



ESCUELA TÉCNICA
SUPERIOR DE INGENIERÍA

uhu.es

Procesamiento del Habla, Visión e Interacción Multimodal.

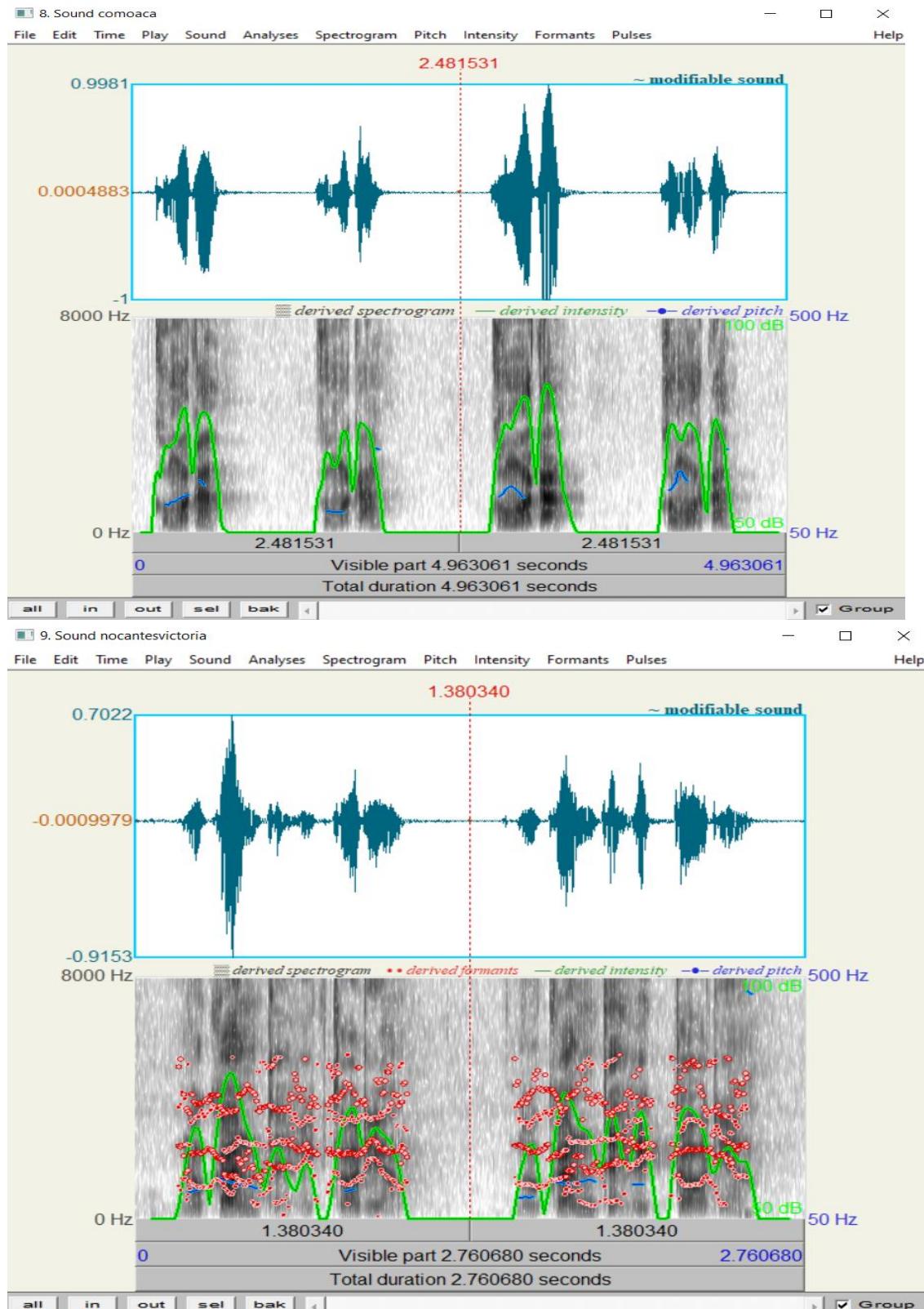
Práctica 4: Prosodia

Grado en Ingeniería Informática

Manuel Ramírez Ballesteros

Ejercicio 1: Estudio de la prosodia.

Para este ejercicio se estudiarán los archivos `nocantesvictoria.wav` y `comoaca.wav` para caracterizar su prosodia:



No -> Desde 0.203 hasta 0.285. F0 = 100.9 Hz. Intensidad media = 66.848 dB (crece unos 5 dB).

Can -> Desde 0.350 hasta 0.507. F0 = 113.6 Hz. Intensidad media = 76.063 dB (primero crece unos 4 dB en 0.04 segundo y luego decrece 19 dB).

Tes -> Desde 0.552 hasta 0.641. F0 = Indefinida. Intensidad media = 62.512 dB (decrece unos 7 dB).

Vic -> Desde 0.641 hasta 0.738. F0 = Indefinida. Intensidad media = 59.164 dB (crece unos 6 dB).

To -> Desde 0.819 hasta 0.927. F0 = 102.37 Hz. Intensidad media = 70.513 dB (crece unos 12 dB en 0.06 segundos y luego decrece 3 dB).

Ria -> Desde 0.942 hasta 1.099. F0 = Indefinida. Intensidad media = 66.354 dB (primero crece unos 4 dB en 0.05 segundos y luego decrece unos 9 dB).

No -> Desde 1.573 hasta 1.661. F0 = 89.94 Hz. Intensidad media = 63.248 dB (crece unos 10 dB y decrece 3 dB en los últimos 0.02 segundos).

Can -> Desde 1.707 hasta 1.896. F0 = 117.99 Hz. Intensidad media = 71.814 dB (crece unos 13 dB en 0.05 segundos y luego decrece unos 13 dB).

Tes -> Desde 1.927 hasta 2.053. F0 = 112.01. Intensidad media = 67.921 dB (crece 5 dB en 0.03 segundos, luego decrece unos 12 dB en 0.07 segundos y crece un par de decibelios al final).

Vic -> Desde 2.054 hasta 2.109. F0 = 113.19. Intensidad media = 70.621 dB (crece cerca de 4 dB en 0.015 segundos, se mantiene estable 0.015 segundos y decrece 4 dB al final).

To -> Desde 2.226 hasta 2.354. F0 = 102.2 Hz. Intensidad media = 70.958 dB (crece 2 dB al principio, se mantiene estable unos 0.04 segundos y decrece 8 dB en 0.07 segundos).

Ria -> Desde 2.382 hasta 2.546. F0 = 472.63 (¿mal?). Intensidad media = 62.423 dB (crece unos 5 dB en 0.06 segundos y decrece 9 dB en 0.1 segundos).

Co -> Desde 0.182 hasta 0.280. F0 = 107.62. Intensidad media = 67.562 dB (crece unos 5 dB).

Mo -> Desde 0.280 hasta 0.379. F0 = 116.01 Hz. Intensidad media = 73.332 dB (crece unos 8 dB).

A -> Desde 0.379 hasta 0.427. F0 = 352.30 Hz. Intensidad media = 71.818 dB (decrece unos 2 dB).

Ca -> Desde 0.478 hasta 0.631. F0 = 156.42 Hz. Intensidad media = 71.81 dB (decrece unos 4 dB).

Co -> Desde 1.402 hasta 1.492. F0 = 93.27 Hz. Intensidad media = 65.577 dB (crece 10 dB).

Mo -> Desde 1.492 hasta 1.597. F0 = 92.41 Hz. Intensidad media = 68.440 dB (crece unos 6 dB).

A -> Desde 1.597 hasta 1.655. F0 = 100.9 Hz. Intensidad media = 66.848 dB (decrece unos 5 dB).

Ca -> Desde 1.689 hasta 1.891. F0 = 223.56 Hz. Intensidad media = 72.172 dB (empieza creciendo unos 9 dB en 0.03 segundos y luego decrece 20 dB).

Co -> Desde 2.727 hasta 2.823. F0 = 245.01 Hz. Intensidad media = 70.863 dB (crece 13 dB).

Mo -> Desde 2.823 hasta 2.951. F0 = 139.51 Hz. Intensidad media = 77.192 dB (crece unos 5 dB).

A -> Desde 2.952 hasta 3.024. F0 = 127.26 Hz. Intensidad media = 80.192 dB (decrece 5 dB).

Ca -> Desde 3.081 hasta 3.357. F0 = 79.02. Intensidad media = 79.027 dB (primero crece unos 12 dB en 0.06 segundos y luego decrece 34 dB).

Co -> Desde 4.010 hasta 4.096. F0 = 147.98 Hz. Intensidad media = 72.665 dB (crece unos 14 dB).

Mo -> Desde 4.122 hasta 4.217. F0 = 165.69 Hz. Intensidad media = 73.630 dB (crece unos 3 dB).

A -> Desde 4.251 hasta 4.319. F0 = 226.56 Hz. Intensidad media = 71.345 dB (decrece unos 6 dB).

Ca -> Desde 4.378 hasta 4.561. F0 = 216.81 Hz. Intensidad media = 72.120 dB (decrece 22 dB).

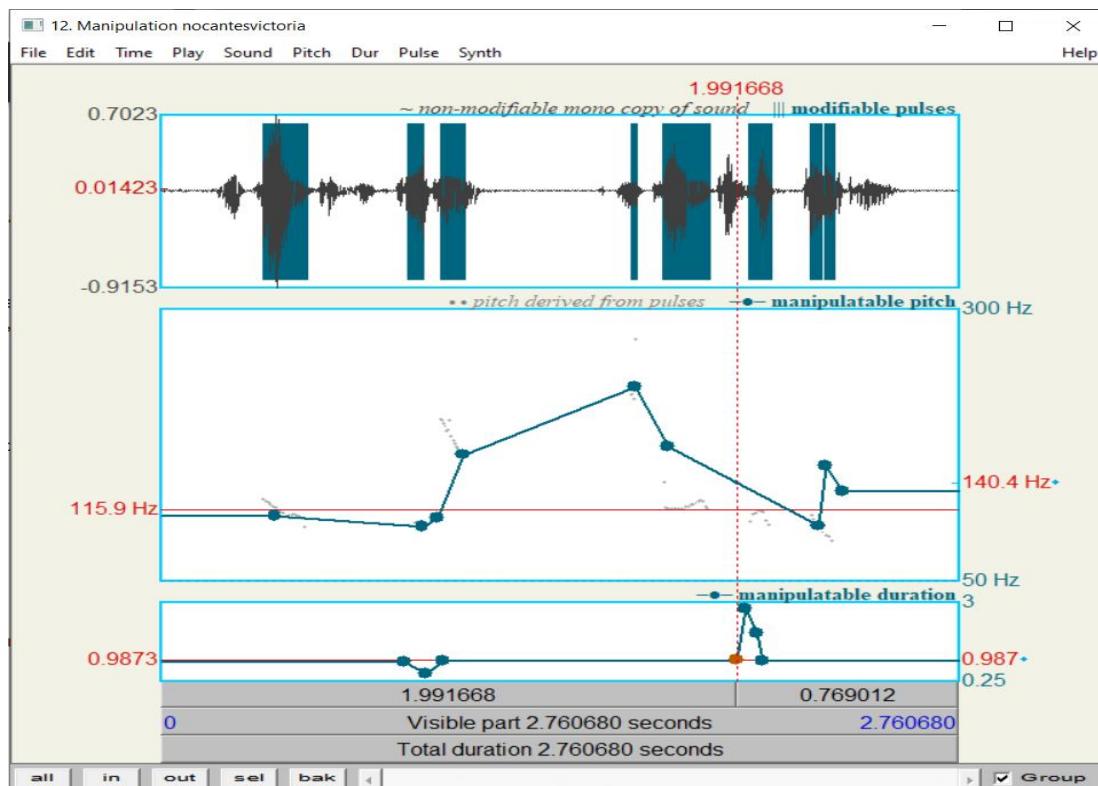
Podemos caracterizar la prosodia en la afirmación de “como acá” en que la mayor variación de intensidad se produce en la sílaba /mo, donde crece unos 8 dB, mientras

que en la pregunta apenas se aprecia crecimiento en dicha sílaba, sin embargo, se observa un gran crecimiento de volumen en el /co y un gran decrecimiento en el /ca.

En cuanto al significado de “igual que acá” o “comer acá”, comparando la forma interrogativa de la segunda y la cuarta pronunciación, vemos que la frecuencia fundamental es más baja en la forma de “igual que acá” y que la duración de /a se alarga algo más en la forma de “comer acá”.

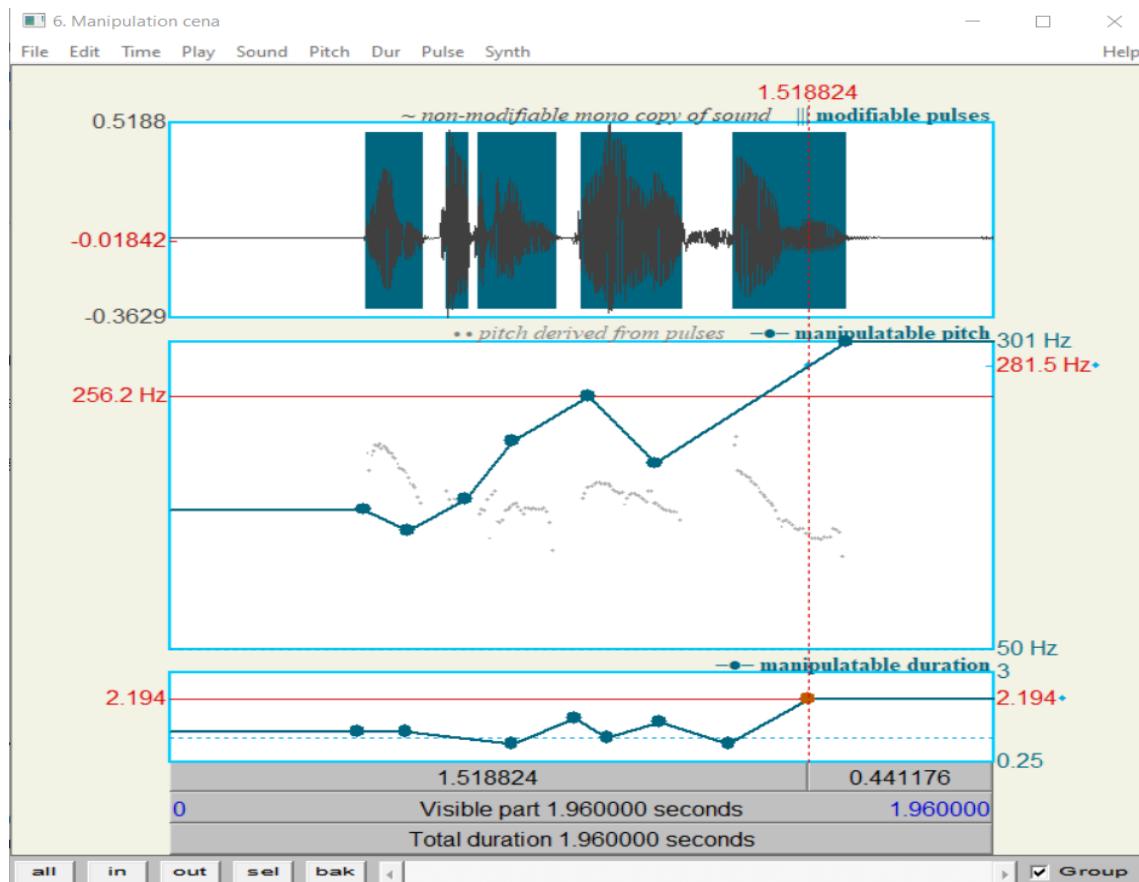
Ejercicio 2: Manipulación de la prosodia.

Para este ejercicio se han seguido los pasos descritos en el guión de la práctica, modificando el audio `nocantesvictoria.wav`, cuyo archivo manipulado queda de la siguiente manera:



Ejercicio 3: Manipulación de la prosodia.

Se ha modificado el fichero cena.wav con el fin de que la frase suene interrogativa con un tono de indignación mediante la manipulación adecuada del pitch y de la duración de los distintos sonidos que componen el audio. El resultado es el siguiente:



Ejercicio 4: Manipulación de la prosodia.

Se ha modificado el fichero lamparita.wav con el fin de modificar la prosodia en la palabra alcanza. El resultado es el siguiente:

