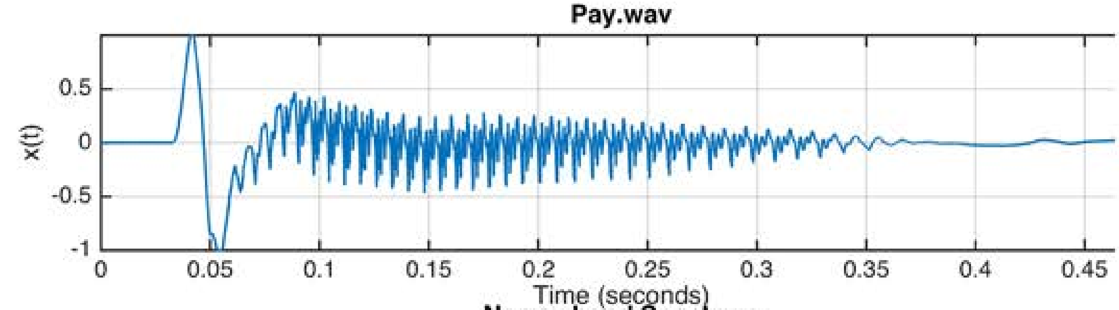
/w/

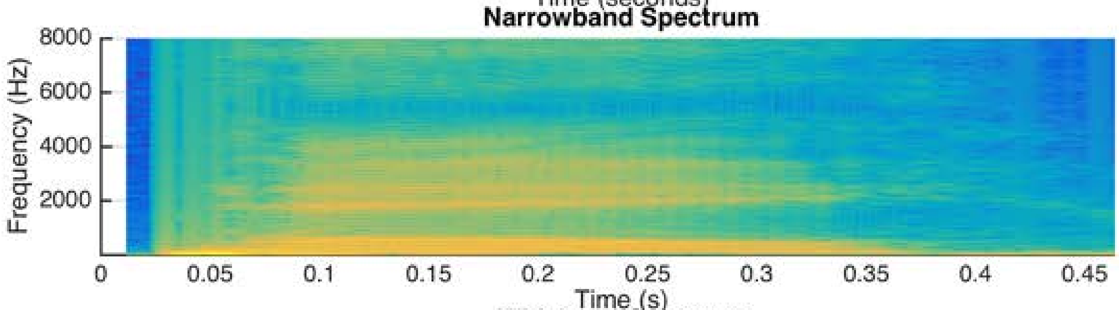
'Out'

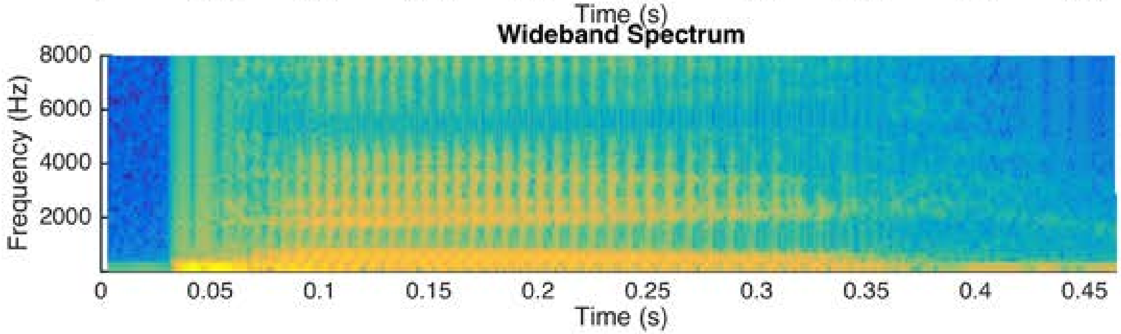

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

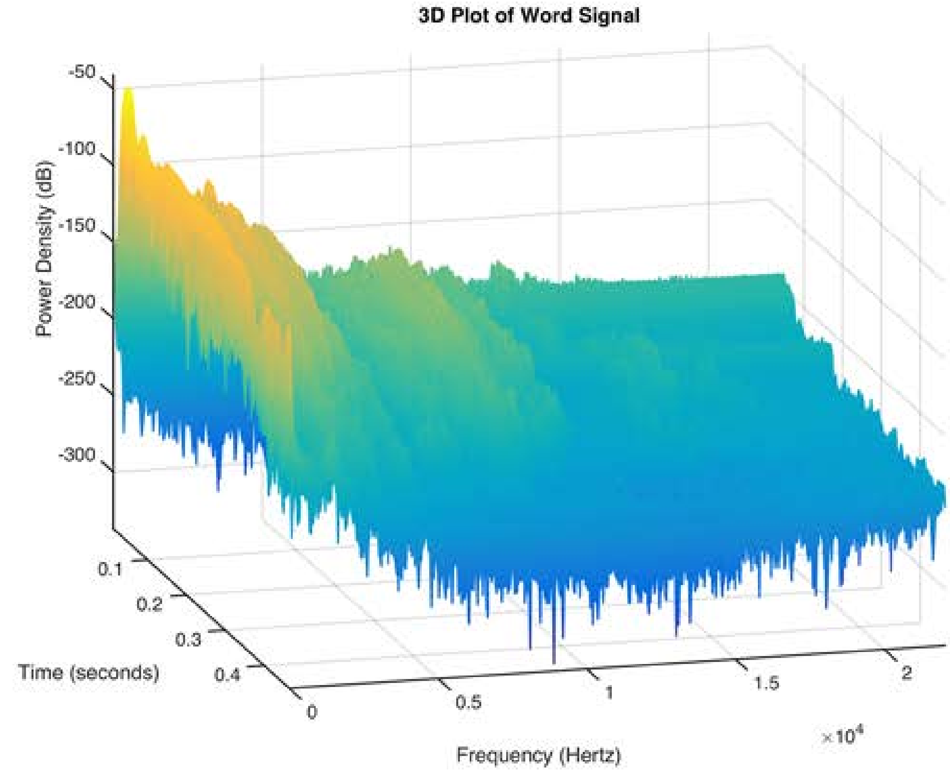

Capital_w.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


Pay.wav
x(t)
Time (seconds)

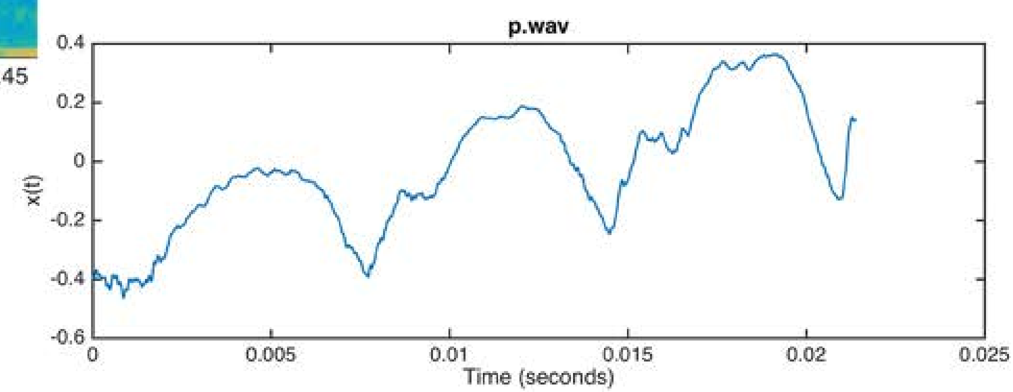

Narrowband Spectrum
Frequency (Hz)
Time (s)

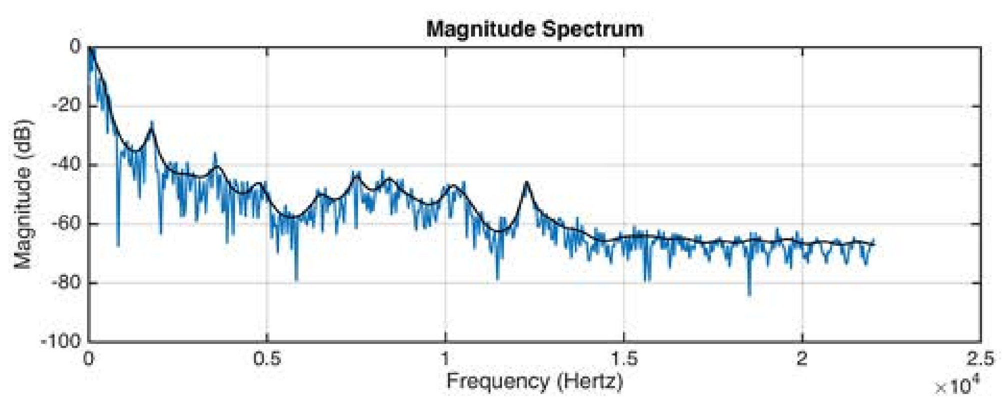

Wideband Spectrum
Frequency (Hz)
Time (s)

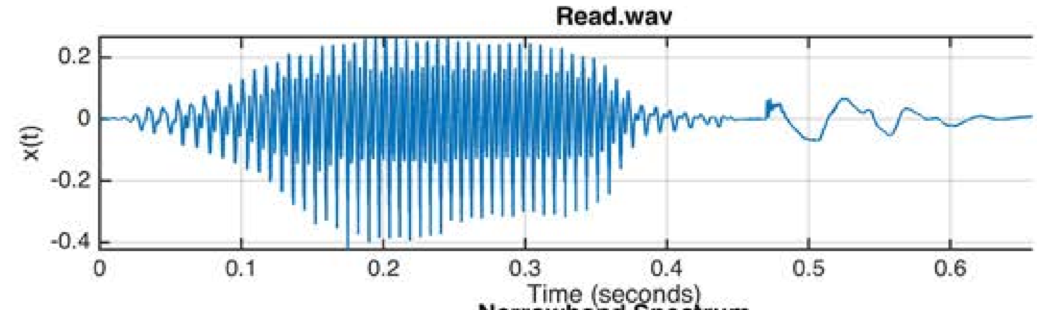

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

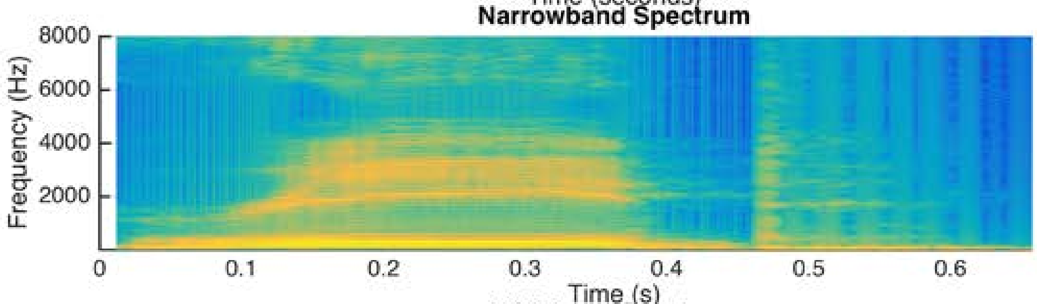
/p/

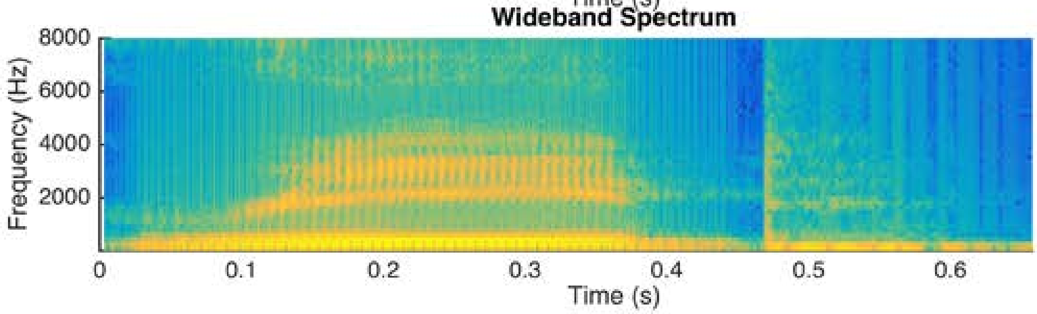
‘Pay’

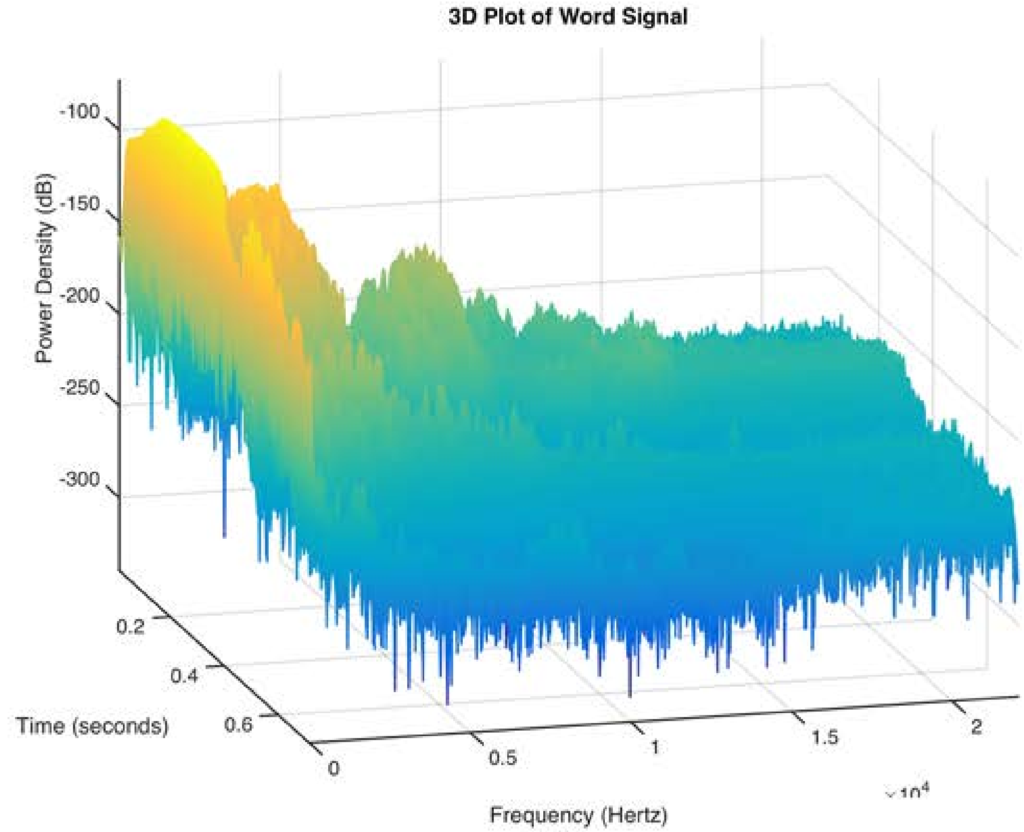

p.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


Read.wav
x(t)
Time (seconds)

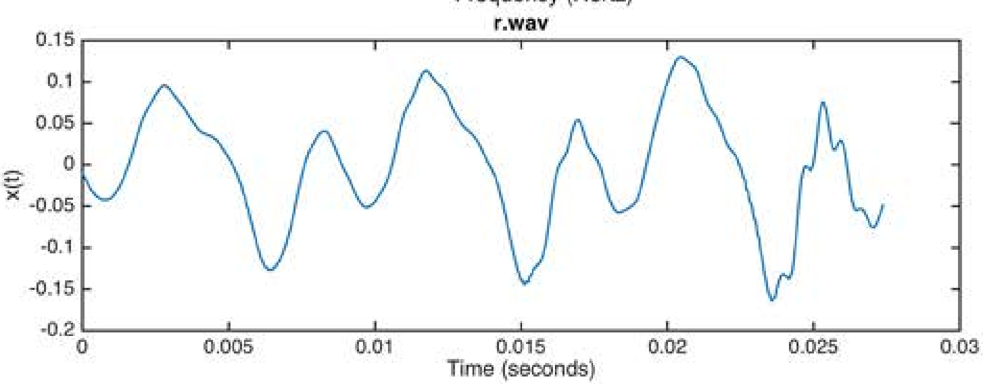

Narrowband Spectrum
Frequency (Hz)
Time (s)

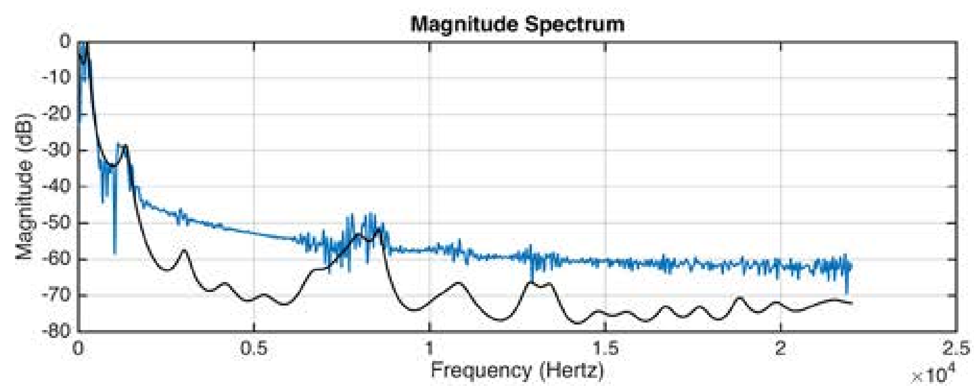

Wideband Spectrum
Frequency (Hz)
Time (s)

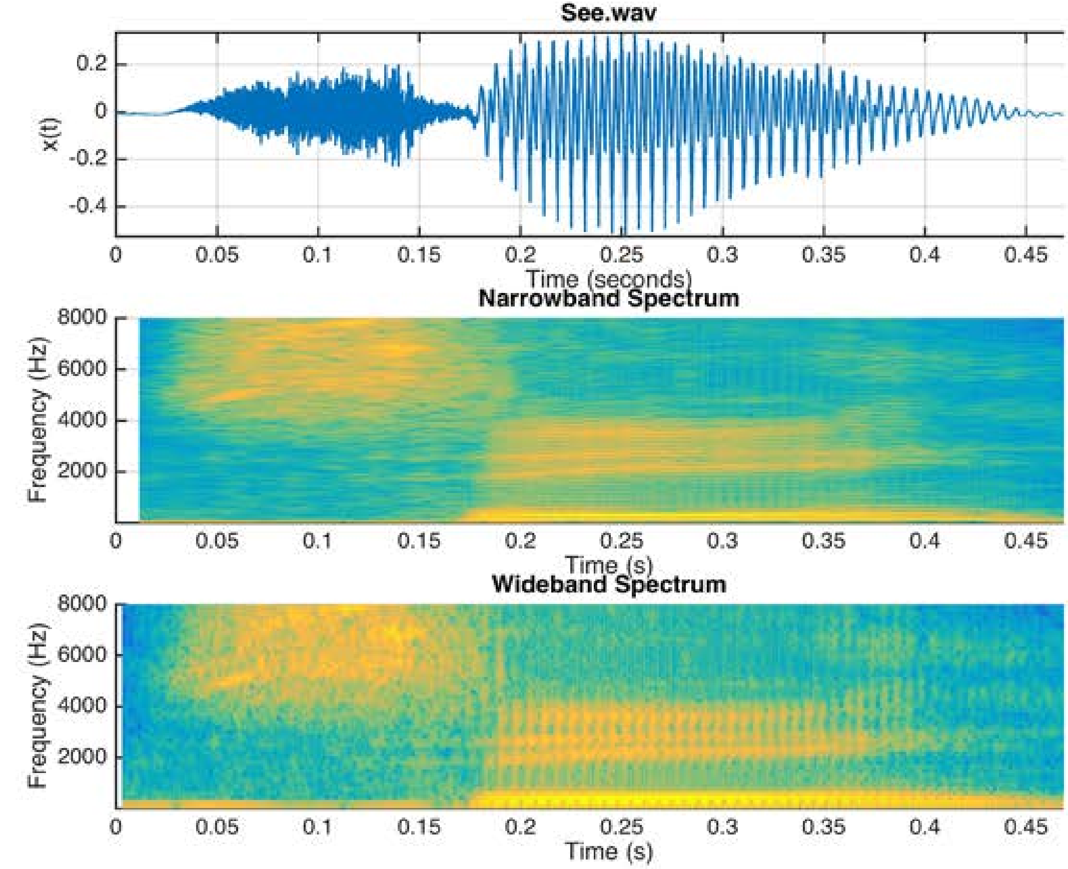

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

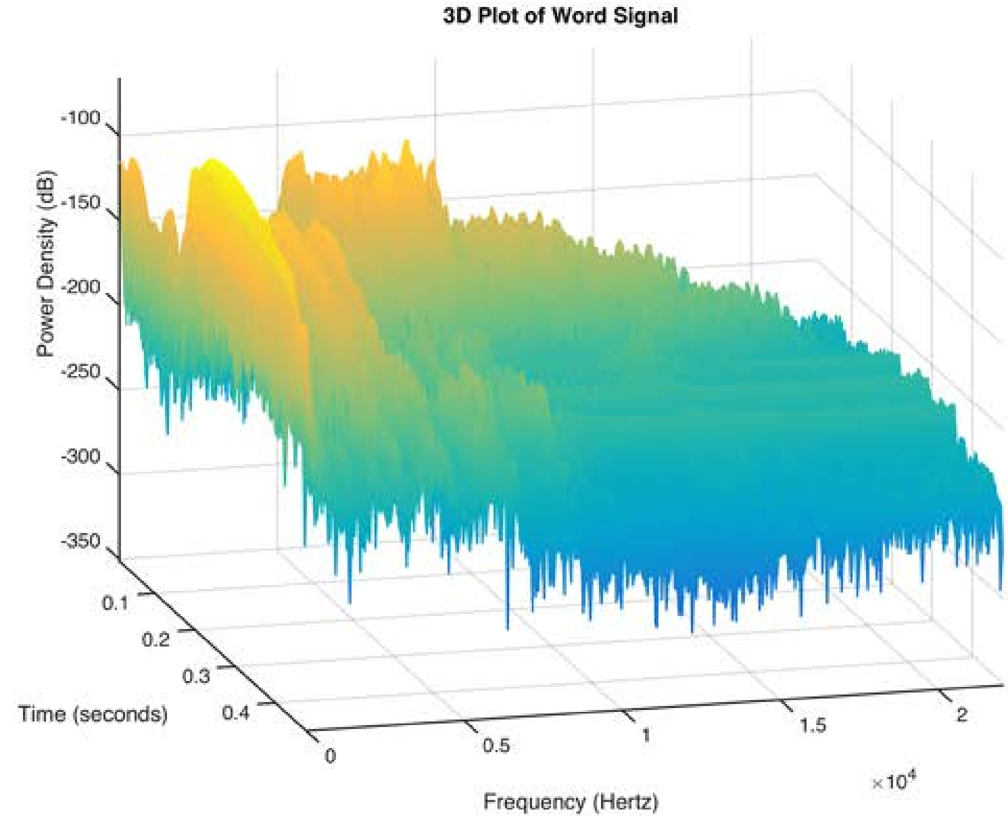
/r/

‘Read’


r.wav
x(t)
Time (seconds)

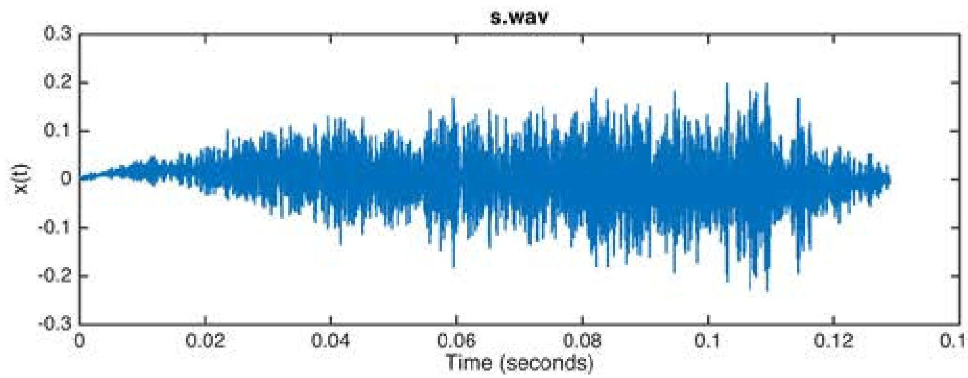

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

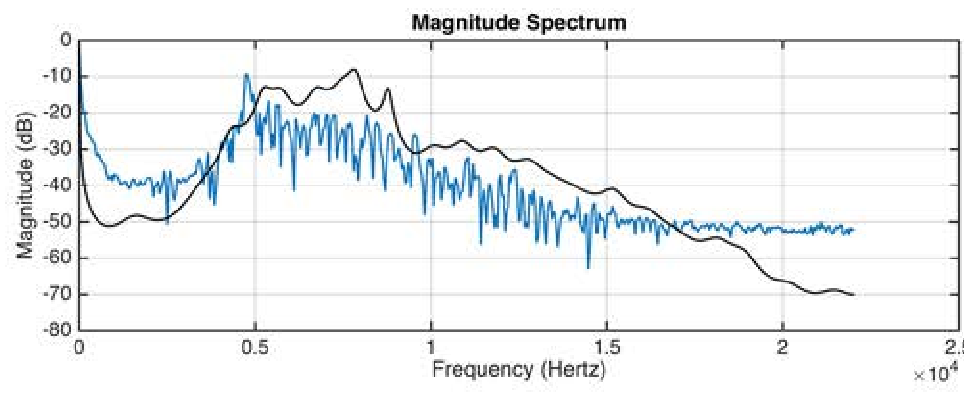

See.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

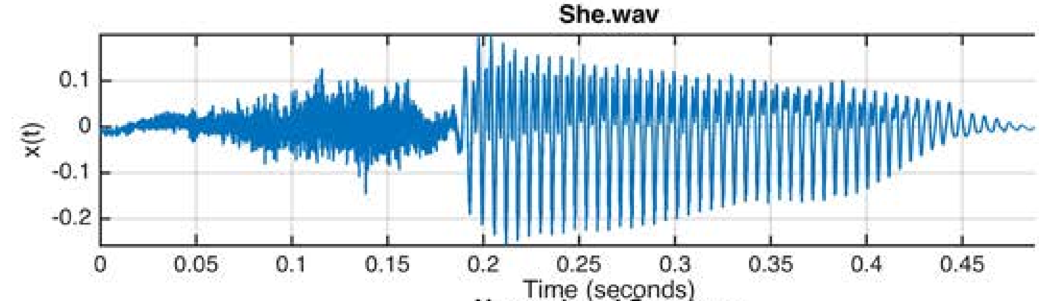

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
$\times 10^4$

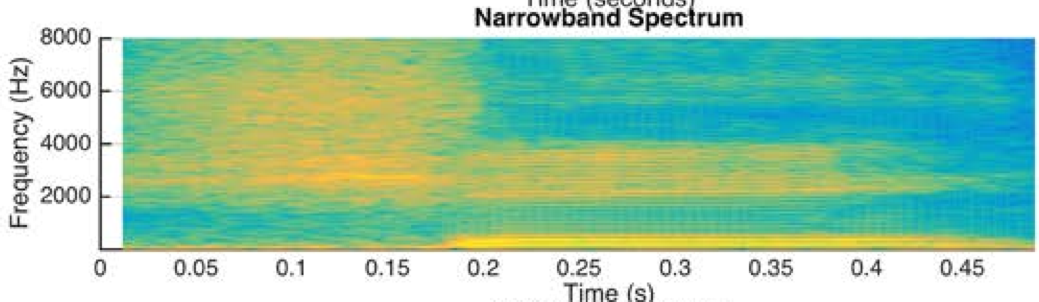
/s/

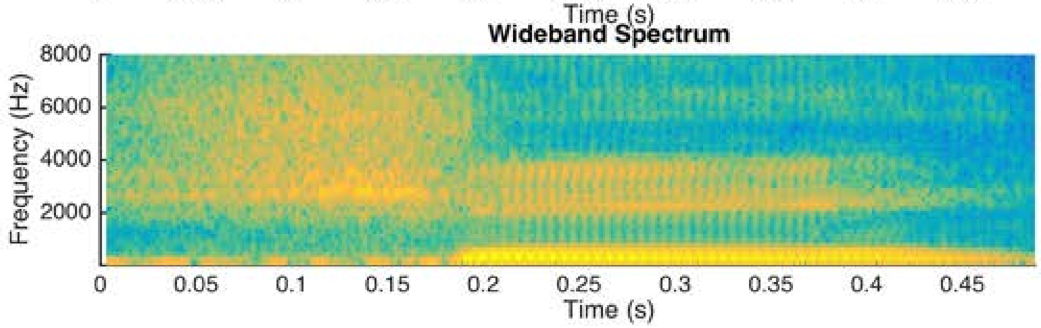
‘See’

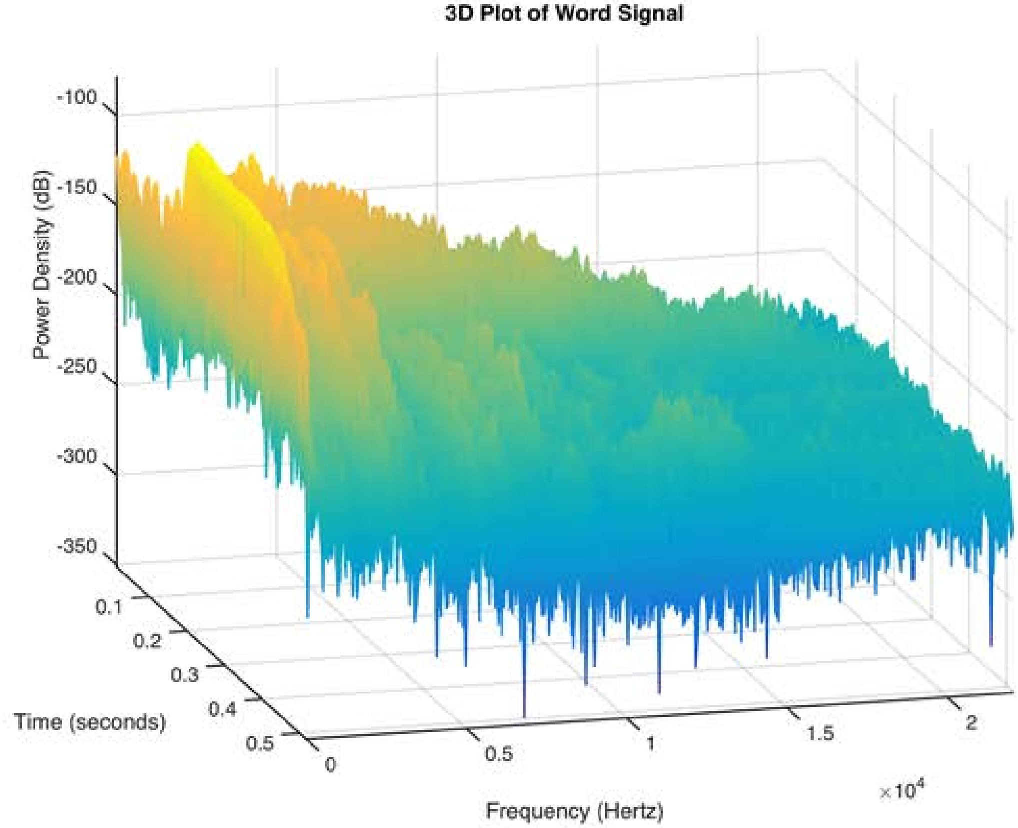

s.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
$\times 10^4$


She.wav
x(t)
Time (seconds)

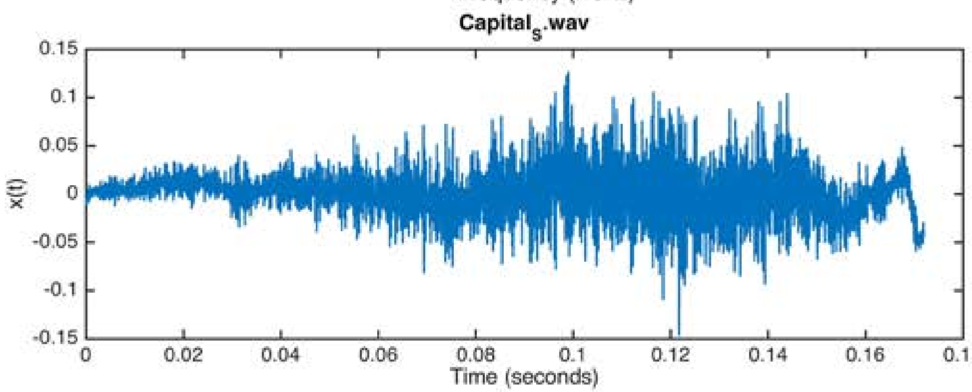

Narrowband Spectrum
Frequency (Hz)
Time (s)

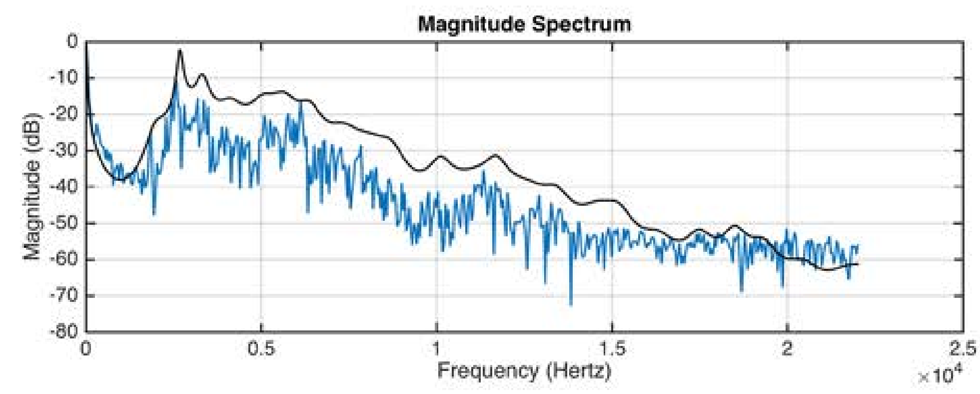

Wideband Spectrum
Frequency (Hz)
Time (s)

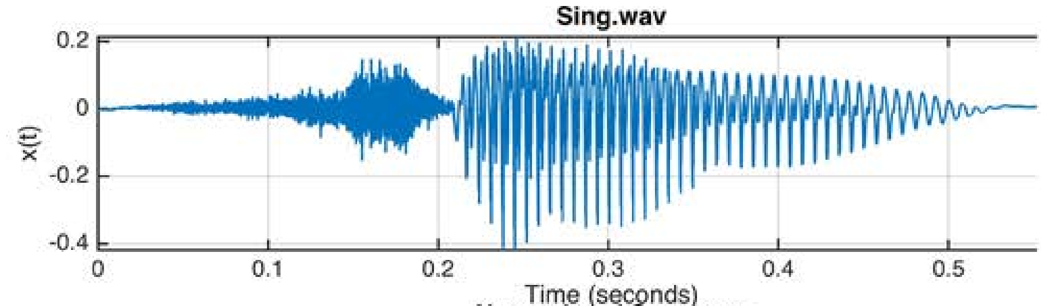

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

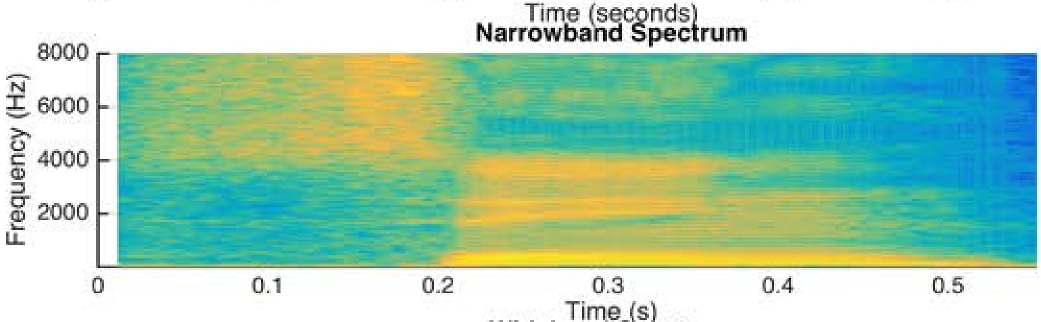
/S/

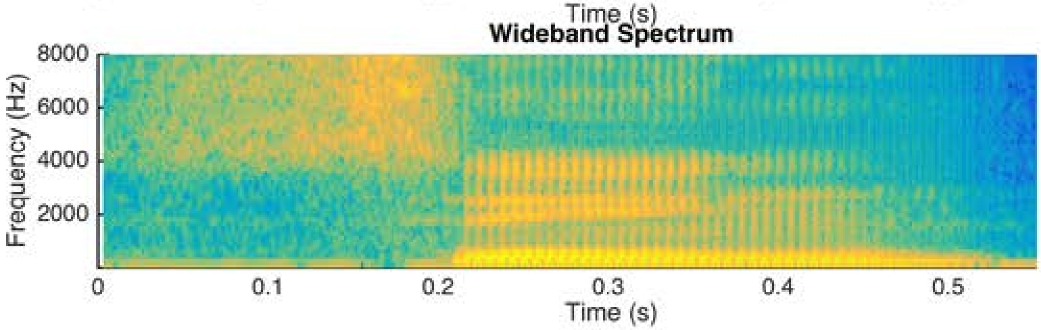
‘She’


Capital_s.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4

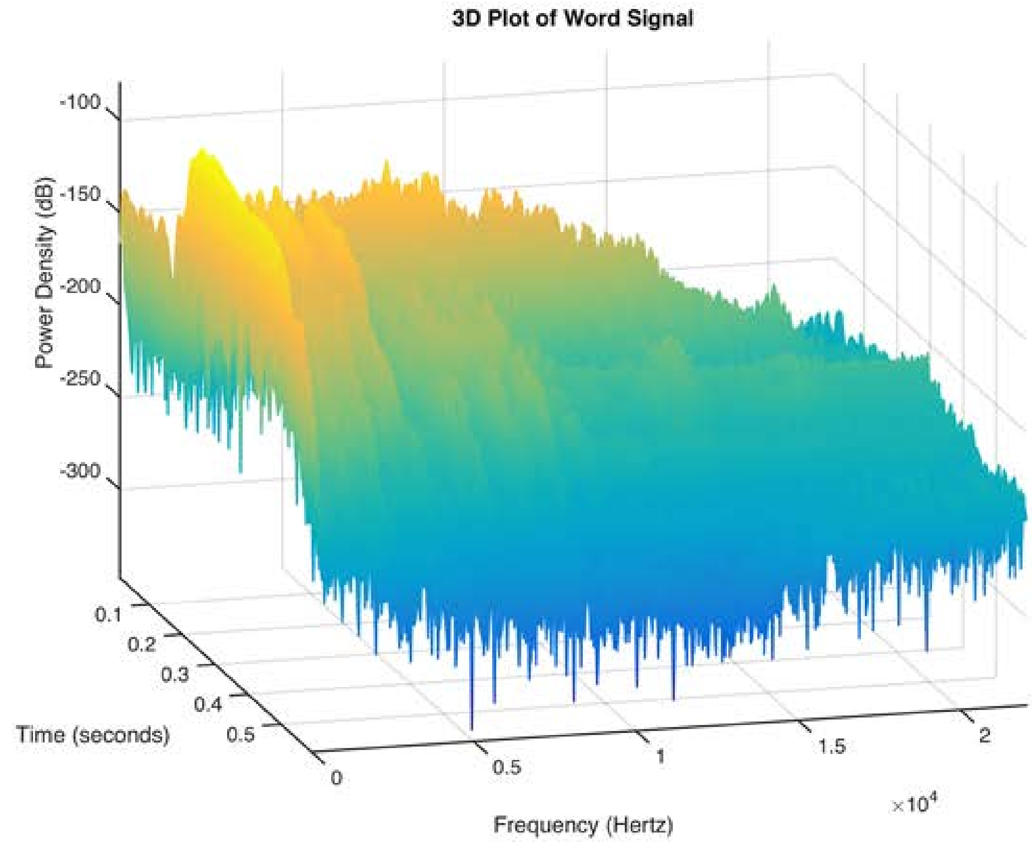

Sing.wav
x(t)
Time (seconds)

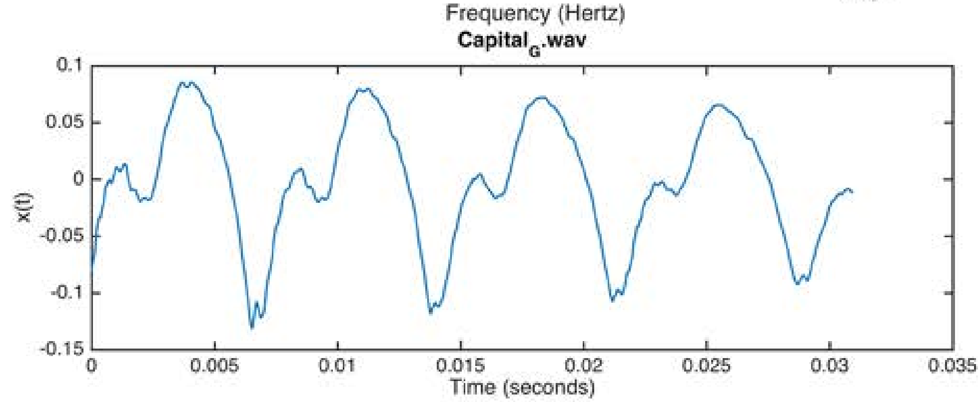

Narrowband Spectrum
Frequency (Hz)
Time (s)

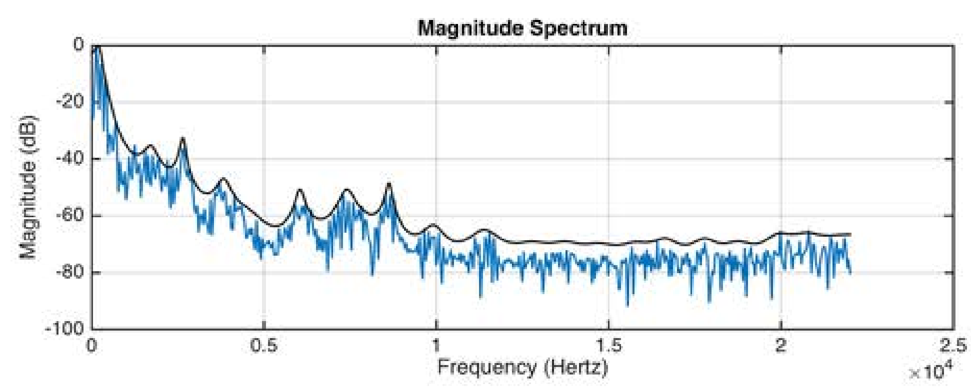

Wideband Spectrum
Frequency (Hz)
Time (s)

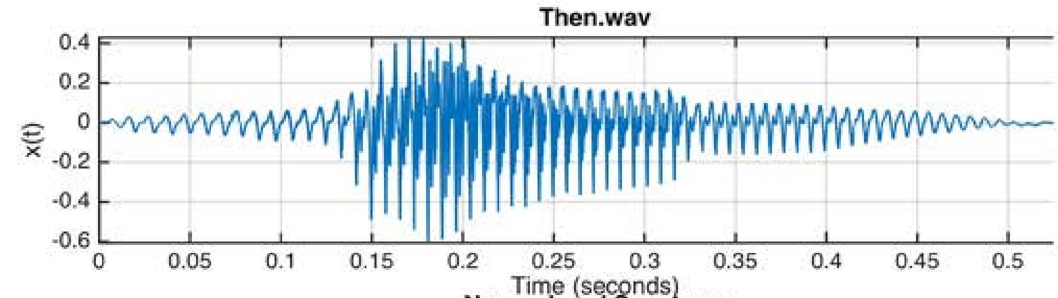
/G/

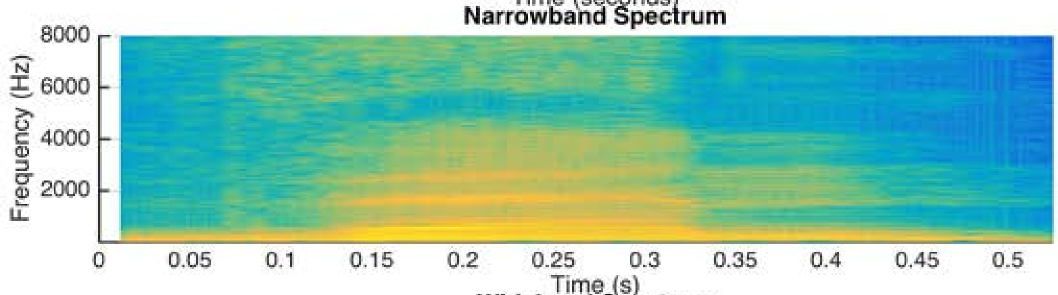
‘Sing’

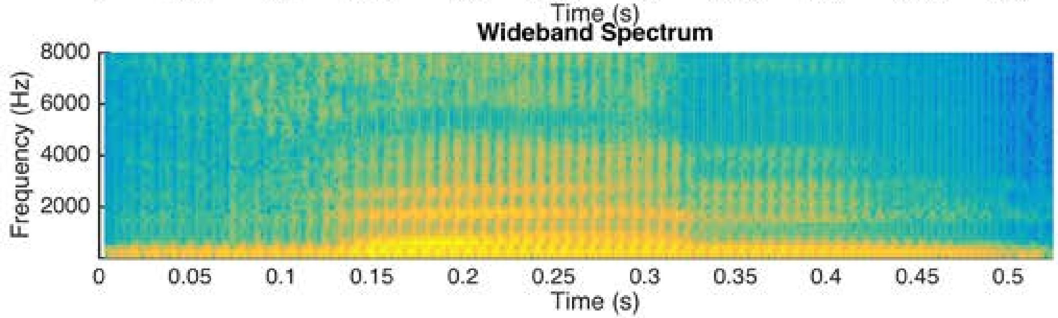

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

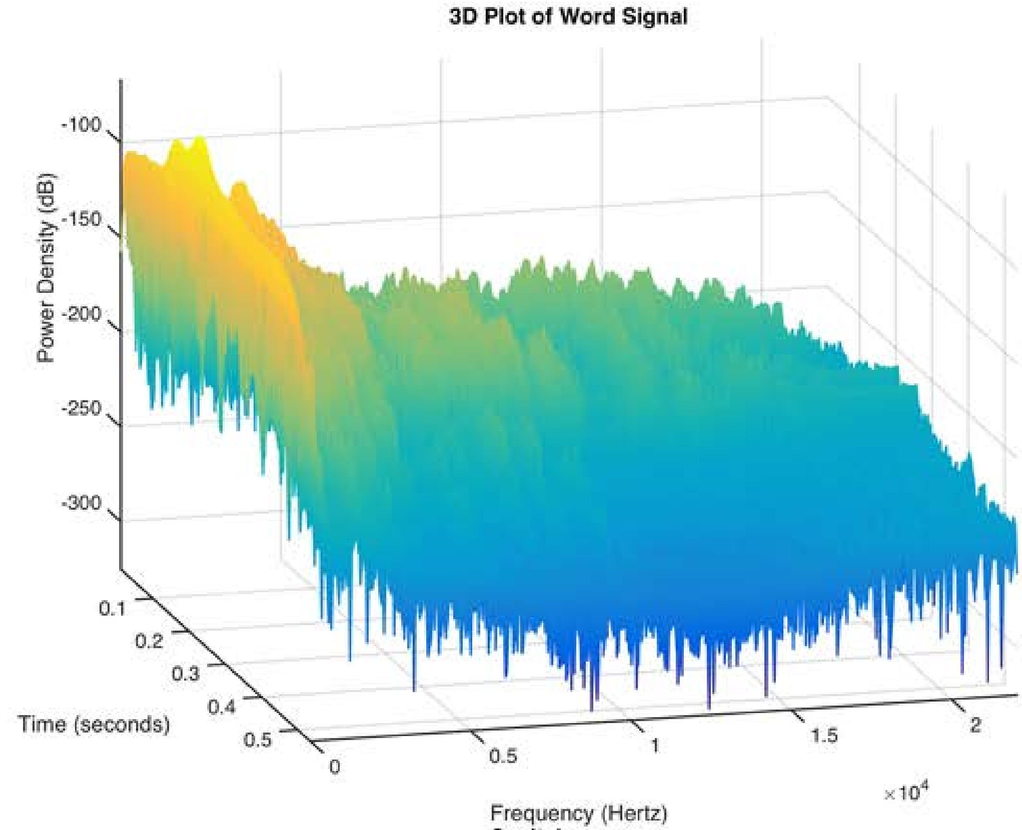

Capital_G.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


Then.wav
x(t)
Time (seconds)

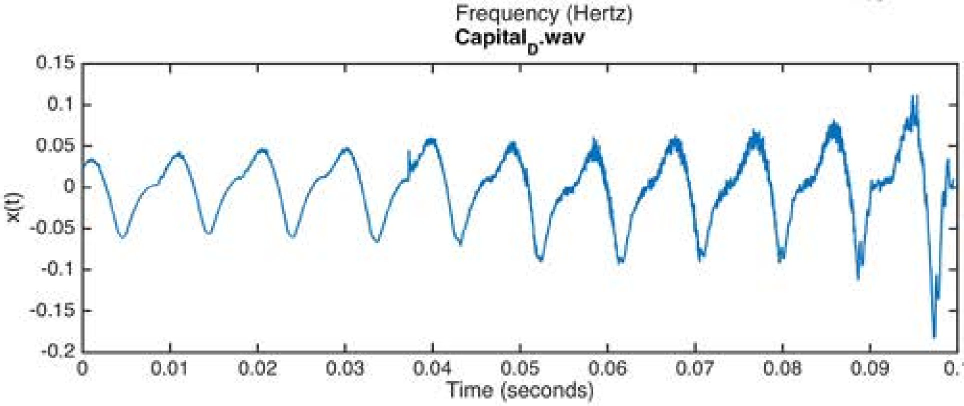

Narrowband Spectrum
Frequency (Hz)
Time (s)

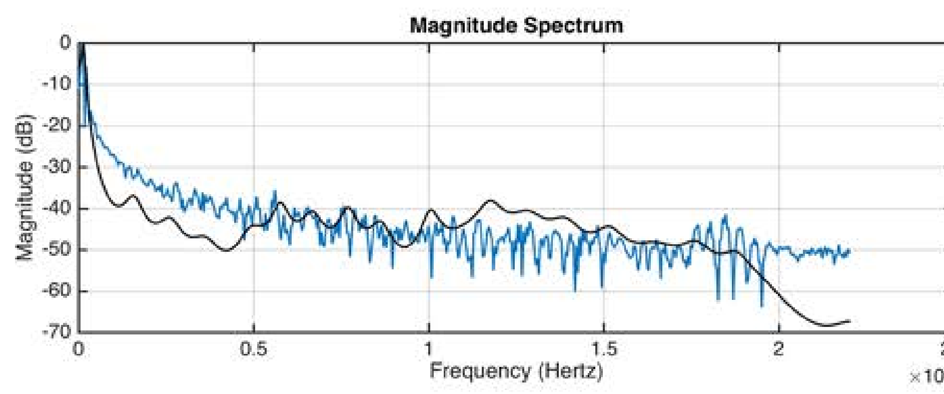

Wideband Spectrum
Frequency (Hz)
Time (s)

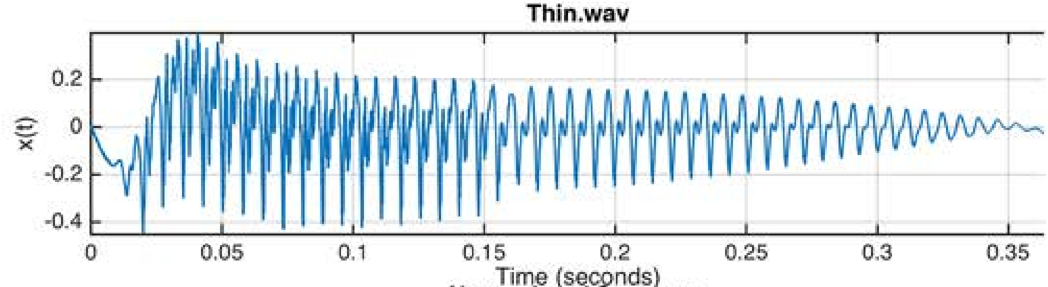

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

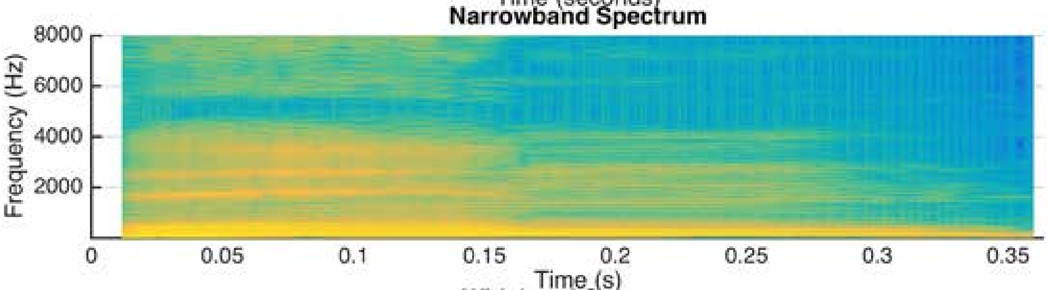
/D/

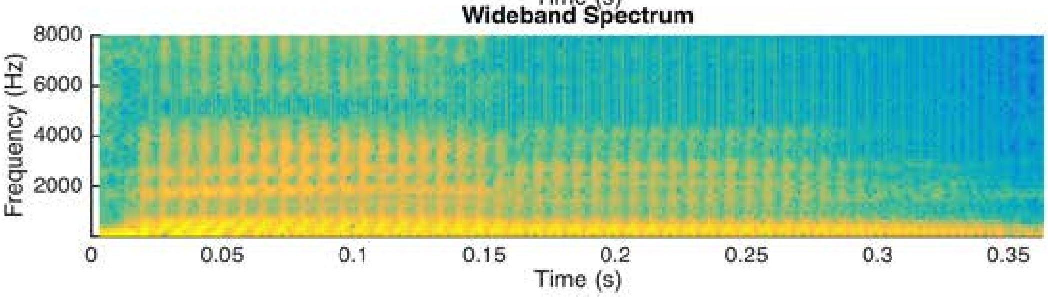
‘Then’

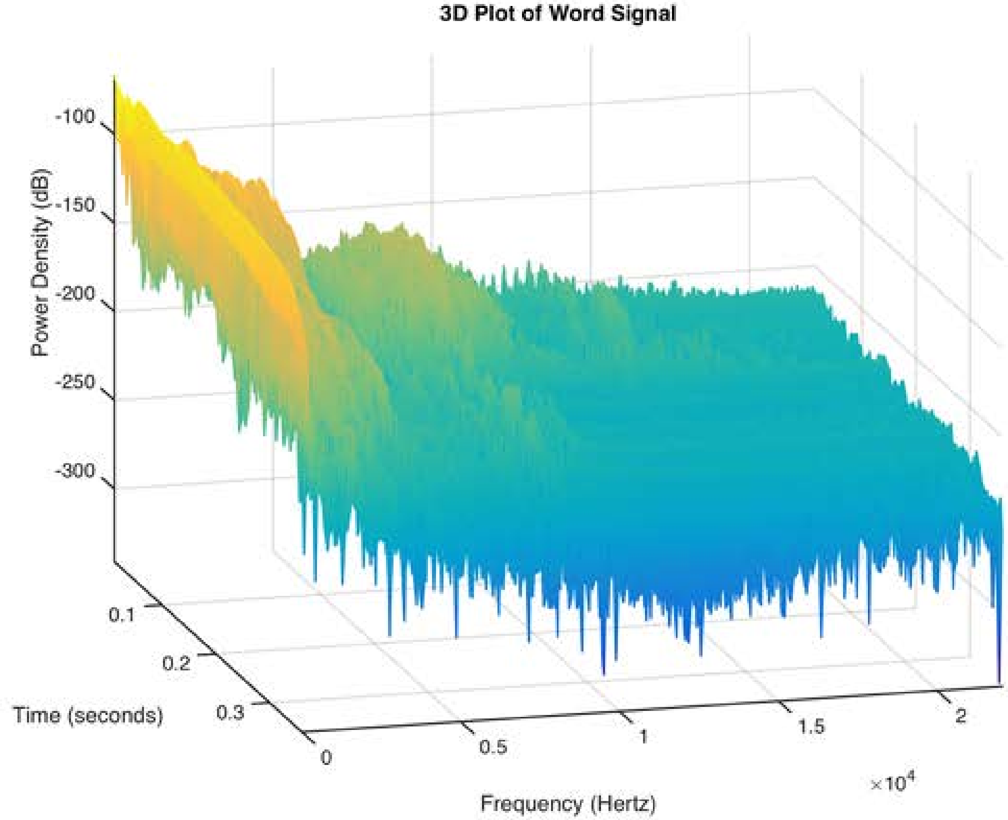

Capital_D.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


Thin.wav
x(t)
Time (seconds)

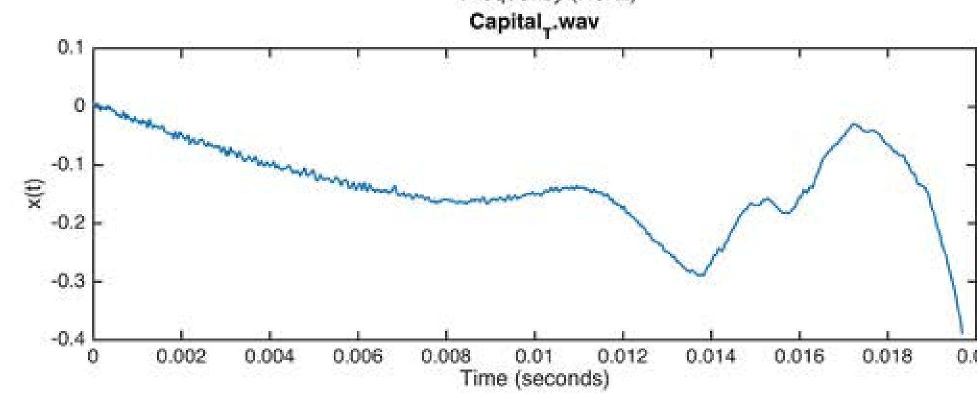

Narrowband Spectrum
Frequency (Hz)
Time (s)

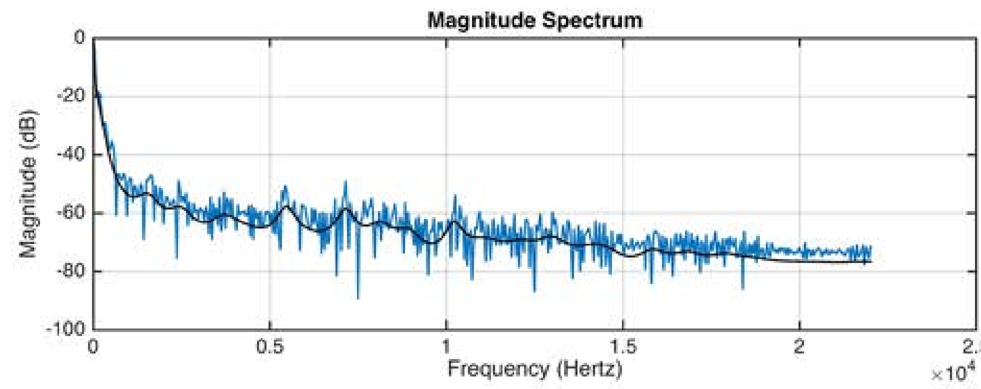

Wideband Spectrum
Frequency (Hz)
Time (s)

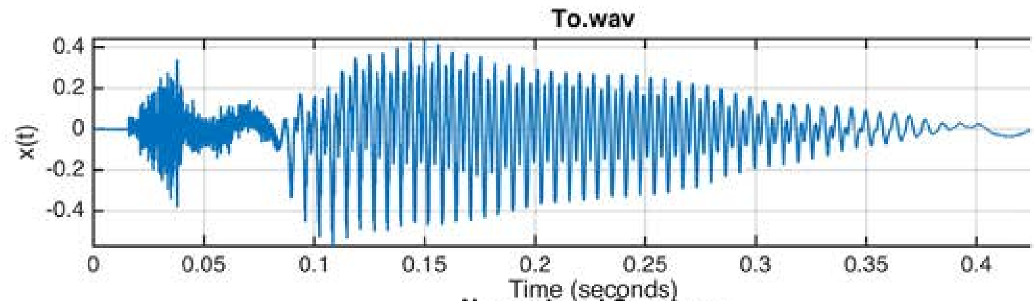

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

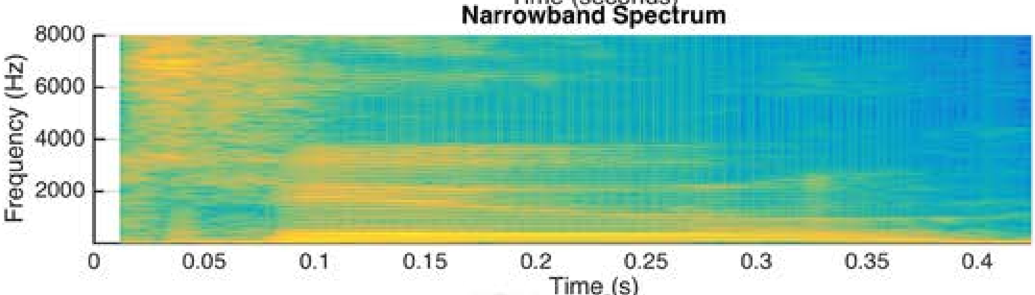
/T/

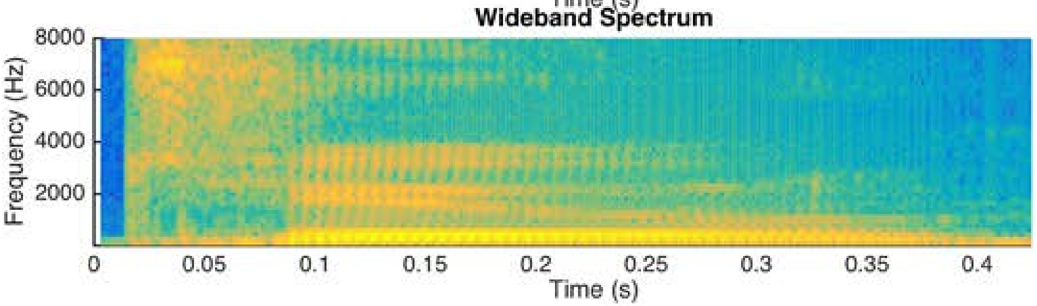
'Thin'

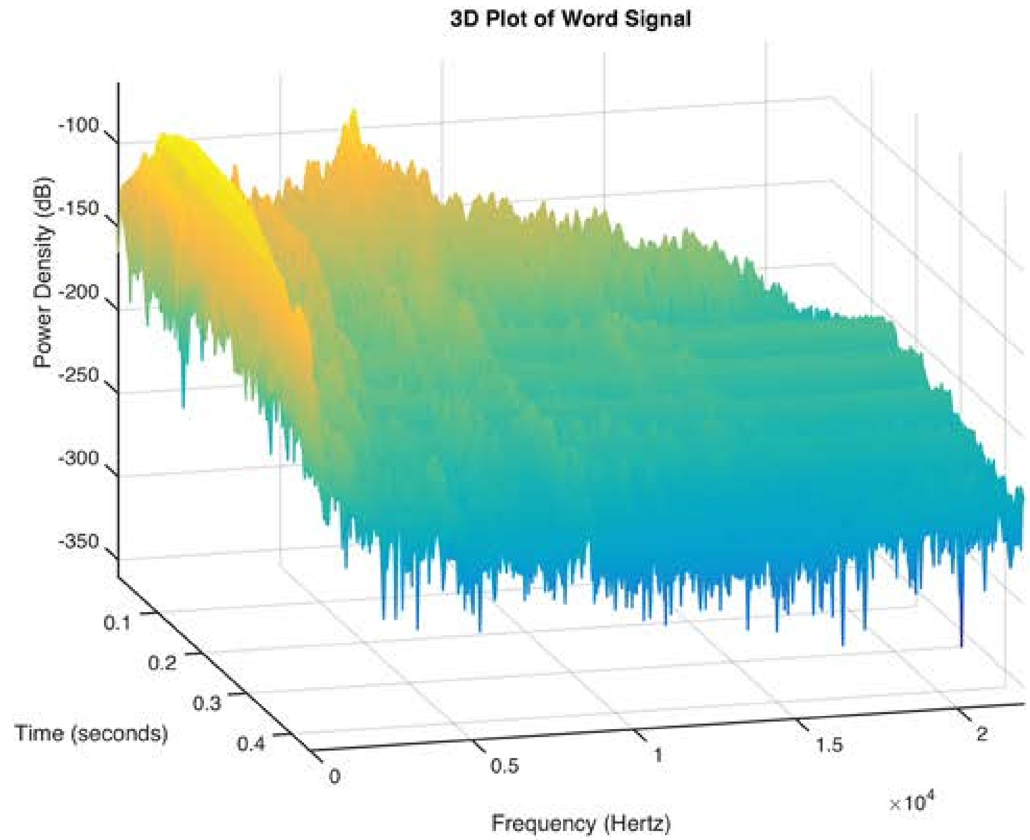

Capital_T.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


To.wav
x(t)
0.4
0.2
0
-0.2
-0.4
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
Time (seconds)

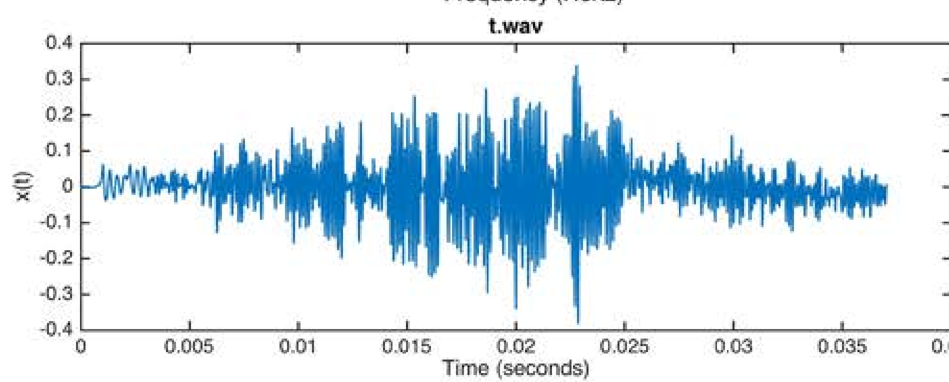

Narrowband Spectrum
Frequency (Hz)
8000
6000
4000
2000
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
Time (s)

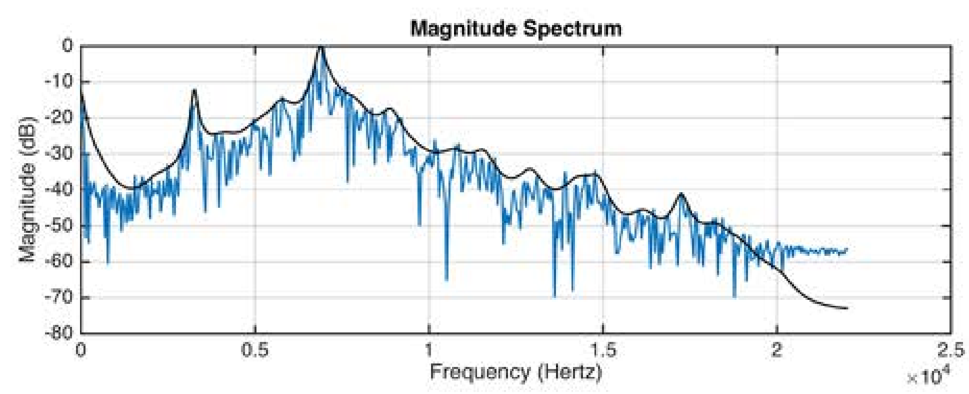

Wideband Spectrum
Frequency (Hz)
8000
6000
4000
2000
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
Time (s)

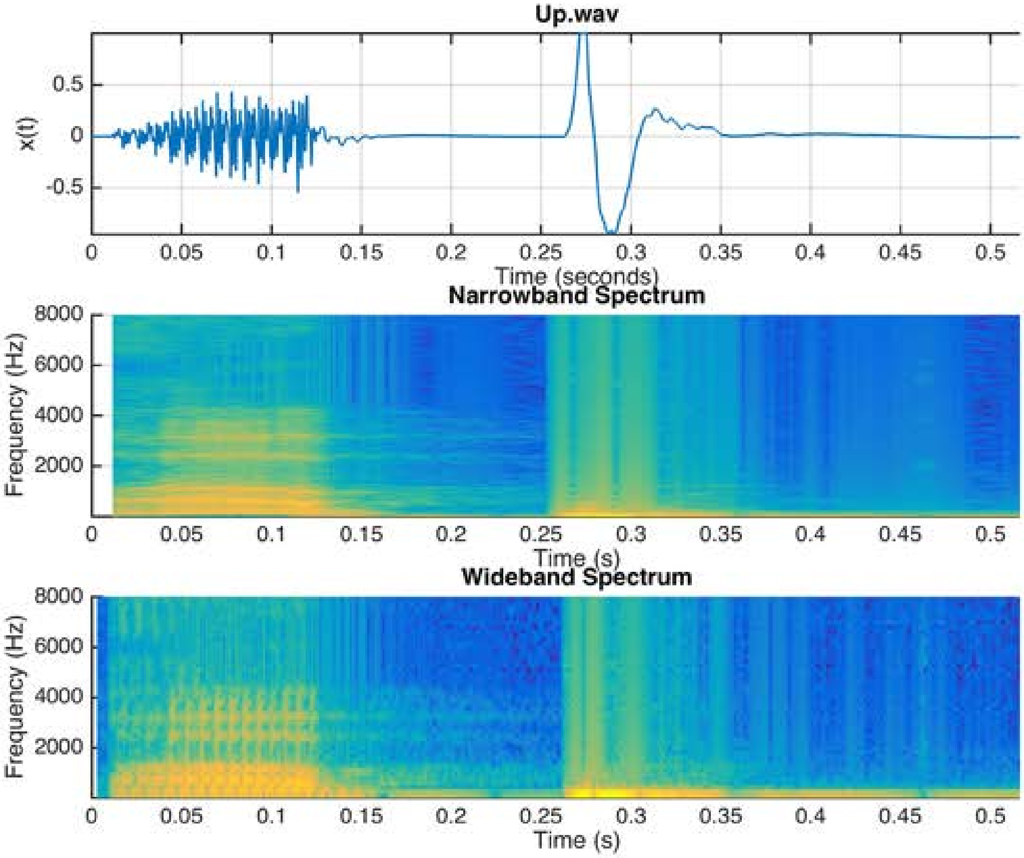

3D Plot of Word Signal
Power Density (dB)
-100
-150
-200
-250
-300
-350
Time (seconds)
0.1
0.2
0.3
0.4
Frequency (Hertz)
0
0.5
1
1.5
2
×10^4

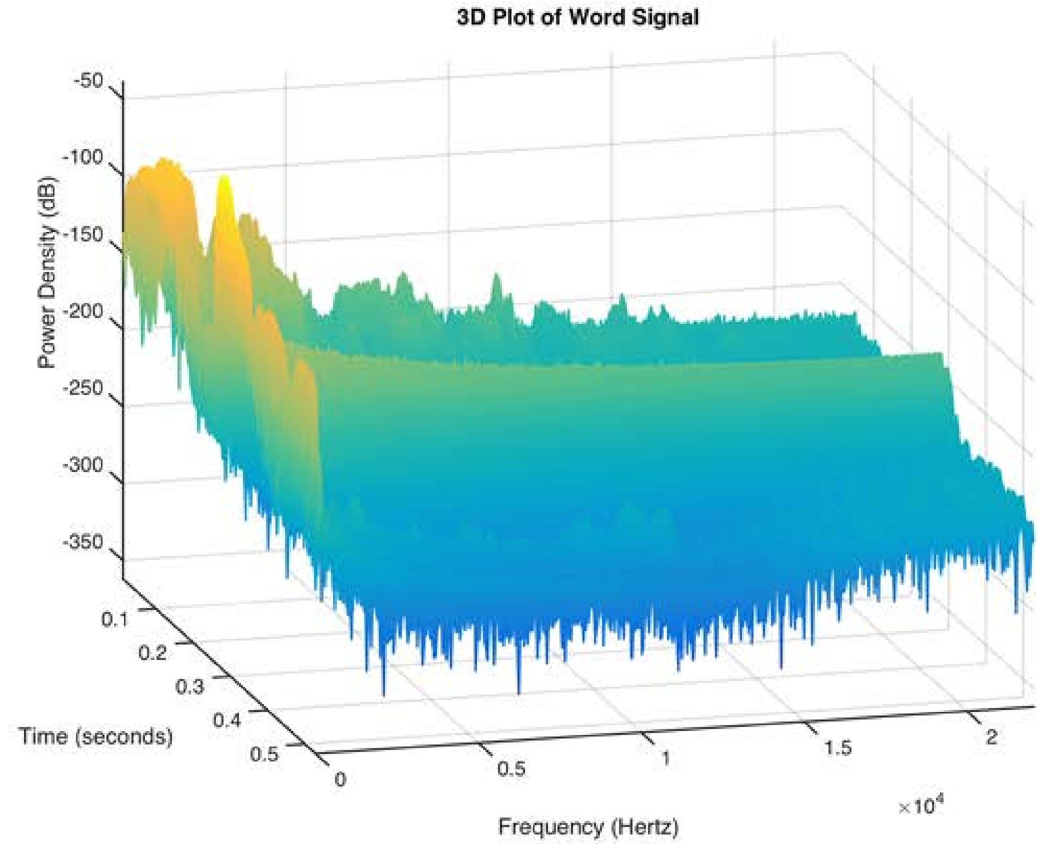
/t/

‘To’


t.wav
x(t)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
Time (seconds)

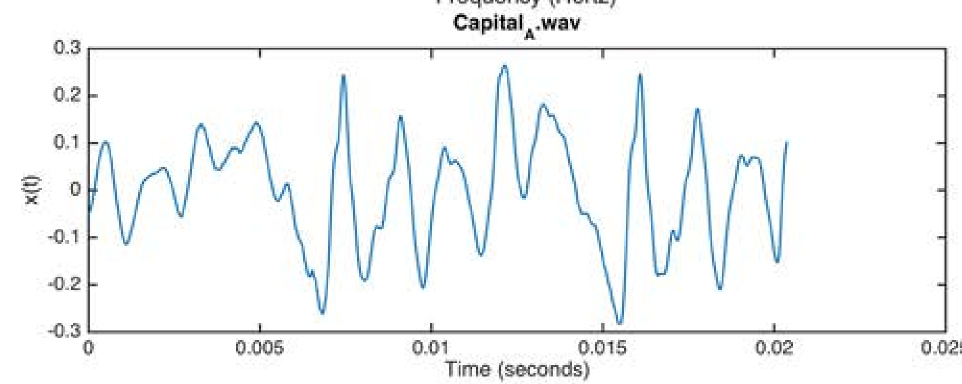

Magnitude Spectrum
Magnitude (dB)
0
-10
-20
-30
-40
-50
-60
-70
-80
0
0.5
1
1.5
2
2.5
×10^4
Frequency (Hertz)

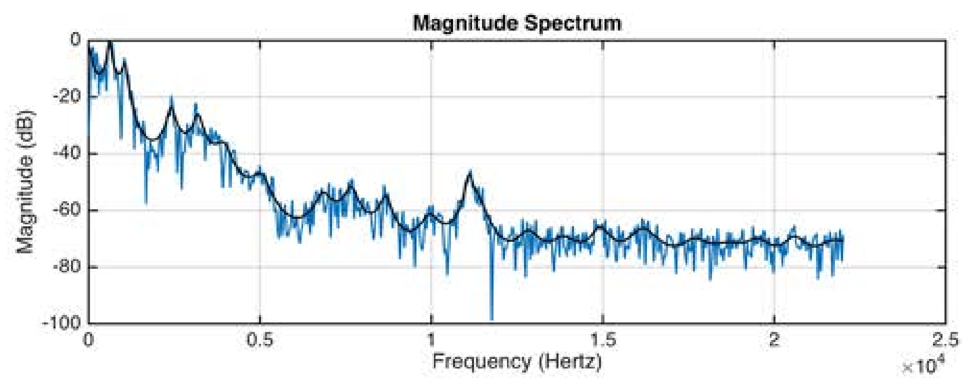

Up.wav
x(t)
Time (seconds)
Narrowband Spectrum
Frequency (Hz)
Time (s)
Wideband Spectrum
Frequency (Hz)
Time (s)

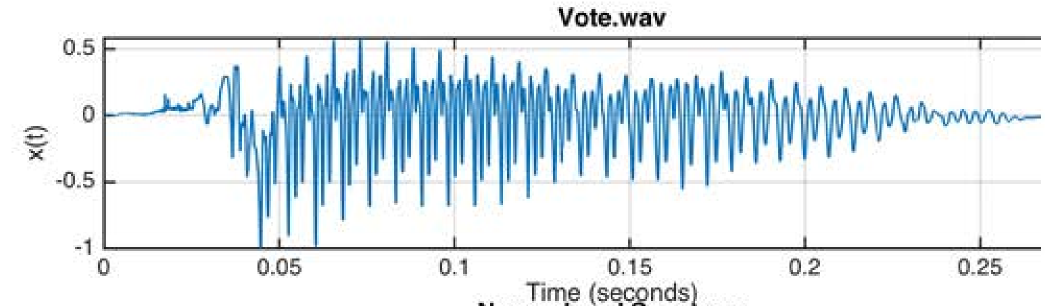

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)
×10^4

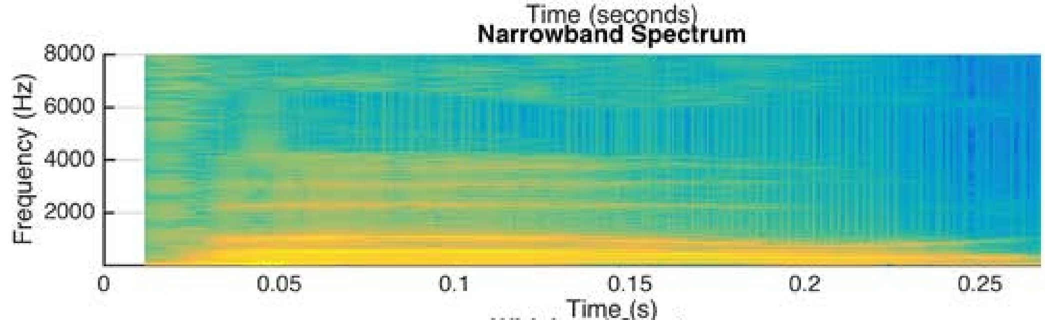
/A/

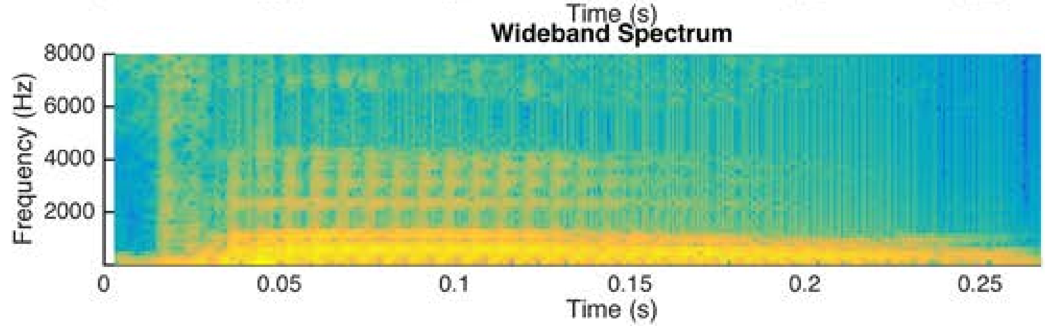
‘Up’

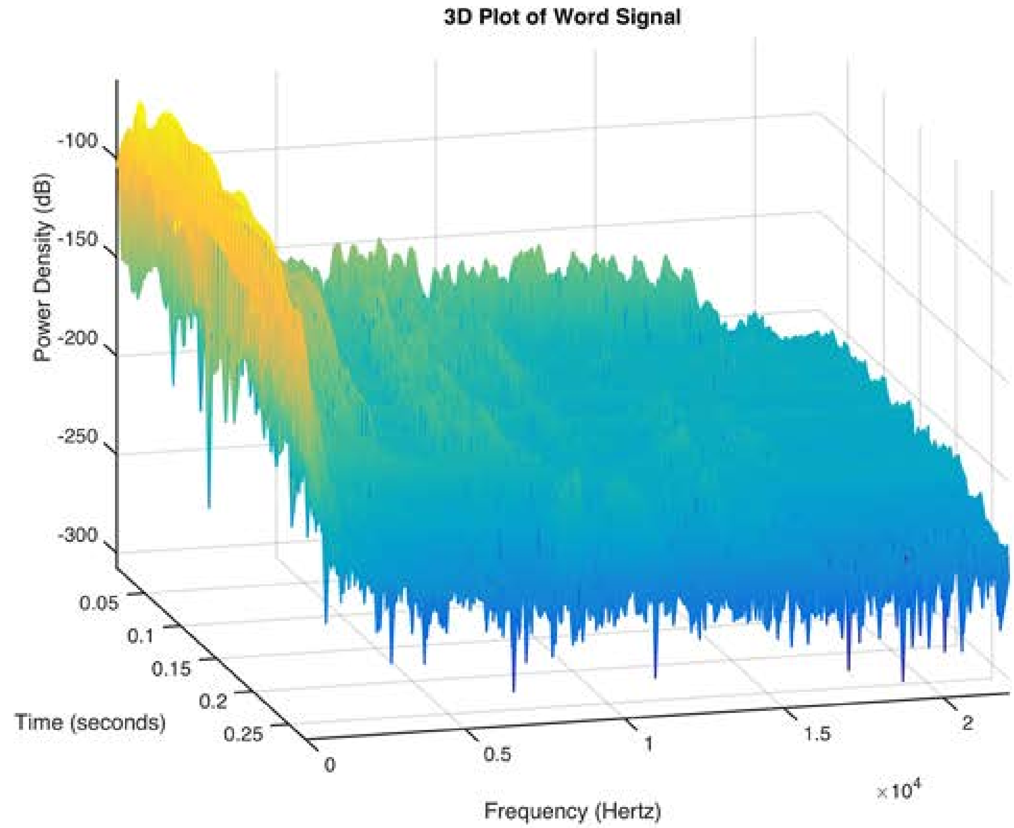

Capital_A.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
×10^4


Vote.wav
x(t)
Time (seconds)

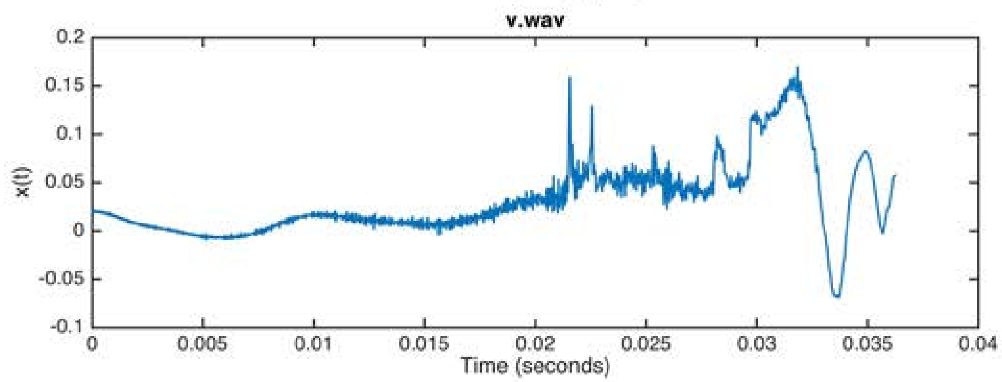

Narrowband Spectrum
Frequency (Hz)
Time (s)

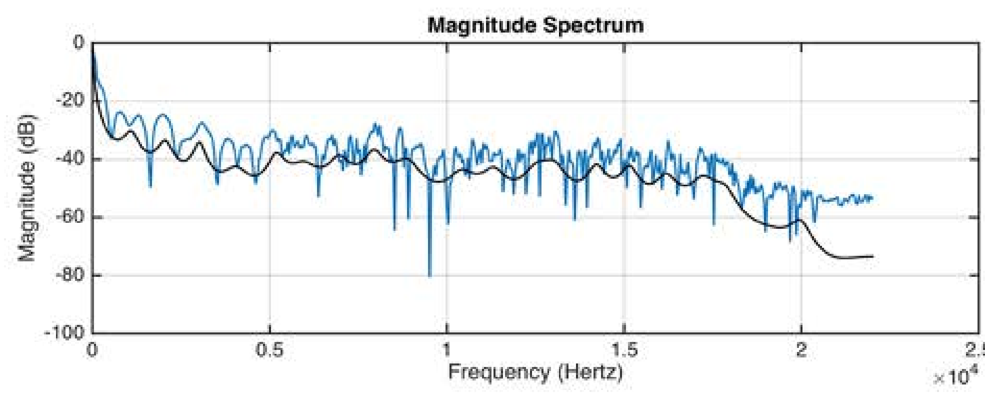

Wideband Spectrum
Frequency (Hz)
Time (s)

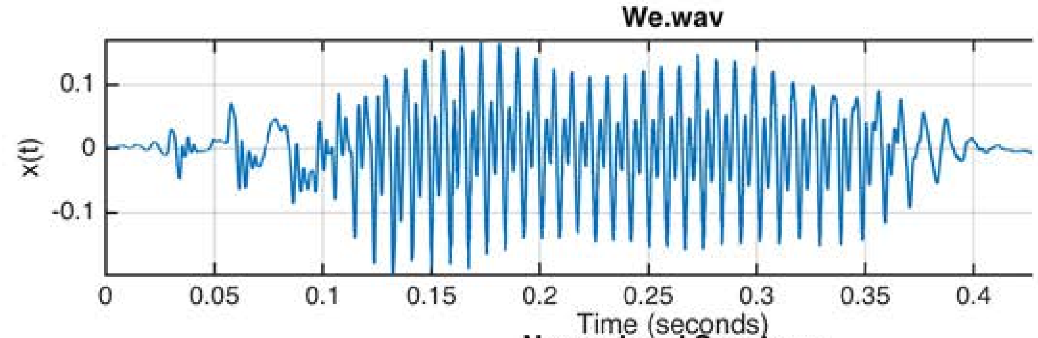

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

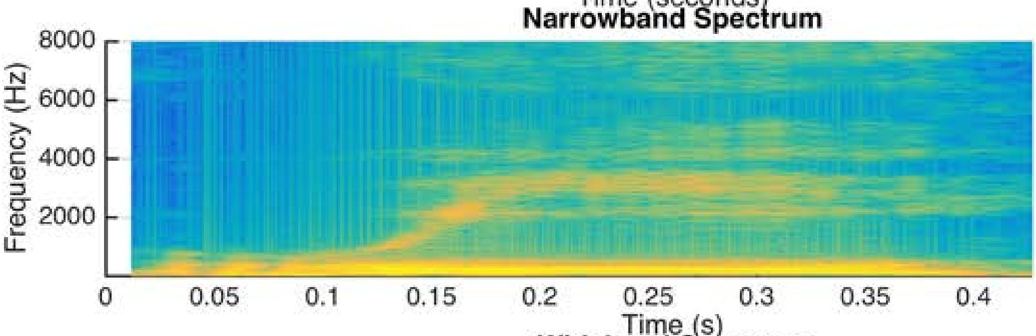
/v/

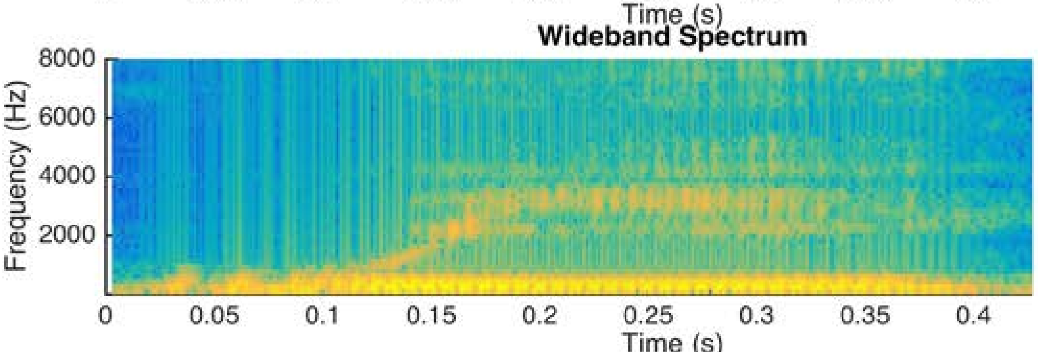
'Vote'


v.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

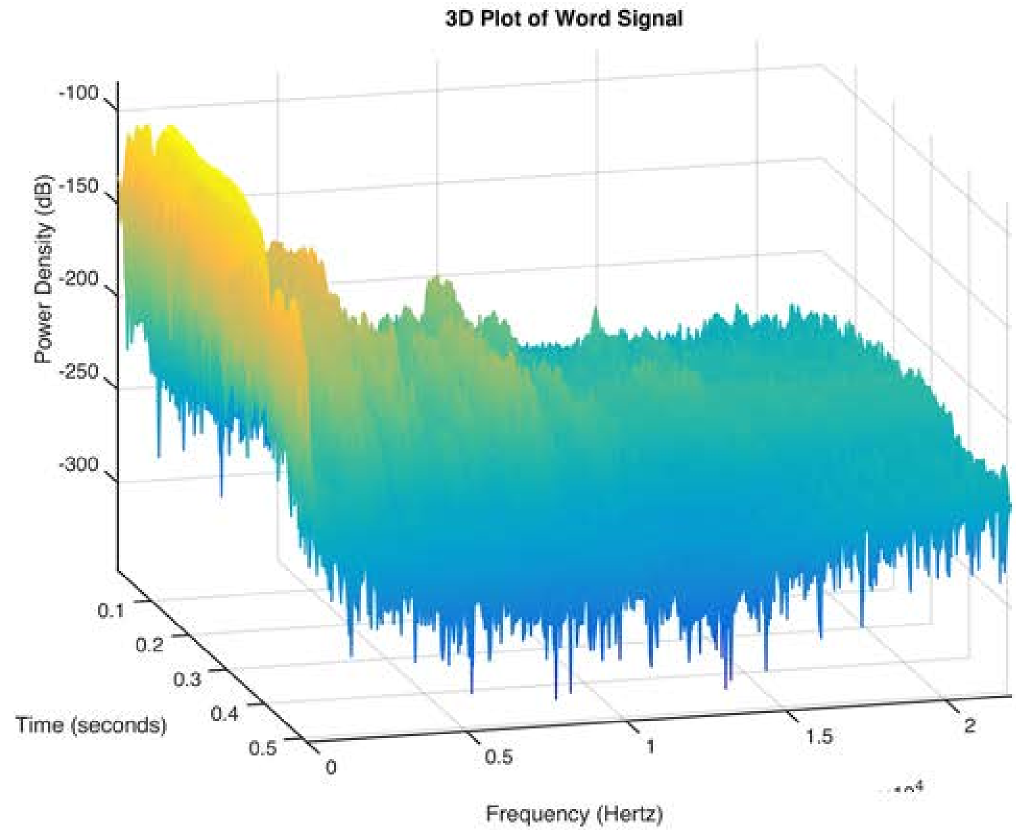

We.wav
x(t)
Time (seconds)

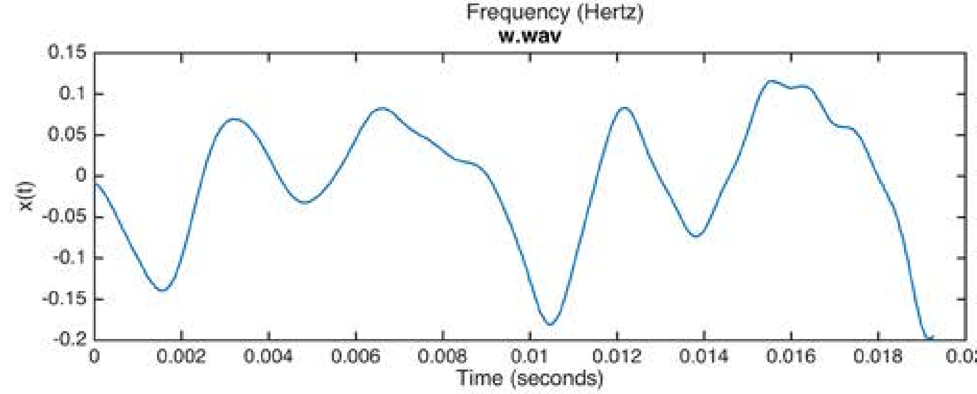

Narrowband Spectrum
Frequency (Hz)
Time (s)

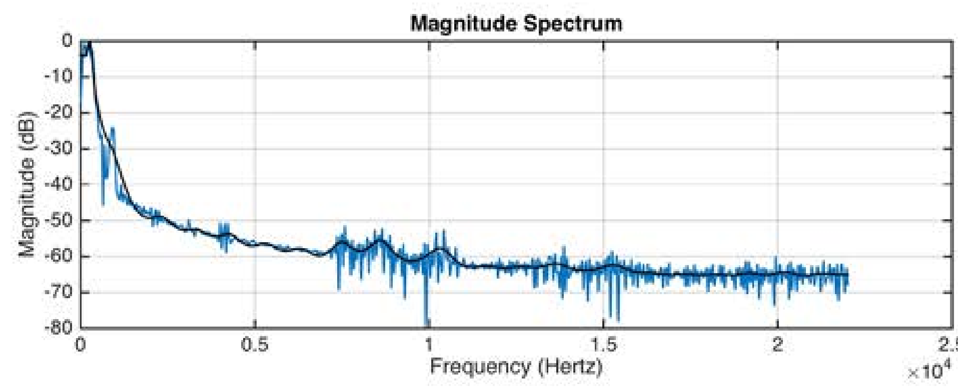

Wideband Spectrum
Frequency (Hz)
Time (s)

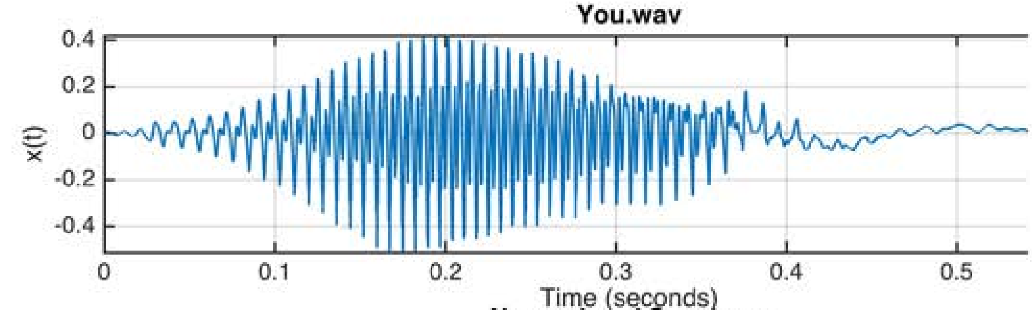
/w/

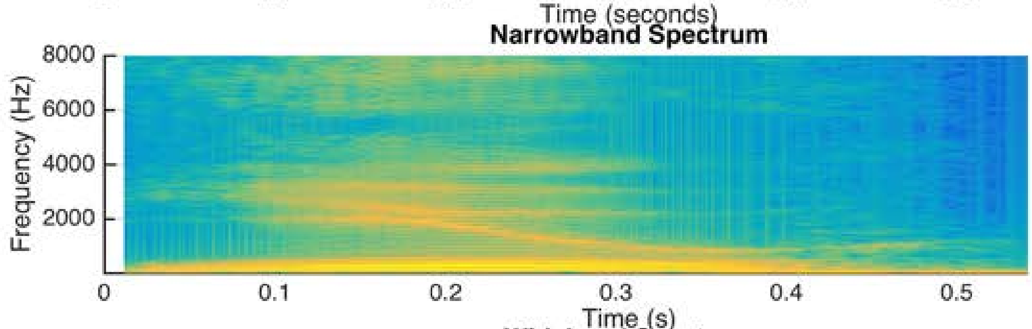
‘We’

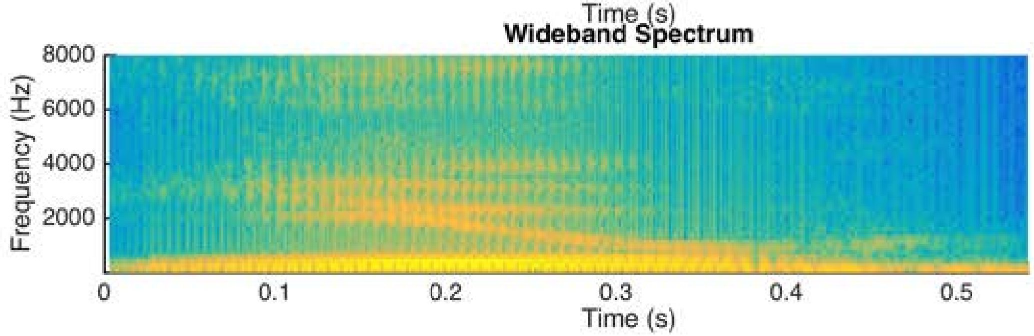

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

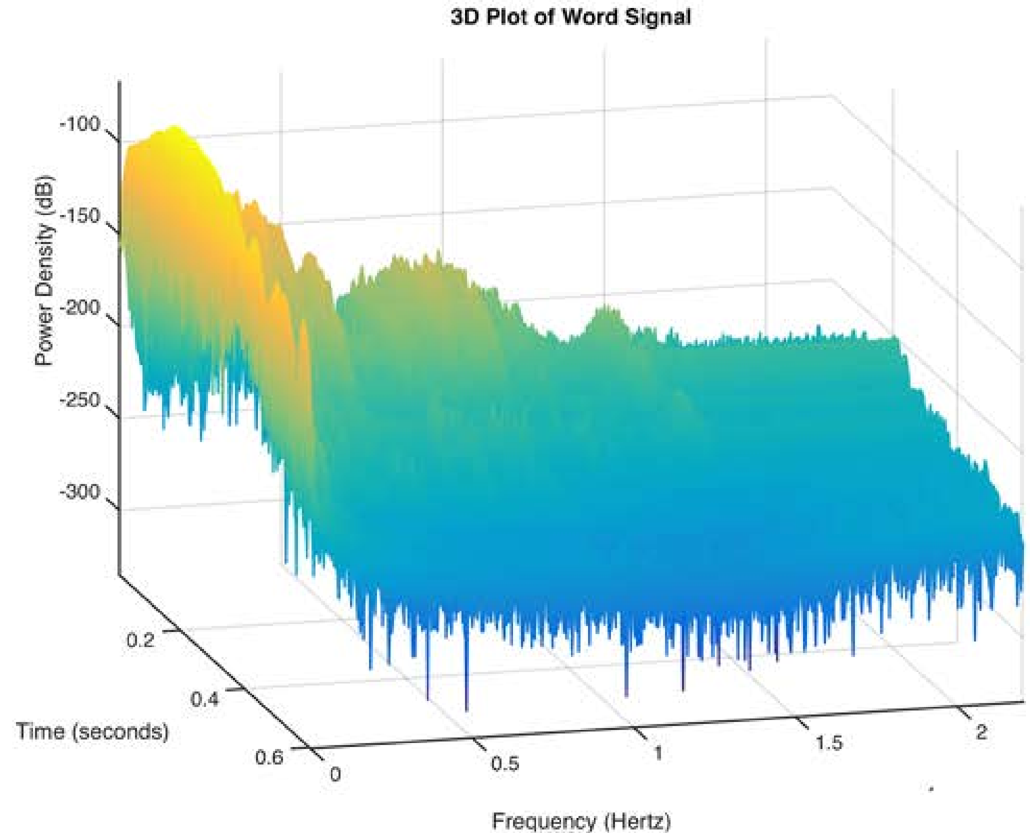

w.wav
x(t)
Time (seconds)

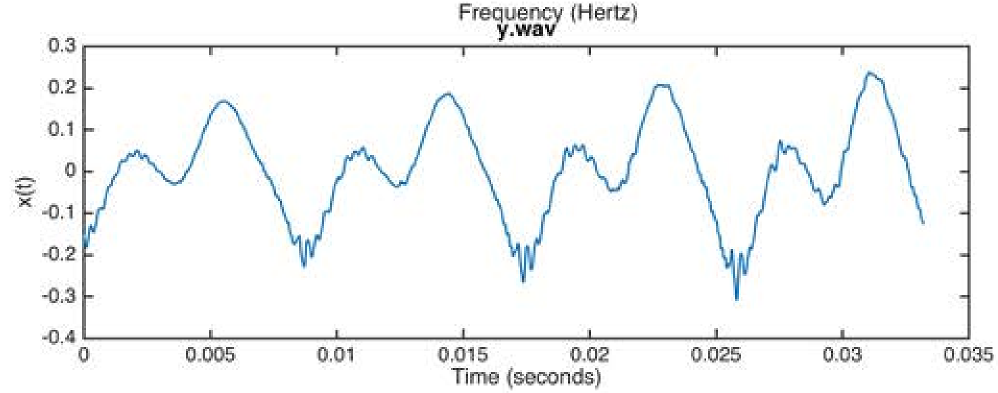

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)


You.wav
x(t)
Time (seconds)


Narrowband Spectrum
Frequency (Hz)
Time (s)

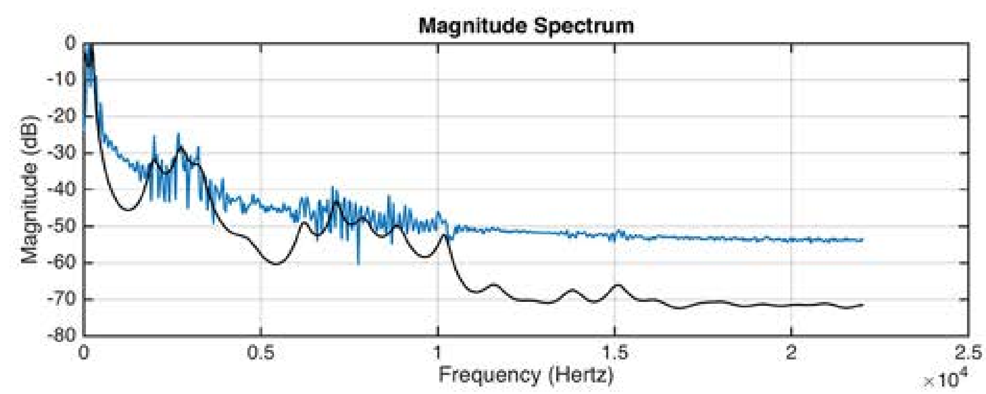

Wideband Spectrum
Frequency (Hz)
Time (s)

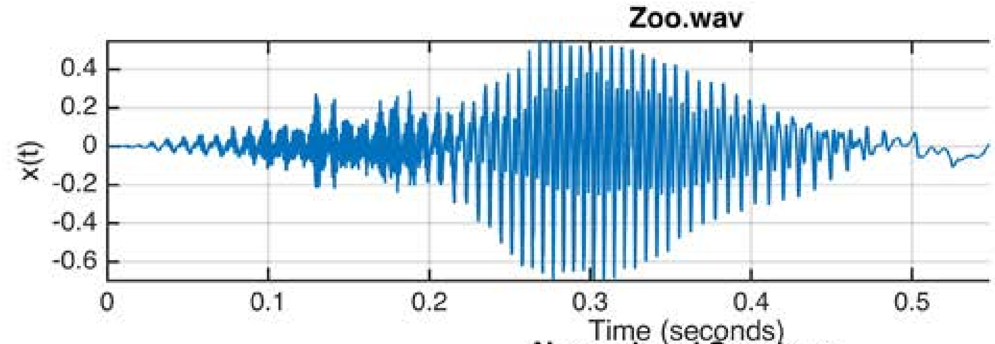

3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)

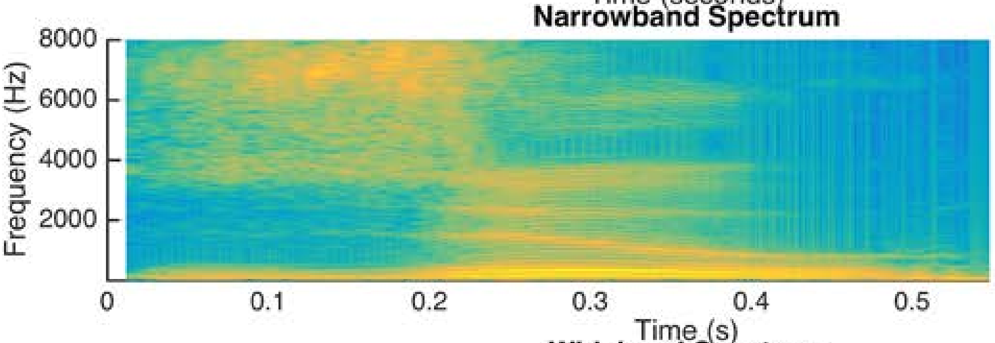

y.wav
x(t)
Time (seconds)

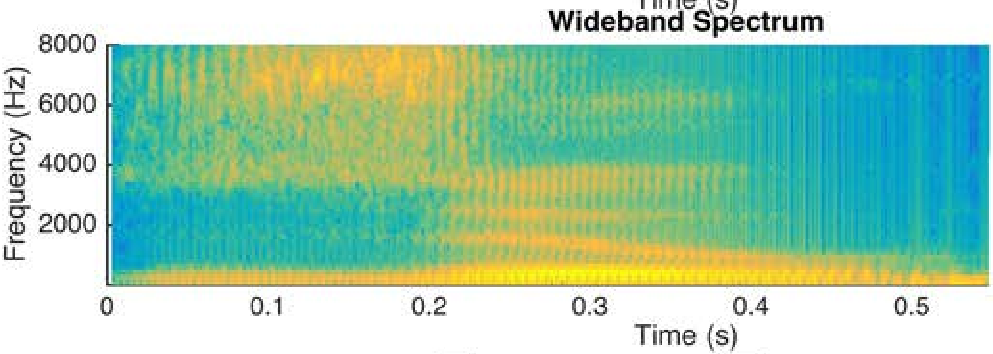
/y/

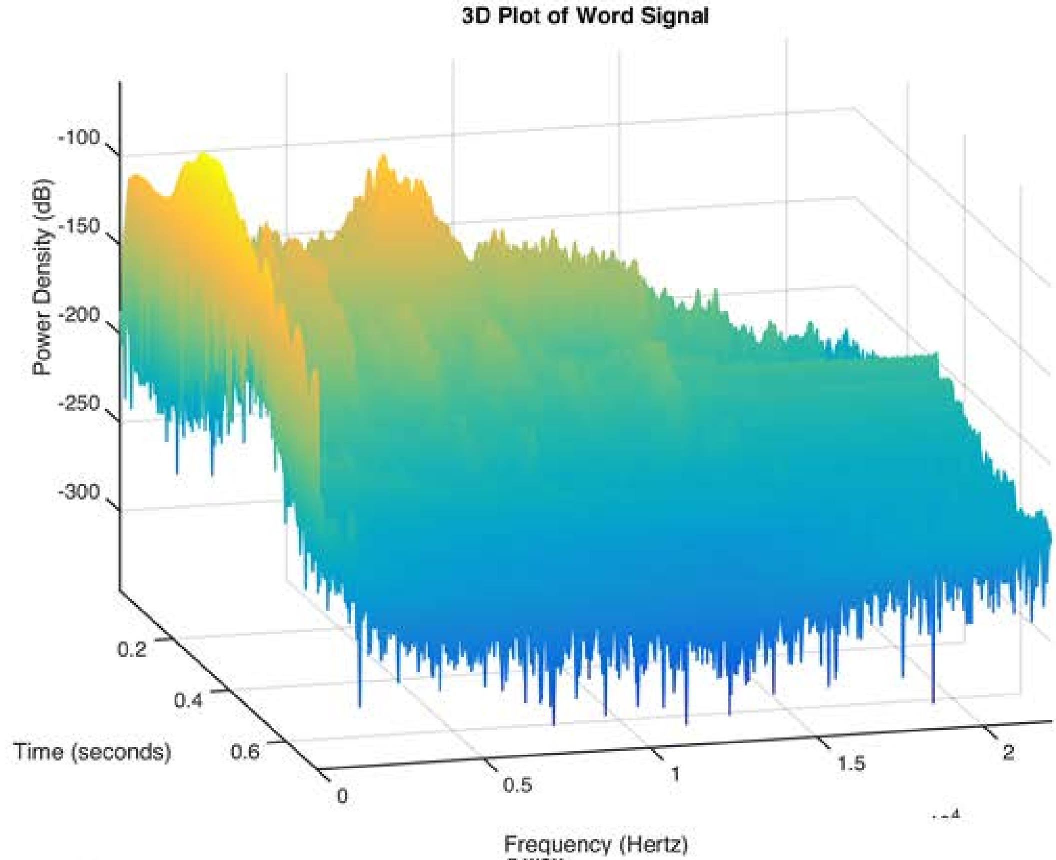
'You'

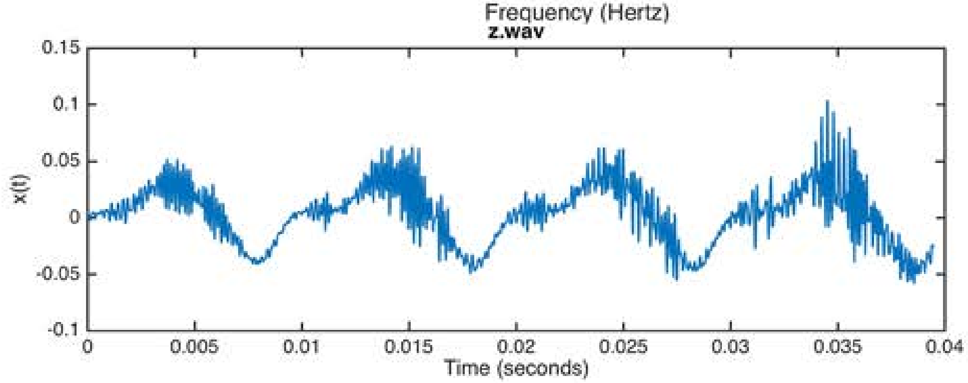

Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)
$\times 10^4$

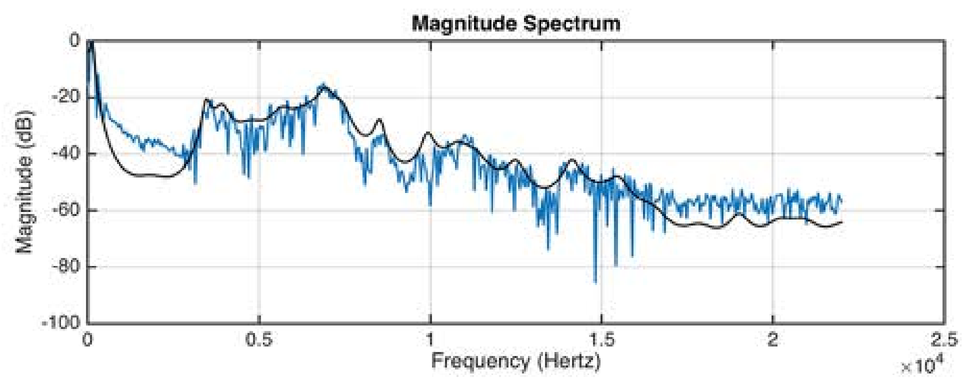

Zoo.wav
x(t)
Time (seconds)


Narrowband Spectrum
Frequency (Hz)
Time (s)


Wideband Spectrum
Frequency (Hz)
Time (s)


3D Plot of Word Signal
Power Density (dB)
Time (seconds)
Frequency (Hertz)


z.wav
x(t)
Time (seconds)


Magnitude Spectrum
Magnitude (dB)
Frequency (Hertz)

/z/

‘Zoo’

Conclusion:

This project combined many concepts learned in Digital Speech and Audio Processing and made each student realize a personal inventory of phonemes. In addition to the acoustics of the human vocal tract, the main concepts reinforced in this project were sampling, windowing, spectrographic filtering, generating linear prediction spectral envelopes, and three-dimensional plotting. The MATLAB code below was run for each word to generate and save the given plots. For higher resolution, the recordings were sampled at 44.1kHz with 24-bits per sample but to highlight the frequencies of the human vocal tract, the y-axes of the narrowband and wideband spectrographs were truncated to 8kHz.

MATLAB Code:

```
clear;
close all
file_name = 'Just.wav';          %Specify the WAV file containing the
spoken word.
pho_name = 'Capital_J.wav';      %Specify the WAV file containing the
phoneme of interest
[s, fs] = wavread(file_name);    %Read in the WAV files and sampling
rates
[ps, fsp] = wavread(pho_name);
t = 0:1/fs:(length(s)-1)/fs;     %Setup time vectors from sampling
frequency
tPhoneme = 0:1/fsp:(length(ps)-1)/fsp;
T = 0.025;                       %Set window size to 25 ms;
N = 2^round(log2(T*fs));
N_pho = 2^round(log2(T*fsp));    %Round the sizes to the closest power
of two for faster computation

figure(1);                       %Plots the waveform & spectrogram of
the word
subplot(311);
plot(t, s); title(file_name); axis tight; grid on;
ylabel('x(t)'); xlabel('Time (seconds)');

subplot(312), spectrogram(s,hamming(N),round(0.97*N),[],fs,'yaxis');
axis ([0 length(s-1)/fs 1 8000]);
title('Narrowband Spectrum');

subplot(313), spectrogram(s,hamming(N/4),round(0.9*N/4),[], fs,
'yaxis');
axis ([0 length(s-1)/fs 1 8000]);
title('Wideband Spectrum');

print -djpeg90 -r300 Just_1       %Saves the figure as a jpeg
```

```
figure(2);                          %Plots the 3-D spectrogram of the word
N2 = 2^round(log2(0.020*fs));
Rounder = round(0.75*N2);
[S,F,T,P] = spectrogram(s, hamming(N2), Rounder, [], fs, 'yaxis');
P = 20*log10(abs(P));
surf(T,F,P,'edgecolor','interp'); axis tight; title('3D Plot of Word
Signal');
xlabel('Time (seconds)'); ylabel('Frequency (Hertz)'); zlabel('Power
Density (dB)');
view(75, 20);

print -djpeg90 -r300 Just_2

figure(3);                          %Plots the phoneme and linearly
predicts its envelope
fourier_phon = fft(ps.*hamming(length(ps)), N_pho);
phon = abs(fourier_phon(1:N_pho/2));
phon = phon./max(phon);
ff = (0:fsp/N_pho:fsp/2-fsp/N_pho);

subplot(211);
plot(tPhoneme, ps); title(pho_name);
ylabel('x(t)'); xlabel('Time (seconds)');

subplot(212);
plot(ff,20*log10(phon)); grid on; hold on;
ph = fsp/1000 + 4;
[a,g] = lpc(ps.*hamming(length(ps)),ph);
lspec = freqz(g, a, ff, fsp);
plot(ff, 20*log10(abs(lspec./max(lspec))), 'k'); title('Magnitude
Spectrum');
xlabel('Frequency (Hertz)'); ylabel('Magnitude (dB)');

print -djpeg90 -r300 Just_3
```