main

October 20, 2022

[1]:
```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import os
import re
import warnings
import matplotlib.pyplot as plt
import seaborn as sns
import librosa
import librosa.display
from sklearn.preprocessing import minmax_scale
import IPython.display as ipd

plt.rcParams['figure.figsize'] = (20,8)
plt.rcParams['font.size'] = 16
sns.set_style('darkgrid')
warnings.filterwarnings("ignore")
```

1 Data Visualization and Analysis (EDA)

[18]:
```
directory = 'input/dysarthria-detection'
data = pd.read_csv("input/dysarthria-detection/torgo_data/data.csv")
data['filename'] = data['filename'].apply(lambda x: os.path.join(directory,x))
data
```

[18]:

	is_dysarthria	gender	\
0	non_dysarthria	female	
1	non_dysarthria	female	
2	non_dysarthria	female	
3	non_dysarthria	female	
4	non_dysarthria	female	
…	…	…	
1995	dysarthria	male	
1996	dysarthria	male	
1997	dysarthria	male	
1998	dysarthria	male	
1999	dysarthria	male	

```
                                                filename
0     input/dysarthria-detection/torgo_data/non_dysa…
1     input/dysarthria-detection/torgo_data/non_dysa…
2     input/dysarthria-detection/torgo_data/non_dysa…
3     input/dysarthria-detection/torgo_data/non_dysa…
4     input/dysarthria-detection/torgo_data/non_dysa…
…                                                   …
1995  input/dysarthria-detection/torgo_data/dysarthr…
1996  input/dysarthria-detection/torgo_data/dysarthr…
1997  input/dysarthria-detection/torgo_data/dysarthr…
1998  input/dysarthria-detection/torgo_data/dysarthr…
1999  input/dysarthria-detection/torgo_data/dysarthr…

[2000 rows x 3 columns]
```

2 Functions of Visualization: Plots and Features

Spectrogram, Waveplot, Zero Crossing Rate, Spectral Rolloff, MFCCs, Mel Spectrogram, and Spectral Centroids

[19]:
```
def show_waveplot(audio_path,label,gender):
    x , sr = librosa.load(audio_path)
    plt.figure(figsize=(20, 6))
    librosa.display.waveshow(x, sr=sr)
    plt.title(f"Waveplot: of Class: {label}, Gender: {gender}")

def show_spectrogram(audio_path,label,gender):
    x , sr = librosa.load(audio_path)
    X = librosa.stft(x)
    Xdb = librosa.amplitude_to_db(abs(X))
    plt.figure(figsize=(20,6))
    librosa.display.specshow(Xdb, sr=sr, x_axis='time',␣
 ↪y_axis='hz',cmap='plasma')
    plt.colorbar()
    plt.title(f"Spectrogram of Class: {label}, Gender: {gender}")

def show_zcr(audio_path,label,gender):
    x , sr = librosa.load(audio_path)
    zero_crossings = librosa.zero_crossings(x)
    print("Sum of zero crossing ", zero_crossings.sum())
    plt.figure(figsize=(20, 5))
    plt.title(f'Zero Crossing Rate of Class: {label}, Gender: {gender}')
    zcrs = librosa.feature.zero_crossing_rate(x)
    plt.plot(zcrs[0])
    plt.show()
```

```
def normalize(x, axis=0):
    return minmax_scale(x, axis=axis)

def show_spectral_centroids(audio_path,label,gender):
    plt.figure(figsize=(20, 5))
    plt.title(f'Spectral Centroids of Class: {label}, Gender: {gender}')
    x , sr = librosa.load(audio_path)
    spectral_centroids = librosa.feature.spectral_centroid(x, sr=sr)[0]
    frames = range(len(spectral_centroids))
    t = librosa.frames_to_time(frames)
    librosa.display.waveshow(x, sr=sr, alpha=0.4)
    plt.plot(t, normalize(spectral_centroids), color='r')
    plt.show()

def show_spectral_rolloff(audio_path,label,gender):
    plt.figure(figsize=(20, 5))
    plt.title(f'Spectral Rolloff of Class: {label}, Gender: {gender}')
    x , sr = librosa.load(audio_path)
    spectral_rolloff = librosa.feature.spectral_rolloff(x, sr=sr,␣
 ↪roll_percent=0.01)[0]
    frames = range(len(spectral_rolloff))
    t = librosa.frames_to_time(frames)
    librosa.display.waveshow(x, sr=sr, alpha=0.4)
    plt.plot(t, normalize(spectral_rolloff), color='r')
    plt.show()

def show_mfccs(audio_path,label,gender):
    plt.figure(figsize=(20, 6))
    plt.title(f'MFCC of Class: {label}, Gender: {gender}')
    x , sr = librosa.load(audio_path)
    mfccs = librosa.feature.mfcc(y=x, sr=sr)
    librosa.display.specshow(mfccs, sr=sr, x_axis='time',cmap='plasma')
    plt.show()

def show_melspectro(audio_path,label,gender):
    plt.figure(figsize=(20, 6))
    plt.title(f'Mel Spectro of Class: {label}, Gender: {gender}')
    x , sr = librosa.load(audio_path)
    melspectro = librosa.feature.melspectrogram(y=x, sr=sr)
    librosa.display.specshow(melspectro, sr=sr, x_axis='time',cmap='plasma')
    plt.show()
```

3 Audio Samples

[20]:
```
data.sample(frac=1).reset_index(drop=True,inplace=True)
male_dysarthric = data[(data['gender']=='male') &␣
 ↪(data['is_dysarthria']=='dysarthria')].sample(1)
female_dysarthric = data[(data['gender']=='female') &␣
 ↪(data['is_dysarthria']=='dysarthria')].sample(1)
male_ndysarthric = data[(data['gender']=='male') &␣
 ↪(data['is_dysarthria']=='non_dysarthria')].sample(1)
female_ndysarthric = data[(data['gender']=='female') &␣
 ↪(data['is_dysarthria']=='non_dysarthria')].sample(1)
```

4 Male and Dysarthric

[21]:
```
ipd.Audio(male_dysarthric.iloc[0].filename)
```

[21]:
```
<IPython.lib.display.Audio object>
```

5 Female and Dysarthric

[22]:
```
ipd.Audio(female_dysarthric.iloc[0].filename)
```

[22]:
```
<IPython.lib.display.Audio object>
```

6 Male and Non Dysarthric

[23]:
```
ipd.Audio(male_ndysarthric.iloc[0].filename)
```

[23]:
```
<IPython.lib.display.Audio object>
```

7 Female and Non Dysarthric

[24]:
```
ipd.Audio(female_ndysarthric.iloc[0].filename)
```

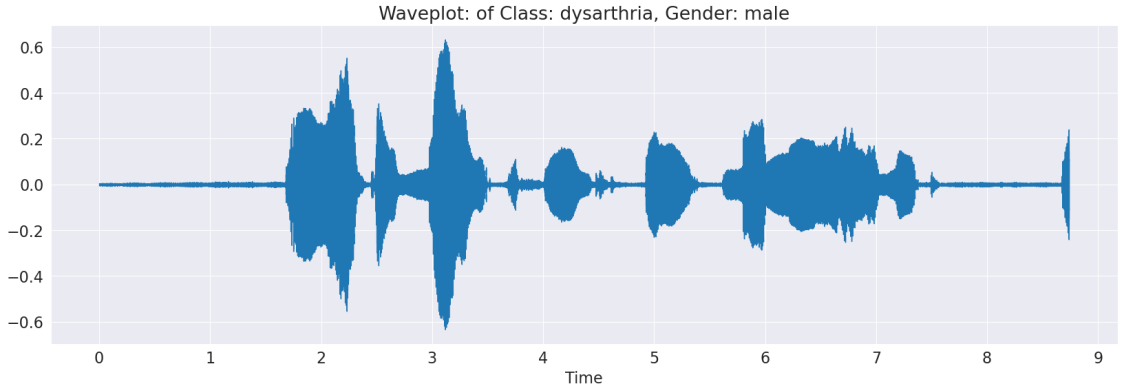
[24]:
```
<IPython.lib.display.Audio object>
```

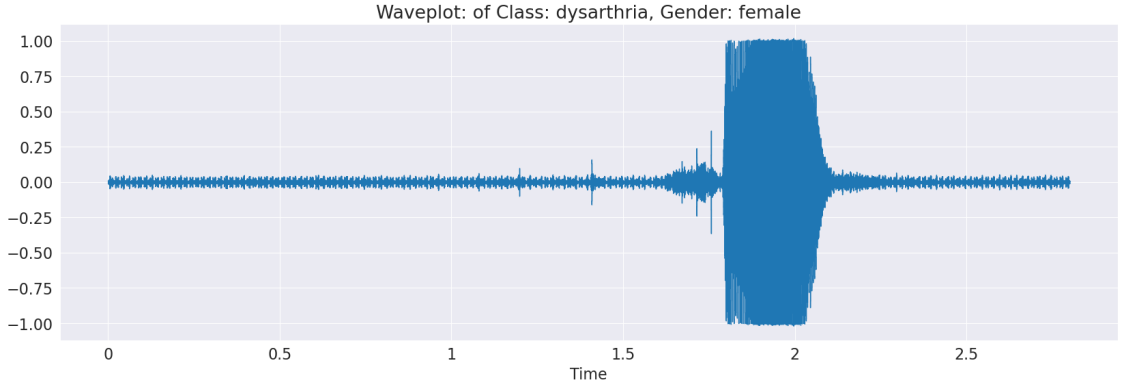
8 wavplot

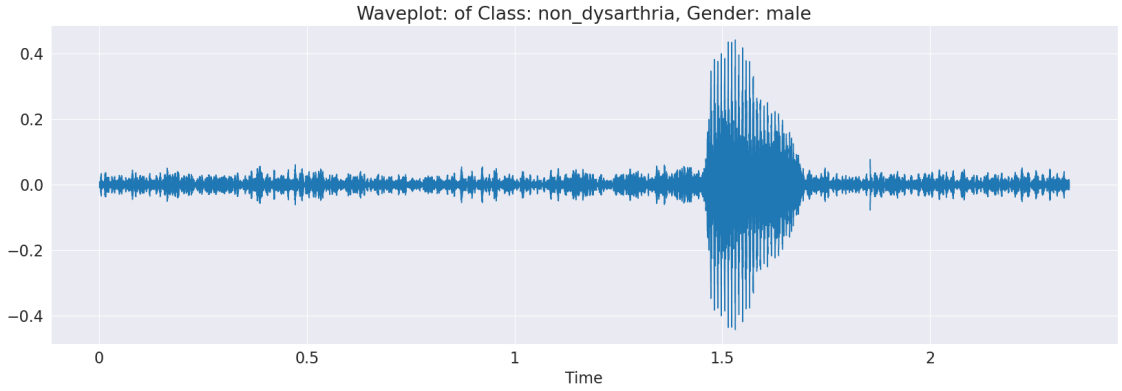
- The waveplot for the dysarthric samples shows that the one male sample used here has slurred speech, while the waveplot for the female sample shows that the speech is fairly rapid and challenging to comprehend since the words overlap, as shown by the box-like patterns emerging in the waveplot.

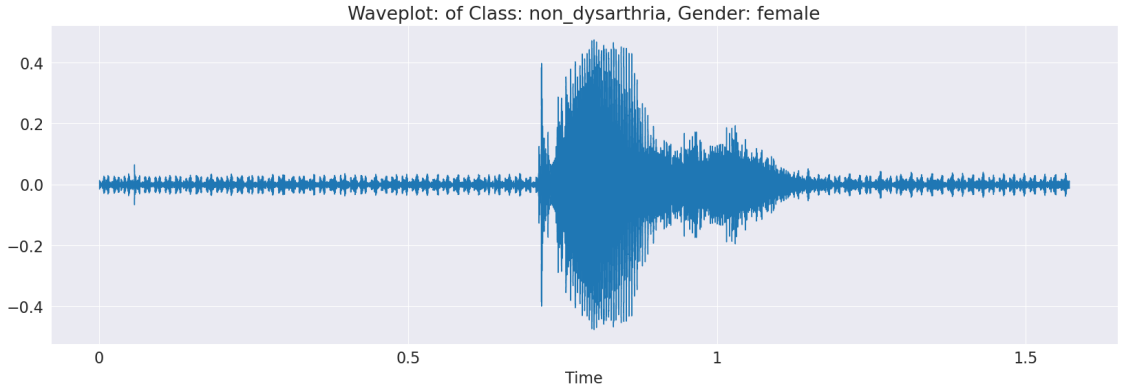
- The non-dysarthric samples exhibit a regular waveplot, which suggests that the speech is regularly timed.

[25]:
```
show_waveplot(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_waveplot(female_dysarthric.iloc[0].filename, female_dysarthric.iloc[0].
 ↪is_dysarthria, female_dysarthric.iloc[0].gender)
show_waveplot(male_ndysarthric.iloc[0].filename, male_ndysarthric.iloc[0].
 ↪is_dysarthria, male_ndysarthric.iloc[0].gender)
show_waveplot(female_ndysarthric.iloc[0].filename, female_ndysarthric.iloc[0].
 ↪is_dysarthria, female_ndysarthric.iloc[0].gender)
```





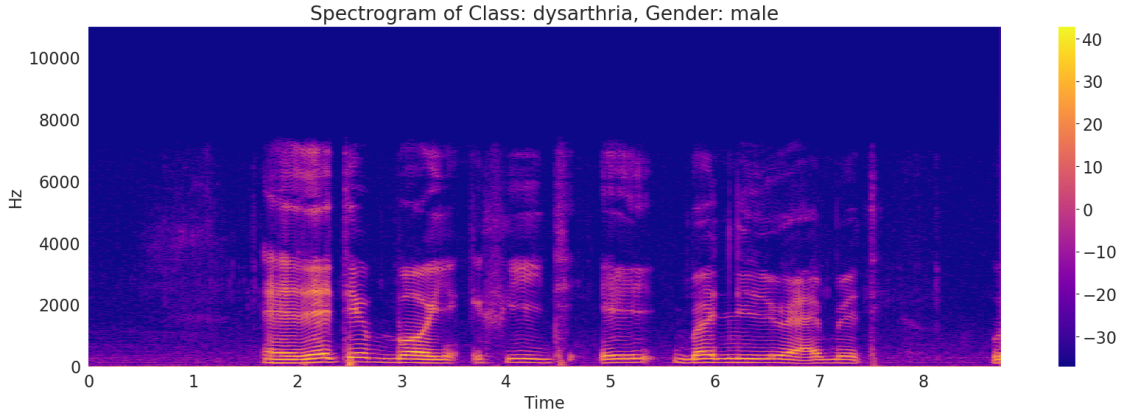


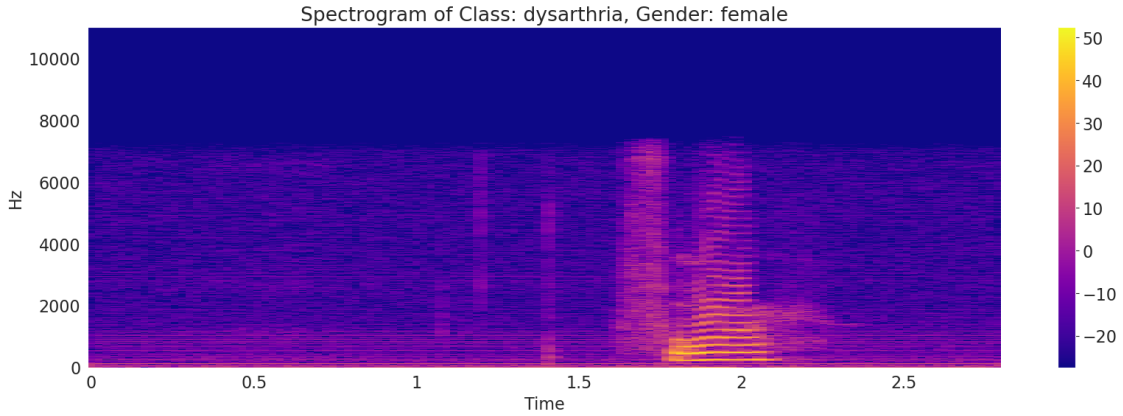


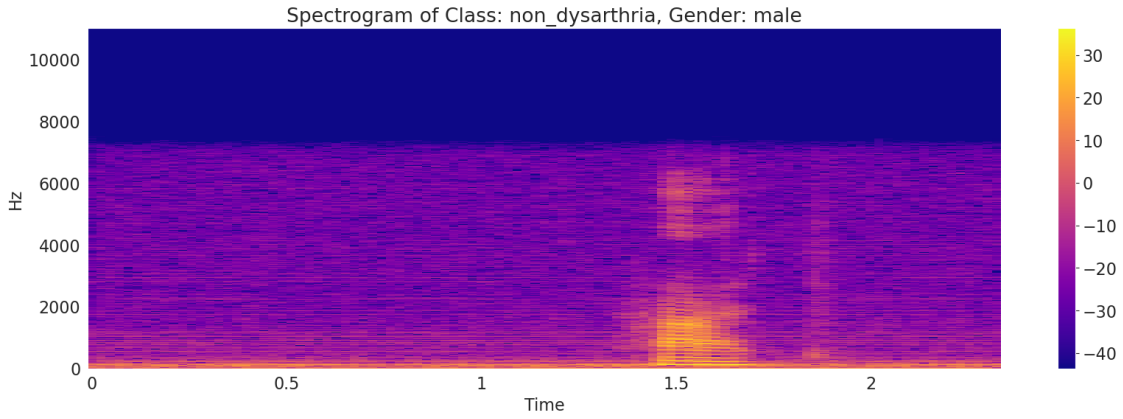
9 Spectrograms

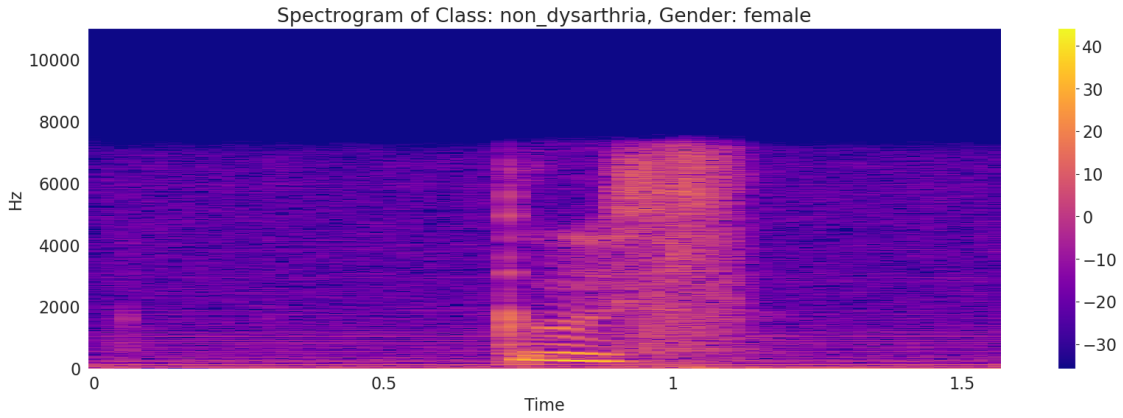
- It is evident from the dysarthric samples that the energy magnitudes of the frequencies are more evenly distributed throughout time, a sign of sluggish, slurred speech or because the words are being said more quickly and overlapping one another. Similar patterns could also be seen in monotonous, dysarthric speech.
- The energy magnitudes for the non-dysarthric samples may be seen to be more restricted to just the areas when they have uttered something, which is regular paced.

[26]:
```
show_spectrogram(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_spectrogram(female_dysarthric.iloc[0].filename, female_dysarthric.iloc[0].
 ↪is_dysarthria, female_dysarthric.iloc[0].gender)
show_spectrogram(male_ndysarthric.iloc[0].filename, male_ndysarthric.iloc[0].
 ↪is_dysarthria, male_ndysarthric.iloc[0].gender)
show_spectrogram(female_ndysarthric.iloc[0].filename, female_ndysarthric.
 ↪iloc[0].is_dysarthria, female_ndysarthric.iloc[0].gender)
```


Spectrogram of Class: dysarthria, Gender: male
Hz
Time
10000
8000
6000
4000
2000
0
0
1
2
3
4
5
6
7
8
40
30
20
10
0
−10
−20
−30


Spectrogram of Class: dysarthria, Gender: female
Hz
Time
10000
8000
6000
4000
2000
0
0
0.5
1
1.5
2
2.5
50
40
30
20
10
0
−10
−20


Spectrogram of Class: non_dysarthria, Gender: male
Hz
Time
10000
8000
6000
4000
2000
0
0
0.5
1
1.5
2
30
20
10
0
−10
−20
−30
−40

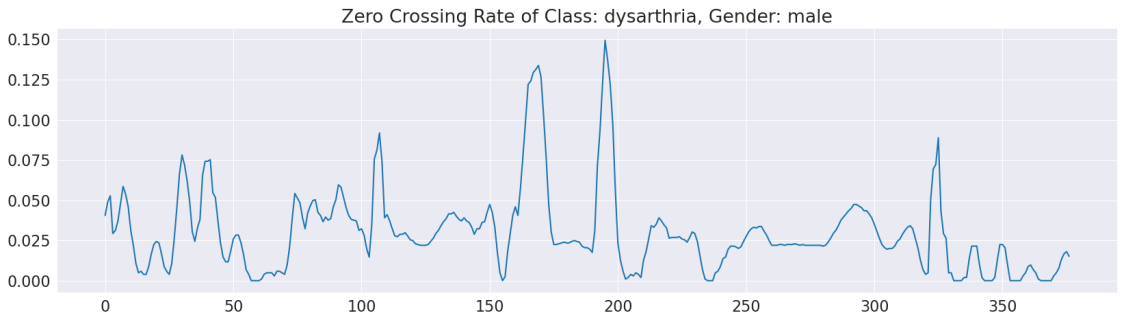


10 Zero Crossing Rate

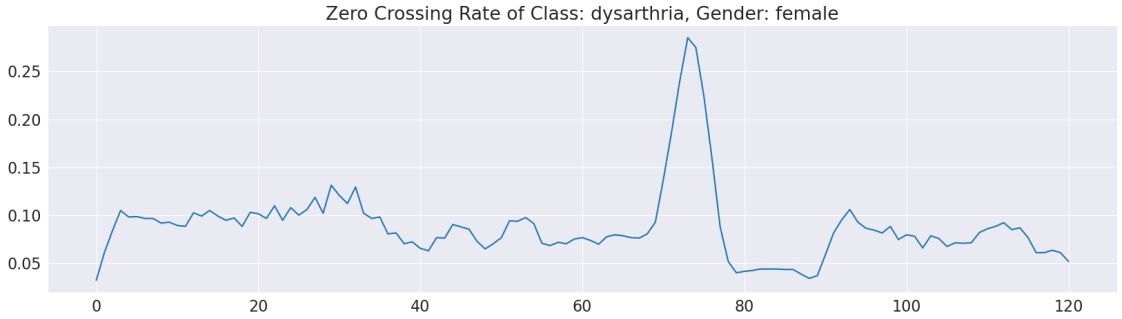
- For dysarthric samples, it can be seen that there were times when some of their words were substantially louder in the audio itself (for the female dysarthric case, for example), leading to several peaks of zero crossing rate in those situations. Their inability to control their tongue and other mouth-related muscles when speaking and their inability to purposefully emphasis words may be the cause of the abrupt loudness of some words.
- In scenarios where the speaker's voice can be heard, the zero crossing rate for non-dysarthric individuals is significant.

[27]:
```
show_zcr(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_zcr(female_dysarthric.iloc[0].filename, female_dysarthric.iloc[0].
 ↪is_dysarthria, female_dysarthric.iloc[0].gender)
show_zcr(male_ndysarthric.iloc[0].filename, male_ndysarthric.iloc[0].
 ↪is_dysarthria, male_ndysarthric.iloc[0].gender)
show_zcr(female_ndysarthric.iloc[0].filename, female_ndysarthric.iloc[0].
 ↪is_dysarthria, female_ndysarthric.iloc[0].gender)
```

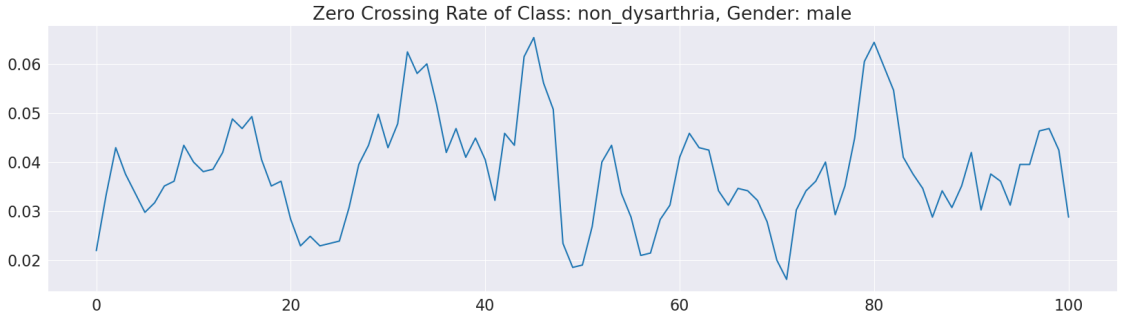
```
Sum of zero crossing  5804
```



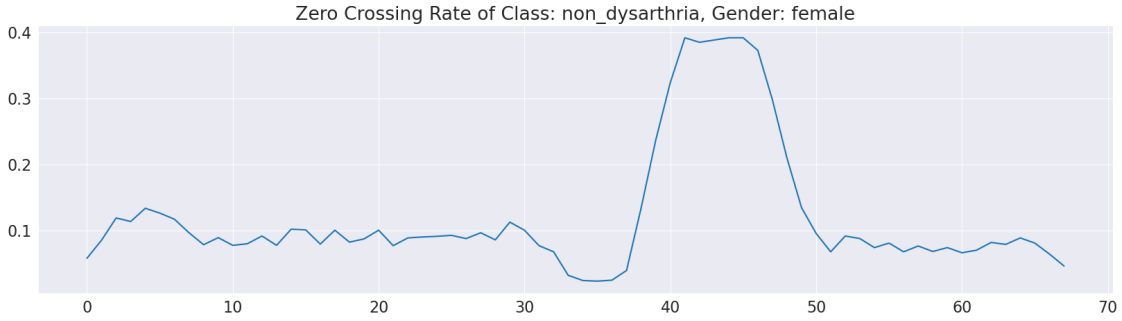
```
Sum of zero crossing  5512
```




```
Sum of zero crossing  1992
```




```
Sum of zero crossing  4240
```

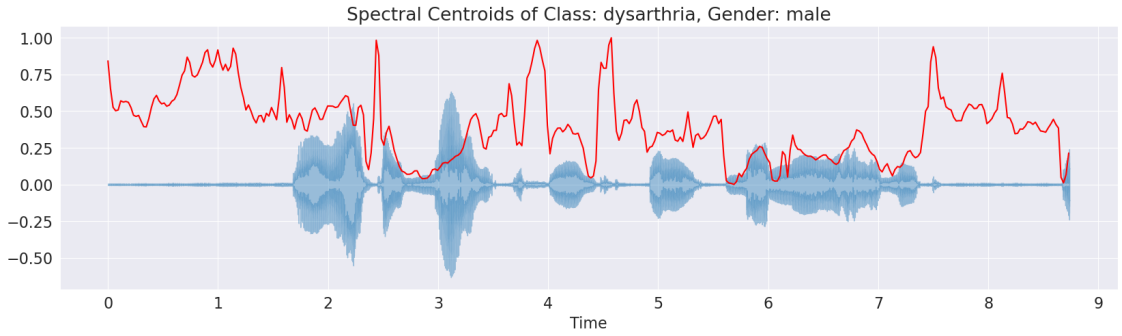


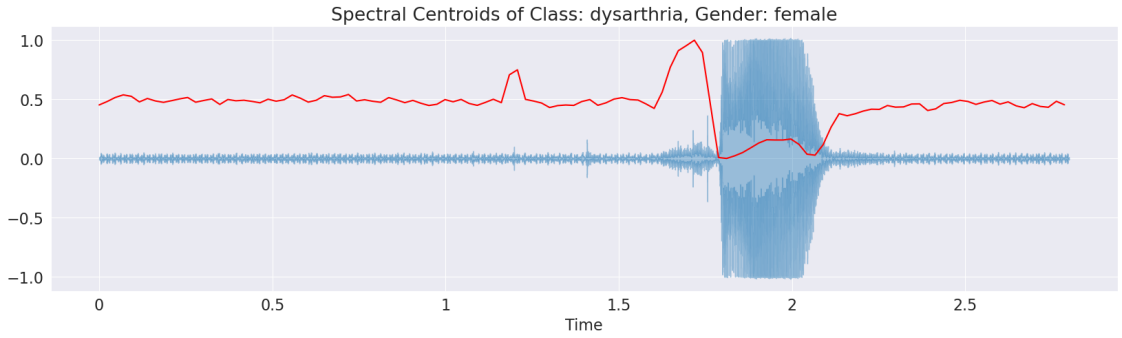
11 Spectral Centroid

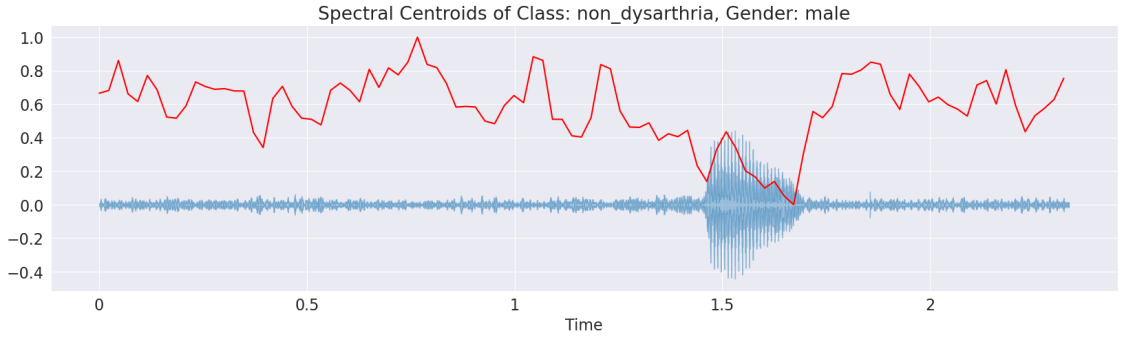
- The location of the spectrum's centre of mass is indicated by spectral centroids.
- The centroids are usually high when there is no speech and become low in those instances of times where speech is present in the audio

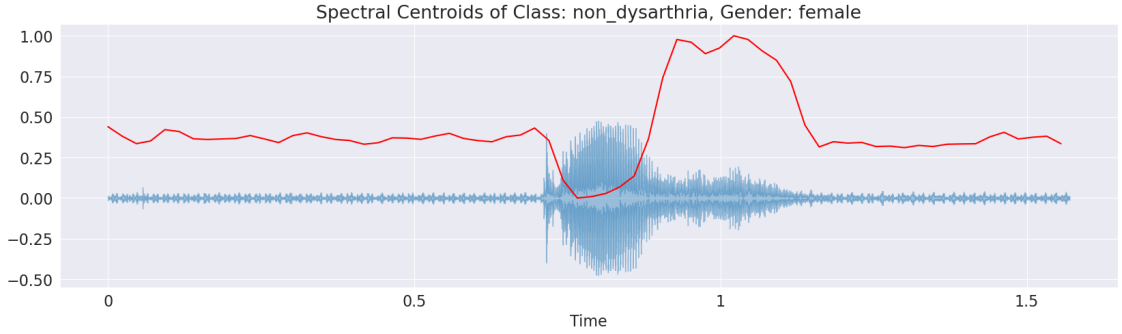
- For dysarthric samples, it can be observed that when the speech is slurred, the centroid values initially decrease and then quickly increase after brief pauses.
- When speech is present in the audio, the centroids for non-dysarthric samples temporarily drop, but as soon as speech is absent, they quickly rise again.

[28]:
```
show_spectral_centroids(male_dysarthric.iloc[0].filename, male_dysarthric.
 ↪iloc[0].is_dysarthria, male_dysarthric.iloc[0].gender)
show_spectral_centroids(female_dysarthric.iloc[0].filename, female_dysarthric.
 ↪iloc[0].is_dysarthria, female_dysarthric.iloc[0].gender)
show_spectral_centroids(male_ndysarthric.iloc[0].filename, male_ndysarthric.
 ↪iloc[0].is_dysarthria, male_ndysarthric.iloc[0].gender)
show_spectral_centroids(female_ndysarthric.iloc[0].filename, female_ndysarthric.
 ↪iloc[0].is_dysarthria, female_ndysarthric.iloc[0].gender)
```





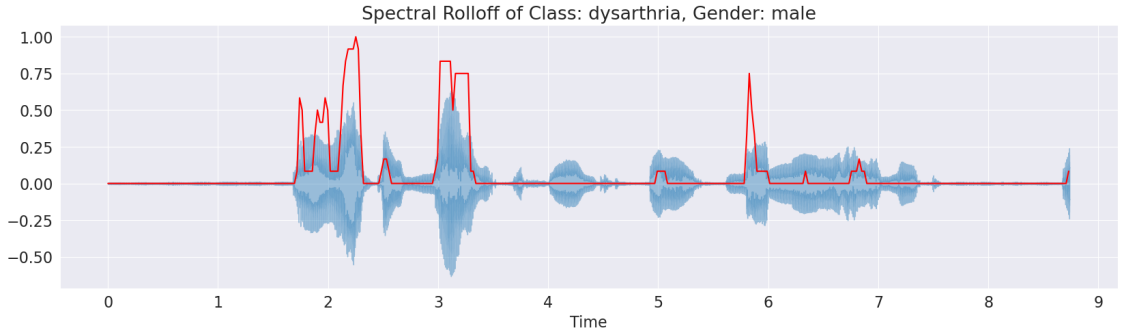


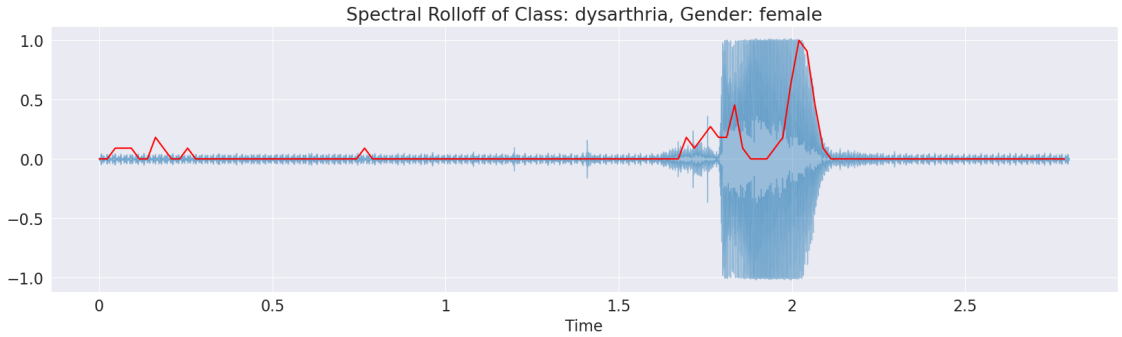


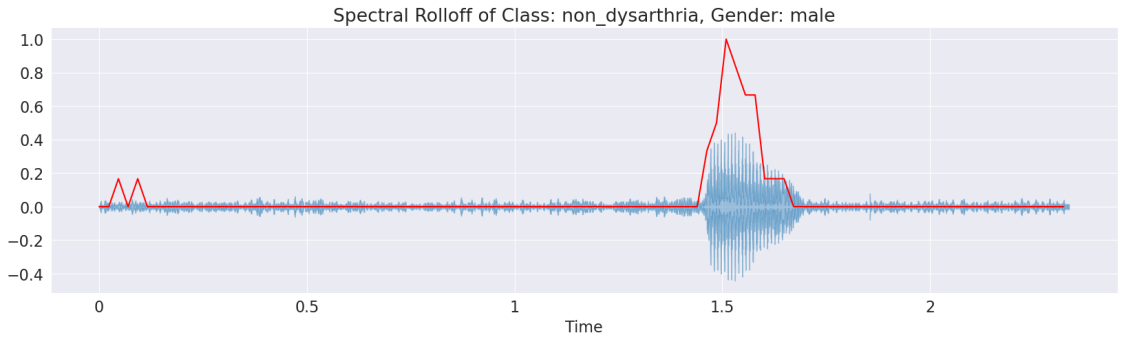
12 Spectral Rolloff

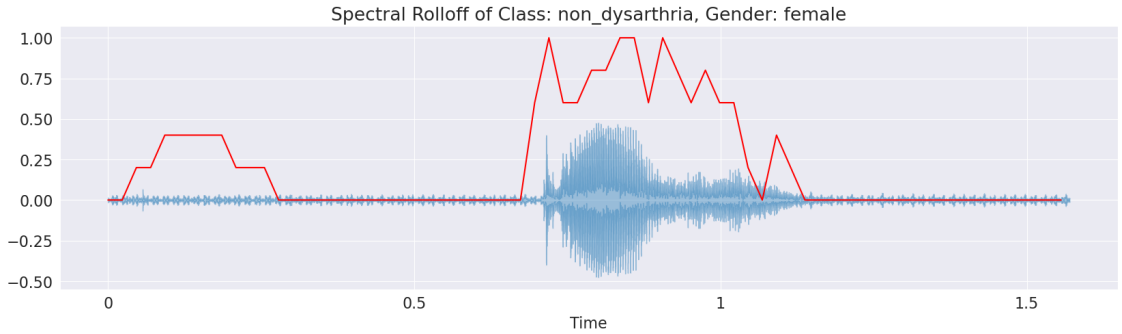
- The frequency below which a certain percentage (librosa: 85%) of the magnitude distribution is concentrated is shown by spectral rolloffs.
- The 85% coverage below the rolloff frequency can be seen surrounding the speech present in the audio signal, which is consistent with our earlier observations in the spectrograms that the energy magnitudes of the frequencies are more dispersed in the case of dysarthric samples.
- For non-dysarthric samples, we can observe that the 85% coverage includes both the speech and some other audio stutters and noise.

[29]:
```
show_spectral_rolloff(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_spectral_rolloff(female_dysarthric.iloc[0].filename, female_dysarthric.
 ↪iloc[0].is_dysarthria, female_dysarthric.iloc[0].gender)
show_spectral_rolloff(male_ndysarthric.iloc[0].filename, male_ndysarthric.
 ↪iloc[0].is_dysarthria, male_ndysarthric.iloc[0].gender)
show_spectral_rolloff(female_ndysarthric.iloc[0].filename, female_ndysarthric.
 ↪iloc[0].is_dysarthria, female_ndysarthric.iloc[0].gender)
```


Spectral Rolloff of Class: dysarthria, Gender: male
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
0
1
2
3
4
5
6
7
8
9
Time


Spectral Rolloff of Class: dysarthria, Gender: female
1.0
0.5
0.0
−0.5
−1.0
0
0.5
1
1.5
2
2.5
Time

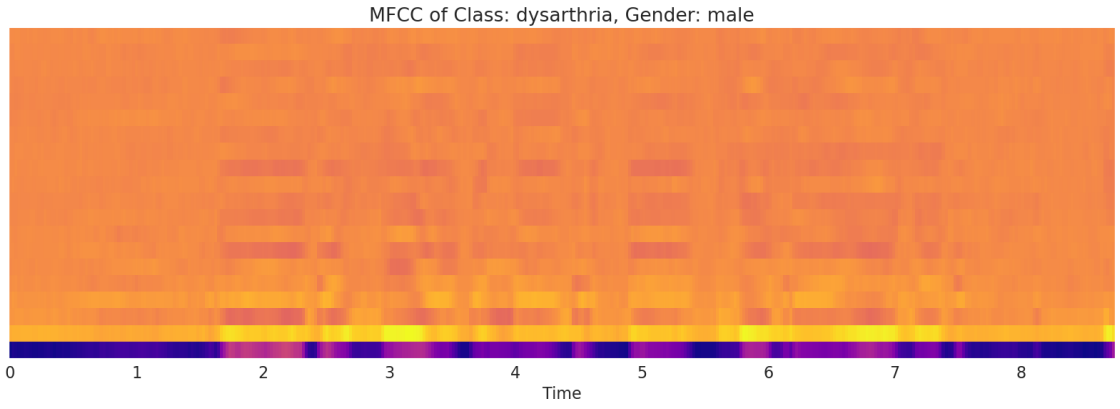

Spectral Rolloff of Class: non_dysarthria, Gender: male
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0
0.5
1
1.5
2
Time

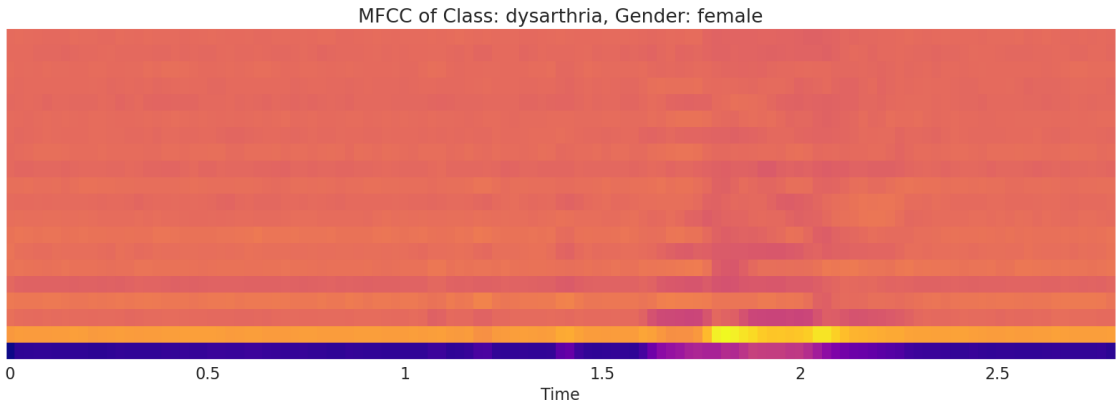

Spectral Rolloff of Class: non_dysarthria, Gender: female
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
0
0.5
1
1.5
Time

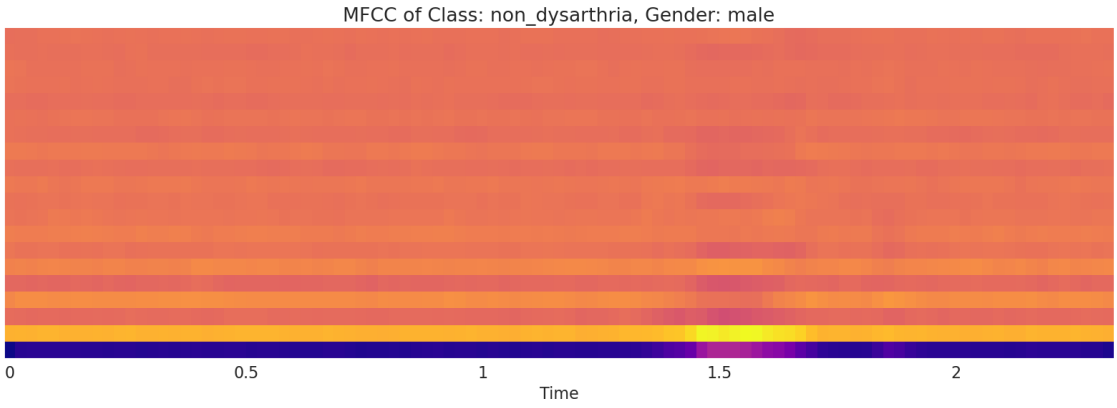
13 MFCCs

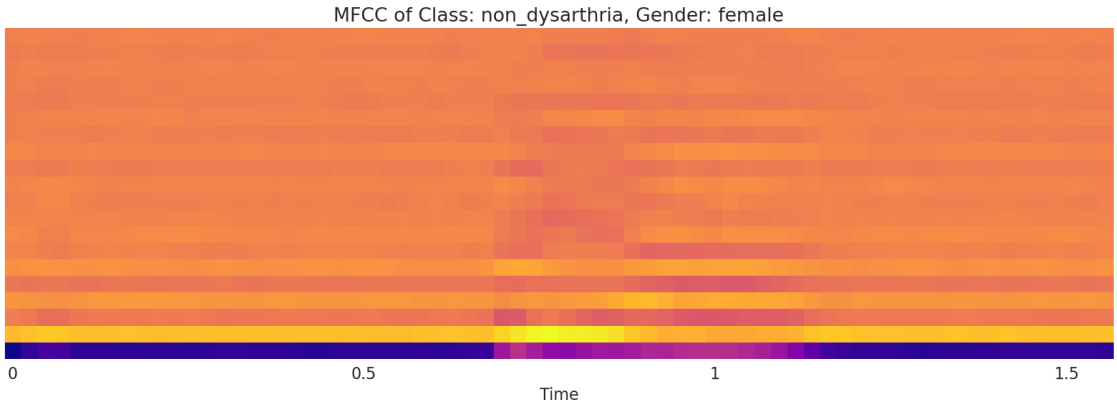
- The MFCC visualisation for dysarthric samples includes significantly darker and more concentrated pixels to show the presence of slurs in their speech.
- Non-dysarthric people experience it as being less concentrated and more dispersed.

[30]:
```
show_mfccs(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_mfccs(female_dysarthric.iloc[0].filename, female_dysarthric.iloc[0].
 ↪is_dysarthria, female_dysarthric.iloc[0].gender)
show_mfccs(male_ndysarthric.iloc[0].filename, male_ndysarthric.iloc[0].
 ↪is_dysarthria, male_ndysarthric.iloc[0].gender)
show_mfccs(female_ndysarthric.iloc[0].filename, female_ndysarthric.iloc[0].
 ↪is_dysarthria, female_ndysarthric.iloc[0].gender)
```





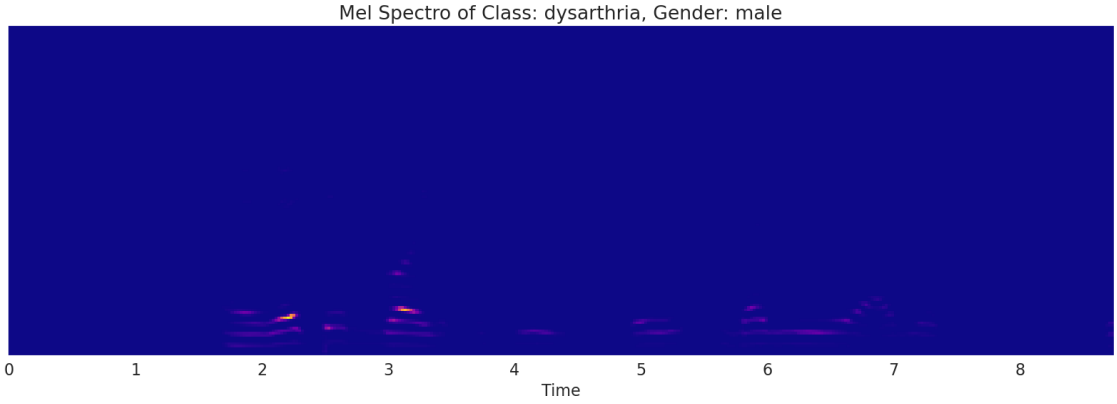


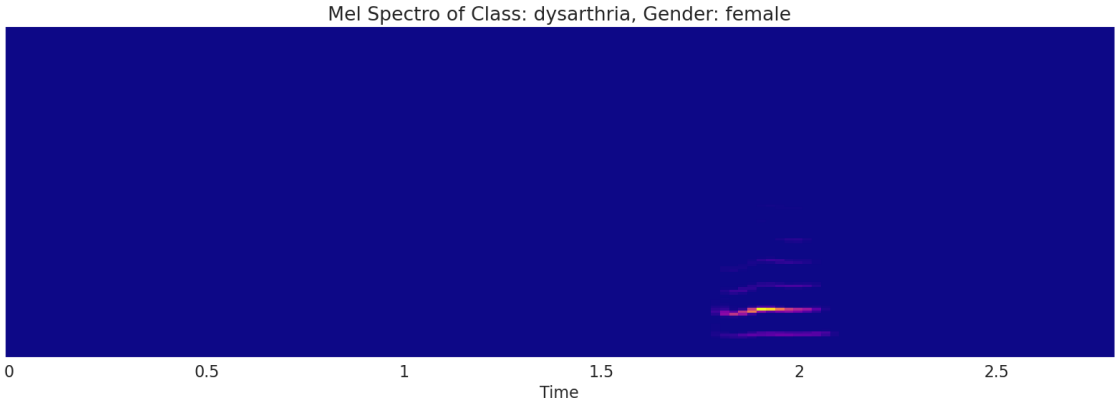


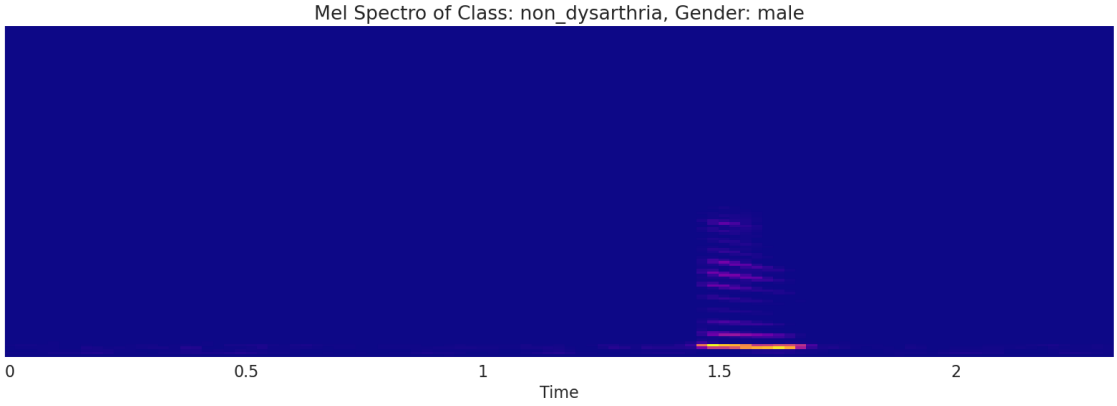
14 Mel Spectrogram

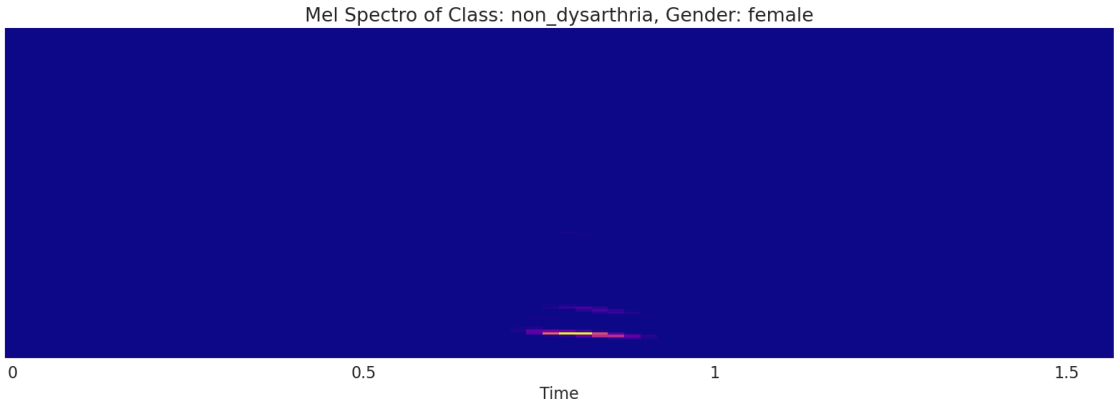
- The Spectrogram converted to Mel scale is showing the magnitudes that are more perceptual to the human ears generally
- Patterns in the dysarthric samples' mel spectrograms are slightly suggestive of the speech slurs.

[31]:
```
show_melspectro(male_dysarthric.iloc[0].filename, male_dysarthric.iloc[0].
 ↪is_dysarthria, male_dysarthric.iloc[0].gender)
show_melspectro(female_dysarthric.iloc[0].filename, female_dysarthric.iloc[0].
 ↪is_dysarthria, female_dysarthric.iloc[0].gender)
show_melspectro(male_ndysarthric.iloc[0].filename, male_ndysarthric.iloc[0].
 ↪is_dysarthria, male_ndysarthric.iloc[0].gender)
show_melspectro(female_ndysarthric.iloc[0].filename, female_ndysarthric.iloc[0].
 ↪is_dysarthria, female_ndysarthric.iloc[0].gender)
```


Mel Spectro of Class: dysarthria, Gender: male
0
1
2
3
4
5
6
7
8
Time


Mel Spectro of Class: dysarthria, Gender: female
0
0.5
1
1.5
2
2.5
Time


Mel Spectro of Class: non_dysarthria, Gender: male
0
0.5
1
1.5
2
Time


Mel Spectro of Class: non_dysarthria, Gender: female
0
0.5
1
1.5
Time

[ ]: