

Trabalho 1: ENG1400

27/05/2020

Lecturer: Marco Antônio Grivet

Scribe: Rafael Vilela

1 Introdução

O trabalho visa evidenciar a diferença da voz de duas pessoas através da análise do espectro de frequência e do espectrograma dos dois sinais. Foi utilizado o software Audacity. As vozes foram capturadas pelo microfone de um telefone no modo mono, depois convertidas no formato mp3, adequado para o software.

Assim como visto em sala de aula, por vezes desejamos modificar o sinal da função do tempo (t) para em função da frequência (f) ou (w). Quando gravado no celular e em outros meios normalmente vemos a forma de onda pelo tempo. Pode-se aplicar a Transformada de Fourier ou nos casos computacionais a FFT (Fast Fourier Transform) a fim de modificar para o domínio da frequência.

2 Diferença entre as vozes

Projeto I - ENG1400

Rafael Vilela Santa Rosa

Rafael	0-10	0-2	2.5-4	4.5-6	6.5-8	8.5-10
Pico 1	121Hz	127Hz	122Hz	110Hz	114Hz	111Hz
Pico 2	255Hz	276Hz	248Hz	229Hz	245Hz	238Hz
Pico 3	400Hz	408Hz	378Hz	460Hz	380Hz	471Hz
Carolina	0-10	0-2	2.5-4	4.5-6	6.5-8	8.5-10
Pico 1	193Hz	193Hz	198Hz	181Hz	184Hz	193Hz
Pico 2	406Hz	413Hz	403Hz	537Hz	386Hz	392Hz
Pico 3	552Hz	835Hz	560Hz	711Hz	582Hz	572Hz

(1)

A frase utilizada foi um palíndromo, uma palavra ou frase que pode ser escrita da mesma forma de trás pra frente: "Socorram-me, subi no ônibus em Marrocos".

Enquanto os 3 formantes da voz masculina variaram entre 120 e 400Hz, sendo o segundo por volta de 250Hz, os da voz feminina variaram entre 190 e 570Hz, sendo o segundo por volta de

400Hz. A voz variou em uma faixa maior de frequências (como é possível ver também no espectrograma), esse trecho é somente o que contém as 3 frequências com maior intensidade.

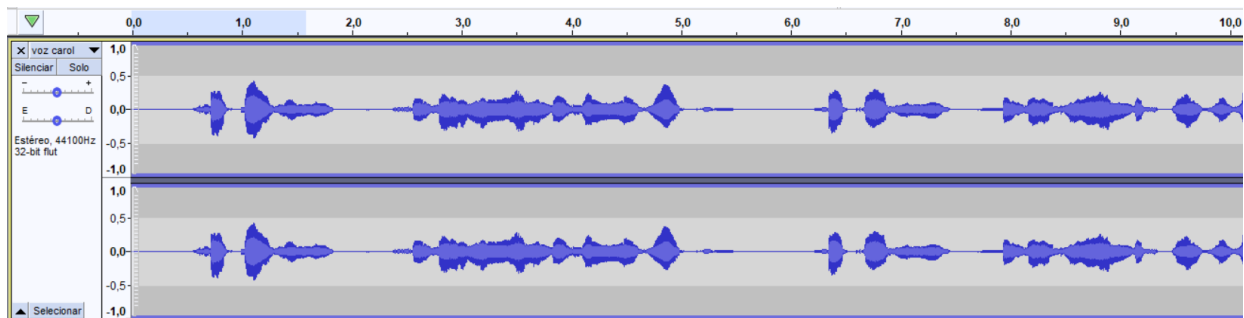
A voz feminina teve maior variação de pico de amplitude no terceiro formante, chegando até 835Hz, enquanto a masculina foi mais constante, não variando muito além de 100Hz da média do terceiro formante.

As frequências analisadas nos 5 trechos disjuntos foram bem próximas às analisadas entre 0-10s, sendo o terceiro formante a que mais variou em relação a média. É interessante notar também que o terceiro formante masculino é parecido com o segundo formante feminino.

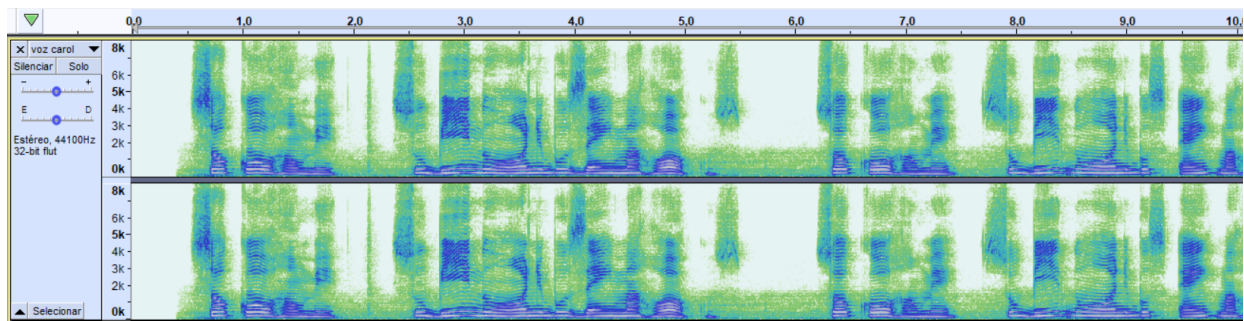
2.1 Voz feminina (Carolina)

Como dito anteriormente, a voz é composta por várias frequências e o espectrograma fornece essa visão relacionando as duas variáveis tempo e frequência. O que mais difere entre os dois sinais são as faixas de frequências e intensidade, mas - como é possível ver na primeira figura - os sons são emitidos quase na mesma cadência.

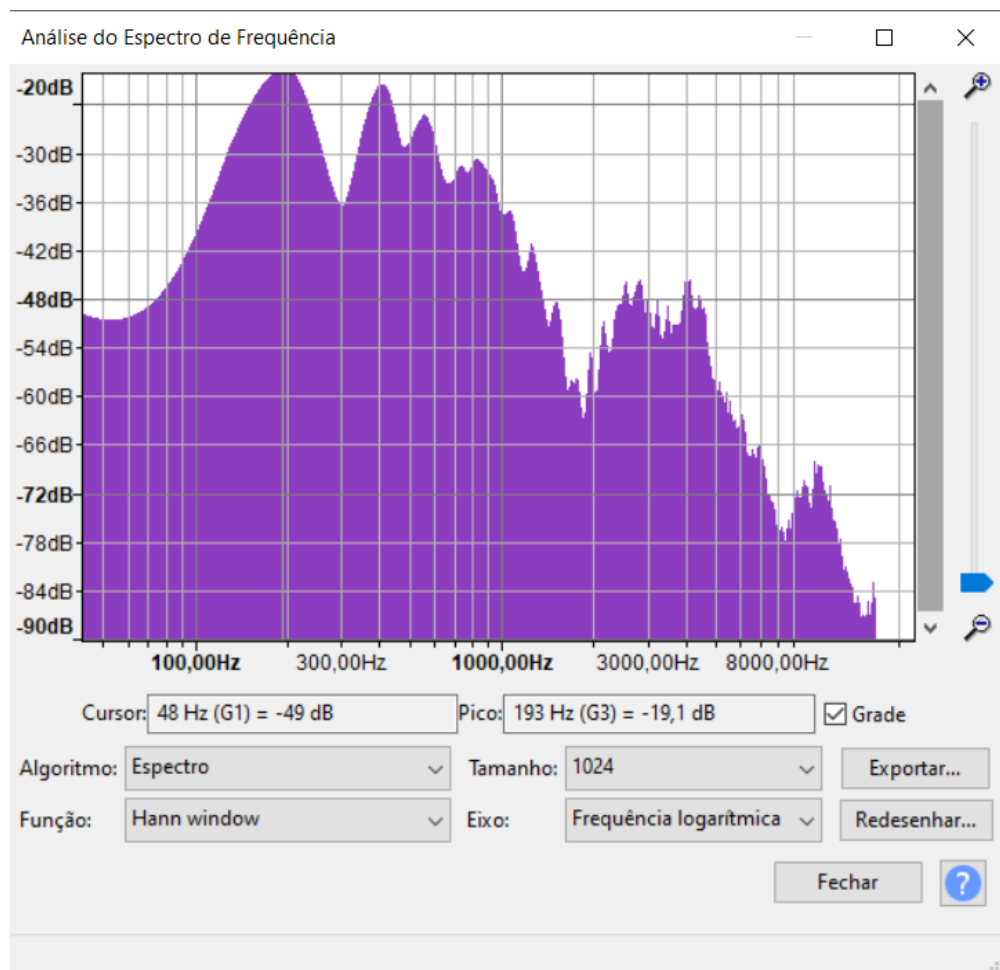
A primeira imagem é o sinal da voz pelo tempo, a segunda o espectrograma e a terceira o espectro de frequência. O espectrograma fornece uma representação da amplitude do sinal pela frequência e tempo. O eixo horizontal representa o tempo, vertical a frequência e o "z" (coloração) a amplitude. Verde representa baixa intensidade, azul alta e branco mais alta ainda.



(2)



(3)

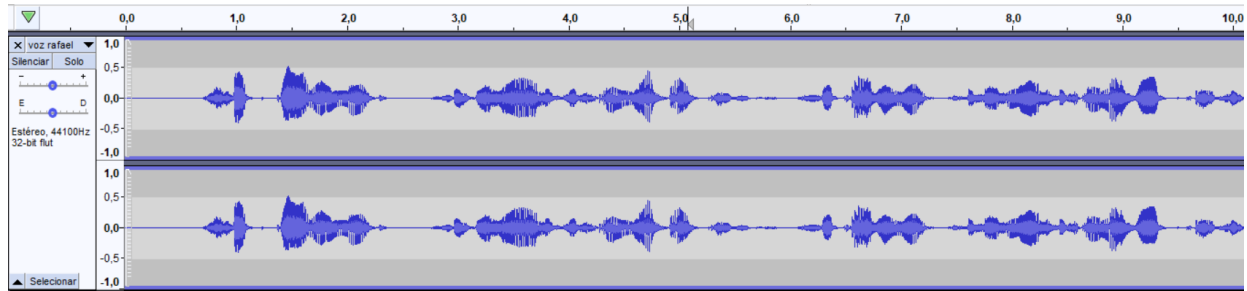


(4)

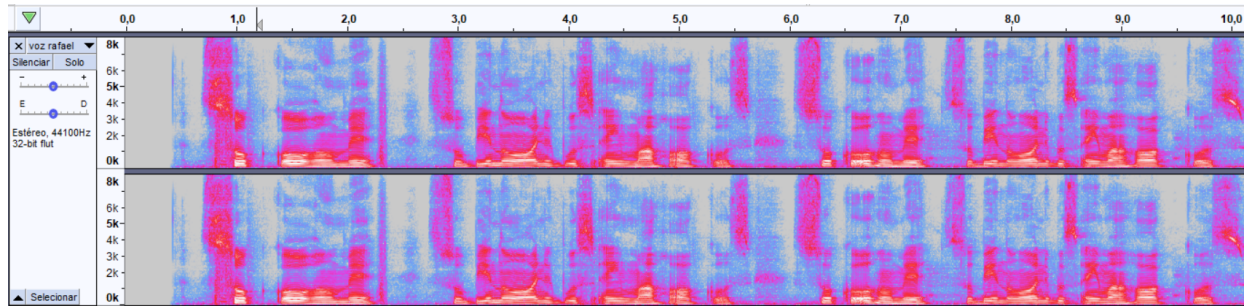
2.2 Voz masculina (Rafael)

A primeira imagem é o sinal da voz pelo tempo, a segunda o espectrograma e a terceira o espectro de frequência. O espectrograma fornece uma representação da amplitude do sinal pela frequência

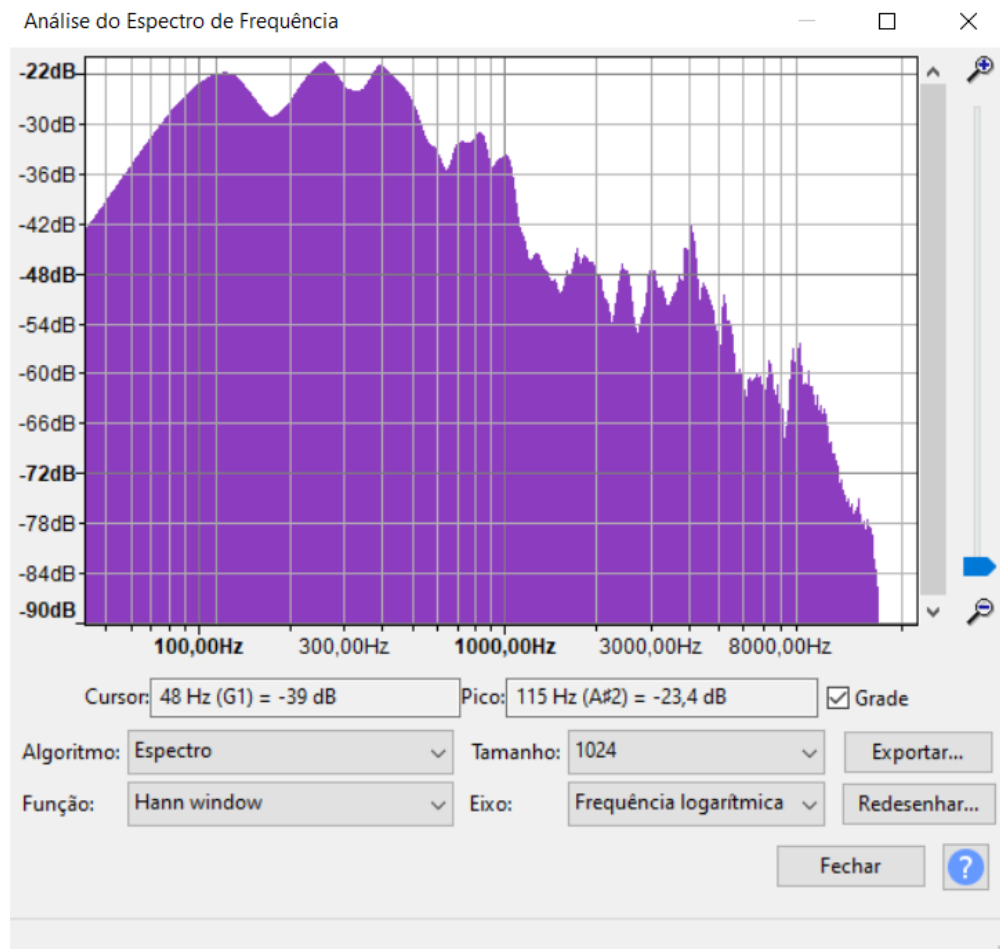
e tempo. O eixo horizontal representa o tempo, vertical a frequência e o eixo "z" (coloração) a amplitude. Azul representa baixa intensidade, vermelho alta e branco mais alta ainda.



(5)



(6)



(7)

Com o espectrograma podemos verificar que ao redor das frequências dos 3 formantes há mais intensidade/concentração de energia. Cada pico de frequência analisado é um formante. A diferença entre os dois espectrogramas representa a diferença de timbre entre as duas vozes, provocadas por - entre outros fatores - a geometria da boca e garganta distintas entre as duas pessoas analisadas.

References

- [1] M. G. Côrtes. *Análise comparativa dos parâmetros espectrográficos da voz antes e depois da fonoterapia*. 2007.
- [2] R. W. S. Lawrence Rabiner. *Digital Processing of Speech Signals*. Prentice Hall, 1978.
- [3] V. K. Madisetti. *The Digital Signal Processing Handbook*. CRC Press, 2010.

[4] A. OPPENHEIM, Alan; WILLSKY. *Sinais e Sistemas*. Pearson, 2010.

[2] [3] [4] [1]