

PRÁCTICA 2

Alba Márquez Rodríguez

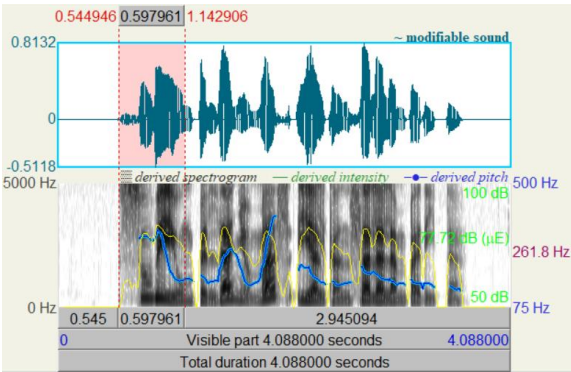


Ingeniería Informática, Universidad de Huelva
Procesamiento del Habla, Visión e Interacción Multimodal

Actividad 1: Intenisdad y nivel tonal en Praat

Fragmento seleccionado: “Subí un”:

El amarillo representa la intensidad mientras el azul del espectrograma representa el Pitch:



Intensity List:

Time_s	Intensity_dB
0.545333	46.221892
0.556000	49.456461
0.566667	53.438007
0.577333	56.522421
0.588000	58.696427
0.598667	60.003357
0.609333	60.726442
0.620000	61.326561
0.630667	61.259839
0.641333	60.386091
0.652000	58.571984
0.662667	57.825930
0.673333	58.592228
0.684000	59.218618
0.694667	59.258461
0.705333	58.479132
0.716000	58.393518
0.726667	64.021900
0.737333	72.252453
0.748000	77.103830
0.758667	79.148776
0.769333	79.778952

77.71824047480882 dB (mean-energy intensity in SELECTION)

Pitch list:

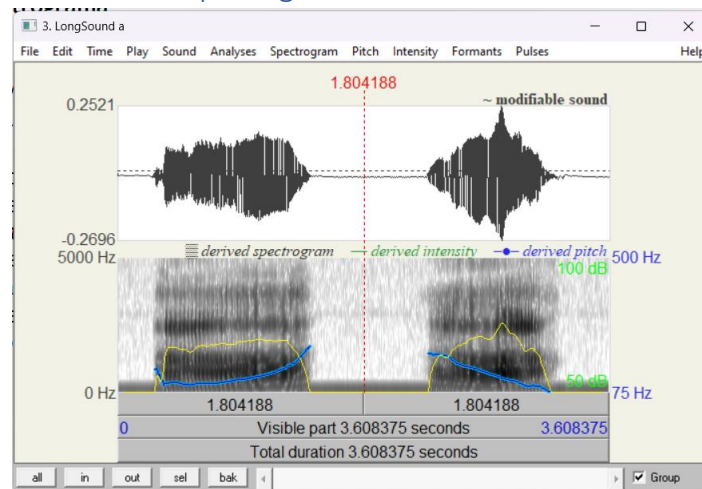
Time_s	F0_Hz
0.554000	--undefined--
0.564000	--undefined--
0.574000	--undefined--
0.584000	--undefined--
0.594000	--undefined--
0.604000	--undefined--
0.614000	--undefined--
0.624000	--undefined--
0.634000	--undefined--
0.644000	--undefined--
0.654000	--undefined--
0.664000	--undefined--
0.674000	--undefined--
0.684000	--undefined--
0.694000	--undefined--
0.704000	--undefined--
0.714000	--undefined--
0.724000	--undefined--
0.734000	313.133510
0.744000	320.623727
0.754000	317.236011
0.764000	319.197135

261.8244987742161 Hz (mean pitch in SELECTION)

La frecuencia fundamental en la primera u: 313.2990993191301 Hz (mean pitch in SELECTION)

En la segunda “u”: 168.2169952836674 Hz (mean pitch in SELECTION)

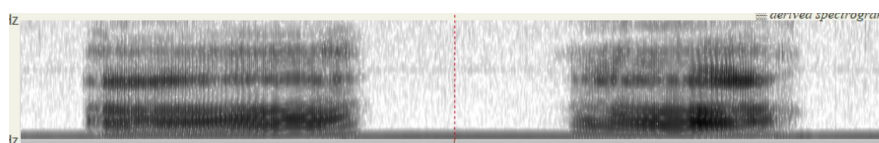
Actividad 2: Pitch Track vs. Espectrograma



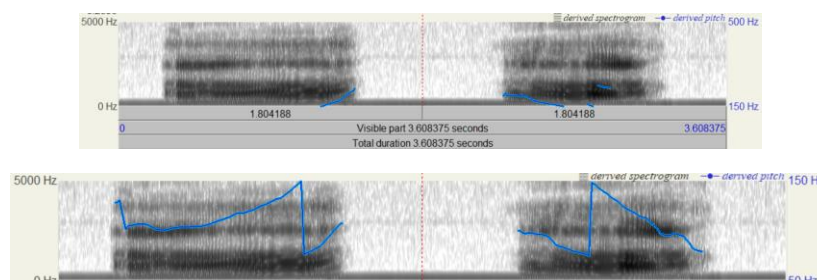
Eje de referencia: En el espectrograma se muestra la frecuencia por lo que el eje de referencia es la frecuencia mínima del hablante del sonido a.wav a la frecuencia máxima del hablante en este audio.

Cuando el tono sube parece que la intensidad es más constante (las zonas oscuras se mantienen constantes, además si nos fijamos en la gráfica de la frecuencia fundamental (línea amarilla de la imagen de arriba). Podemos ver como una vez empieza el sonido, la frecuencia efectivamente se mantiene aproximadamente constante.

Por otro lado cuando baja hay más picos y cambios de intensidad, sobre todo se nota al final cuando hay más zonas oscuras. Lo podemos verificar si analizamos otra vez la línea amarilla. Hay un pico a partir de la última mitad y luego sigue bajando:



Si cambiamos el Pitch setting a 150-500 Hz, sólo vemos las partes con el tono más alto (la última parte de donde el audio va de tono bajo a alto. Y la primera parte del audio donde va de tono alto a tono bajo. Es lo que se podía esperar:



Si ponemos el rango 50-150 Hz pasa lo contrario, solo visualizamos las zonas con el tono más bajo. Además como el rango es inferior (100 Hz vs los 350 Hz de la imagen anterior) podemos notar más las subidas y bajadas de tono. Por otro lado hay algunos picos que no cuadran (final de la primera parte e inicio de la segunda parte). Esto es lo que se conoce en el primer caso como pitch doubling y en el segundo como pitch halving.

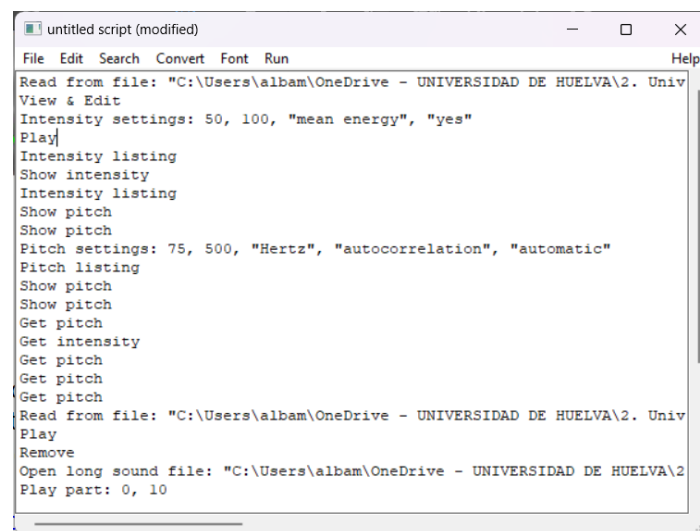
Actividad 3: Scripting

Duration

```
> .\praat duration.praat lamparita.wav  
Curso\1. 1º Cuatrimestre\Procesamiento del  
> 4.0880
```

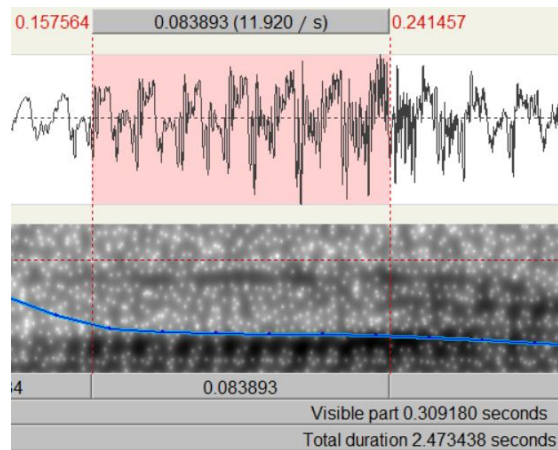
Acoustic:

```
500 .\praat acoustics.praat lamparita.wav 0.5 1.0 75  
PS C:\Users\albam\OneDrive - UNIVERSIDAD DE HUELVA\2. Universidad\4º Curso\1. 1º Cuatrimestre\Procesamiento del Habla, V  
isión e Interacción Multimodal\Prácticas\Práctica 2\archivos-practica2> SECONDS:0.500  
F0_MAX:341.629  
F0_MIN:247.602  
F0_MEAN:311.274  
F0_MEDIAN:317.807  
F0_STDV:22.470  
ENG_MAX:83.361  
ENG_MIN:46.801  
ENG_MEAN:69.707  
ENG_STDV:11.353  
VCD2TOT_FRAMES:0.532
```



Actividad 4: Intensidad y Nivel Tonal en Praat

Primer hola:

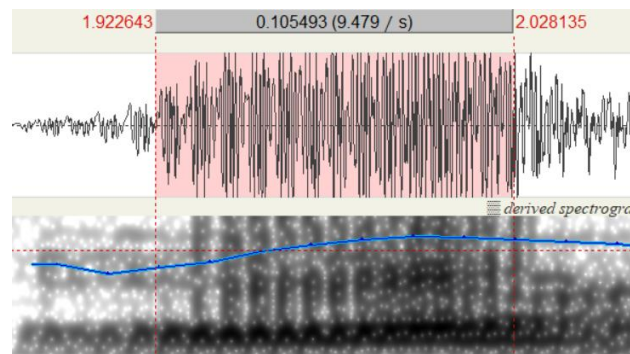


```
. \praat acoustics.praat hola.wav 0.157 0.241 50
300
PS C:\Users\albam\OneDrive - UNIVERSIDAD DE HUELVA\2. Universidad\4º Curso\1. 1º Cuatrimestre\Procesamiento del Habla, V
isión e Interacción Multimodal\Prácticas\Práctica 2\archivos-practica2> SECONDS:2.473
F0_MAX:135.703
F0_MIN:67.247
F0_MEAN:95.998
F0_MEDIAN:87.836
F0_STDV:22.531
ENG_MAX:90.844
ENG_MIN:42.893
ENG_MEAN:60.541
ENG_STDV:14.250
VCD2TOT_FRAMES:0.329
```

En la primera parte de la “o”: 63.488279668361365 dB

En la última parte de la “o”: 66.99494227300268 dB

Último hola:



```
. \praat acoustics.praat hola.wav 1.992 2.028 50
300
PS C:\Users\albam\OneDrive - UNIVERSIDAD DE HUELVA\2. Universidad\4º Curso\1. 1º Cuatrimestre\Procesamiento del Habla, V
isión e Interacción Multimodal\Prácticas\Práctica 2\archivos-practica2> SECONDS:1.000
F0_MAX:129.237
F0_MIN:84.226
F0_MEAN:98.382
F0_MEDIAN:94.665
F0_STDV:14.460
ENG_MAX:87.160
ENG_MIN:44.846
ENG_MEAN:58.980
ENG_STDV:13.100
VCD2TOT_FRAMES:0.270
```

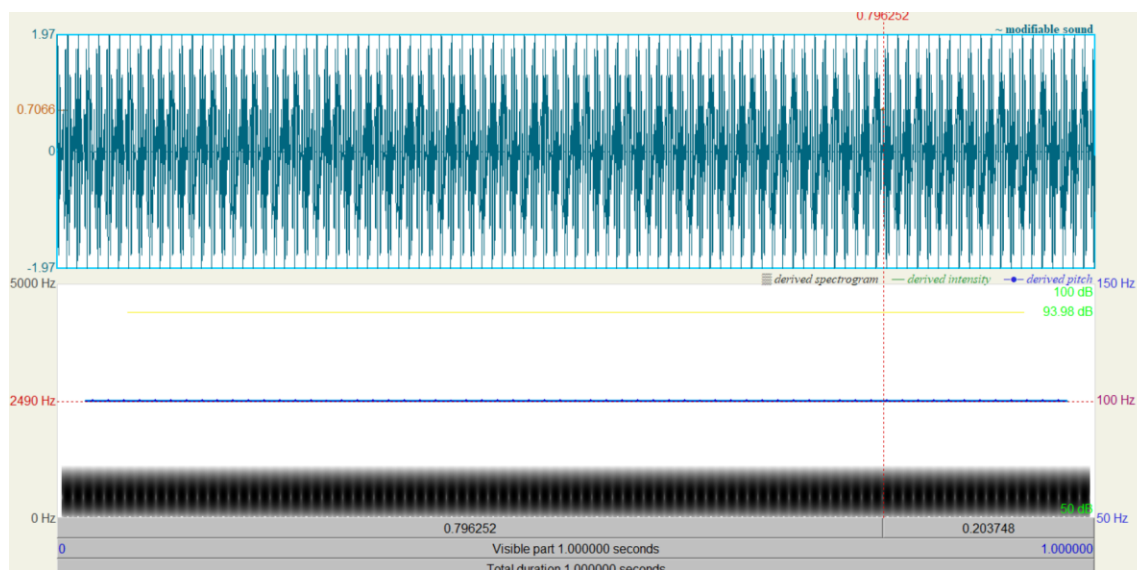
En la primera parte de la “o”: 86.99851613573856 dB

En la última parte de la “o”: 90.17512245710095 dB

La intensidad aumenta a lo largo del audio, tanto al principio como al final de la “o” como para los diferentes “hola”

Actividad 5: Intensidad y nivel tonal en Praat

Tras crear el sonido obtenemos que la f_0 es = 100 Hz, que además coincide con el Pitch



Actividad 6: Intensidad y nivel tonal en Praat

Audio do-re-mi-fa-sol-la-si. Este es el espectrograma del audio:

El pitch solo se puede percibir en algunas zonas del espectrograma:

