



Procesamiento del Habla, Visión e Interacción Multimodal.

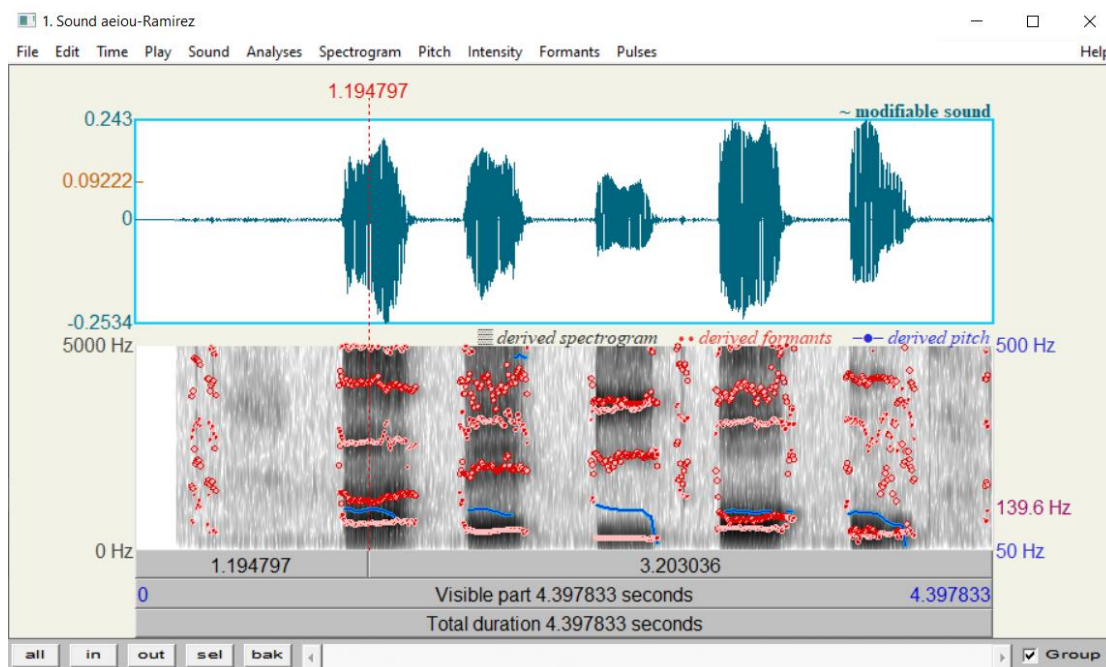
Práctica 3: Fonética y Fonología

Ingeniería Informática

Manuel Ramírez Ballesteros

Ejercicio 1: Formantes

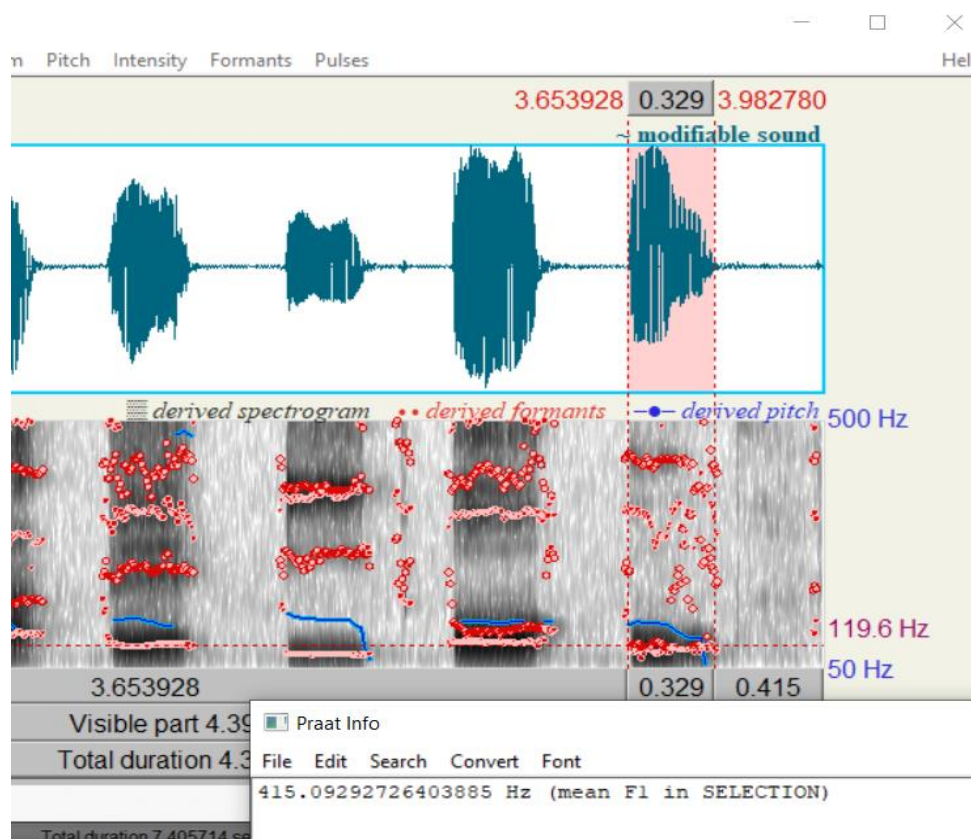
Para este ejercicio se ha empleado el audio solicitado en la práctica anterior aeiou_Ramirez.wav con el que se va a estudiar las características que presentan los espectrogramas de las distintas vocales, prestando especial atención en los formantes:



Como puede apreciarse a simple vista, existen ciertas diferencias entre las distintas vocales. Para un análisis más detallado, se considerará que los dos primeros formantes nos permiten identificar las vocales que emitimos, mientras que los formantes f3, f4 y f5 nos aportan información sobre el color (timbre) de la propia voz. También se considerará que f1 es proporcional a la apertura de la mandíbula (a más apertura, más frecuencia para el primer formante) e inversamente proporcional a la altura de la lengua (al subir, f1 disminuye). En cuanto al segundo formante, éste variará en función de su cercanía a los dientes (a mayor cercanía, mayor frecuencia).

Considerando este, vemos que la /i/ y la /u/ son las que presentan un formante f1 más bajo, seguidas de /e/ y /o/ (donde apenas hay movimiento en la altura de la mandíbula al cambiar de un fono al otro) y la que presenta un primer formante mayor es la /a/.

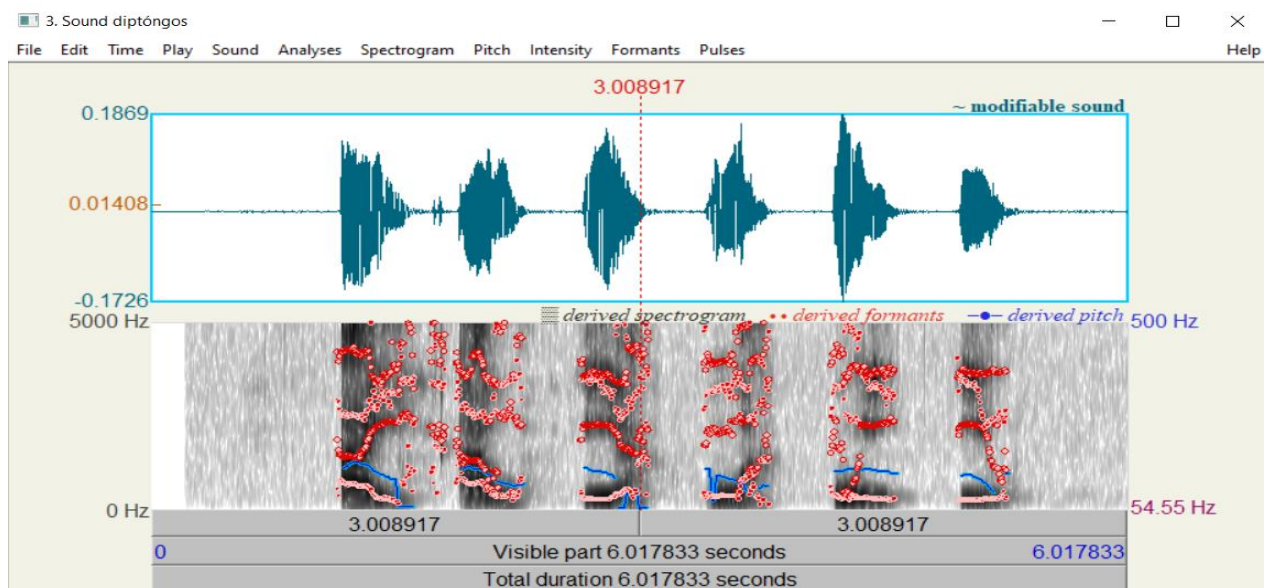
En cuanto al segundo formante, el que tiene mayor frecuencia es la /i/, lo que implica que es quien se genera con la lengua más cercana a los dientes, después tenemos la /e/, seguida de la /a/, la cual es considerada una vocal central, ya que en torno a los 1300Hz suele representarse la lengua en el centro de la boca. Por último, tendremos la /o/ y, aunque cueste distinguir el formante en el espectrograma, la /u/, que praat si es capaz de identificar:



Ejercicio 2: Diptongos, triptongos, sordas y sonoras.

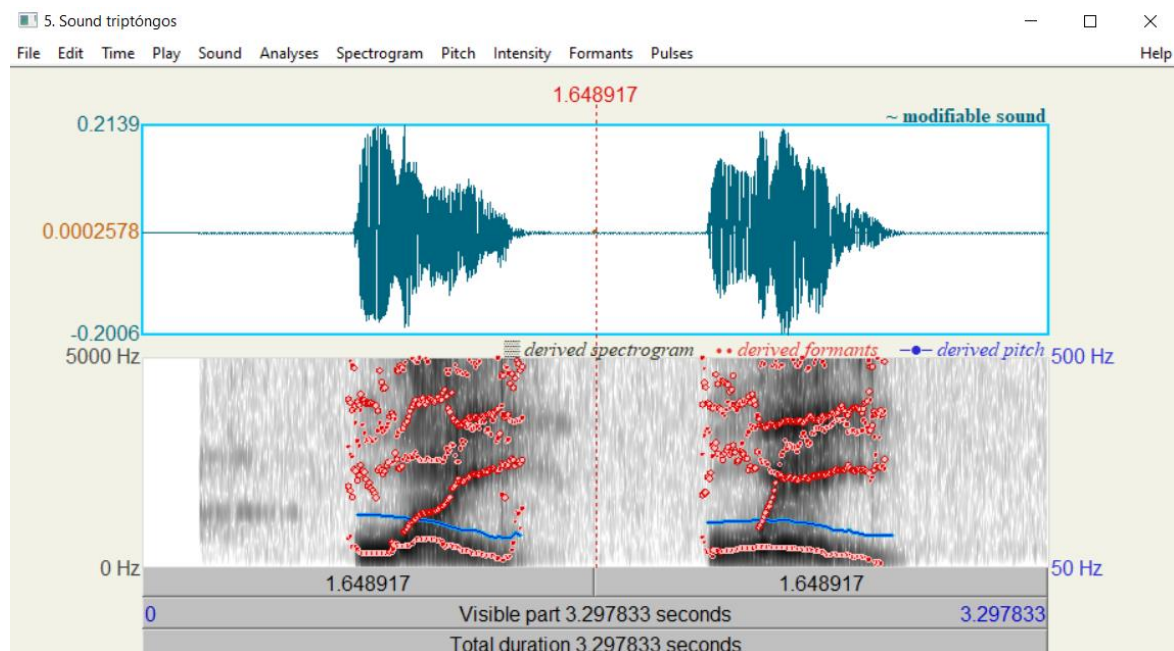
Para este ejercicio se han generado tres audios: diptongos.wav (con una versión más espaciada), triptongos.wav y articulación.wav.

Para el primero, se han grabado los diptongos /ai/, /au/, /ia/, /ua/, /ui/ y /iu/:

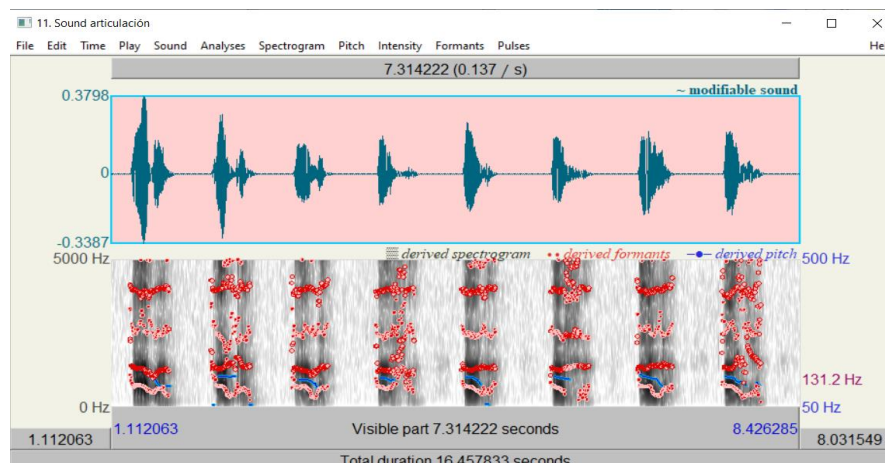


Podemos apreciar que existen ciertas transiciones tanto de la f_0 como de los formantes al pasar de una vocal a otra en todos los sonidos.

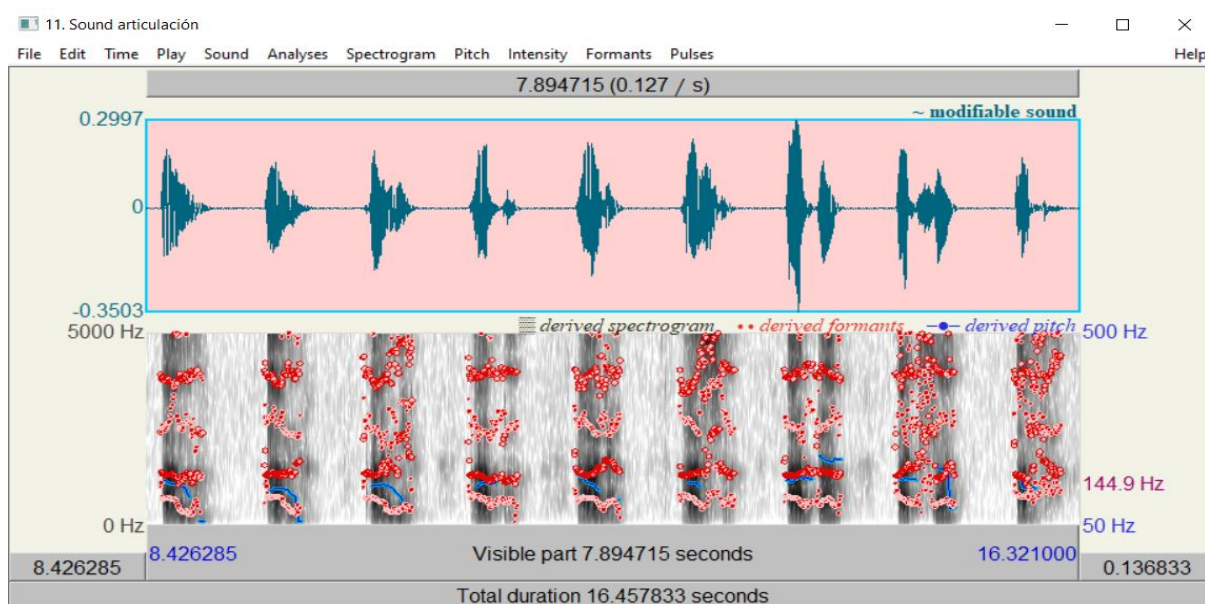
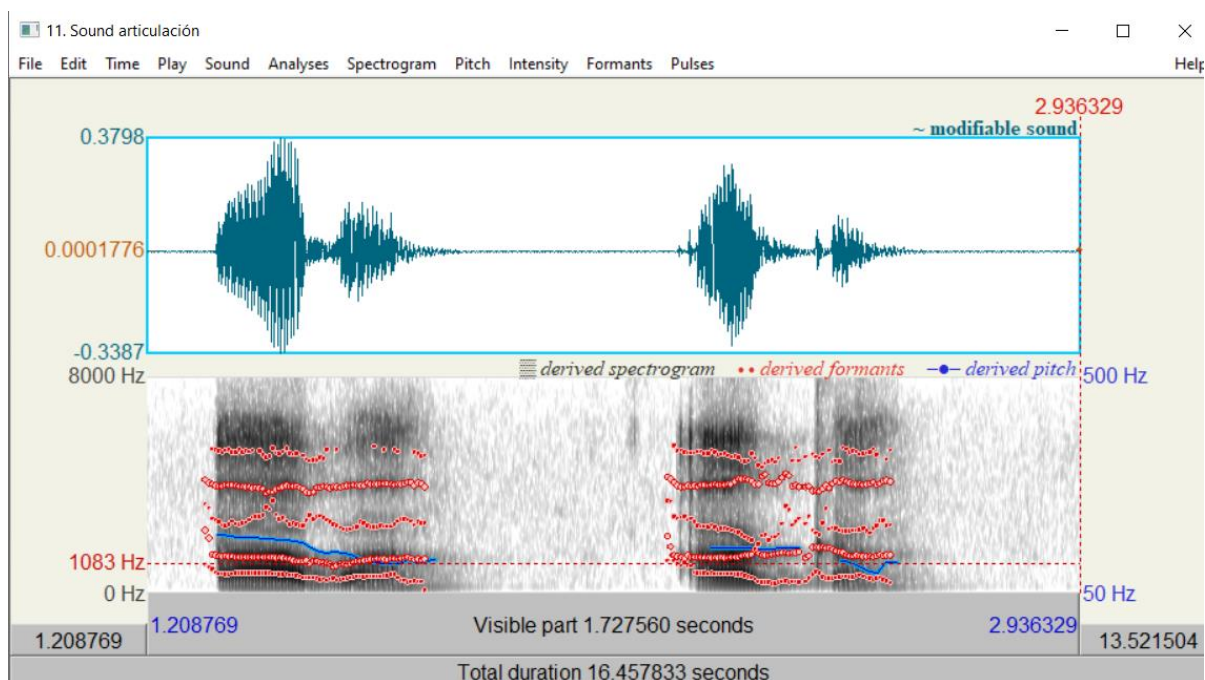
Para el segundo, se han grabado los triptongos /uai/ y /uei/:



Para el tercero, se han grabado las articulaciones /aba/, /aca/, /ada/, /afa/, /aga/, /aja/, /ala/, /alla/, /ama/, /ana/, /aña/, /apa/, /ara/, /asa/, /ata/, /axa/ y /aza/:



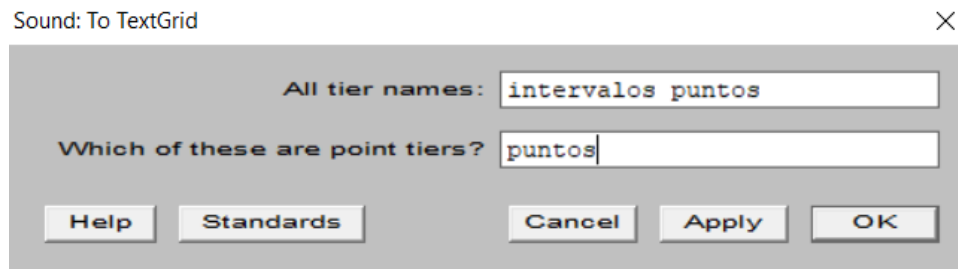
En esta primera parte, podemos ver las articulaciones desde /aba/ hasta /alla/. Para un mejor análisis, distinguiremos entre consonantes sordas, las cuales no producen vibración de las cuerdas vocales, y sonoras, que sí. Analizando los espectrogramas en detalle, obtenemos que las sonoras son /b/, /d/, /j/, /g/, /l/ y /ll/, mientras que las sordas son /c/ y /f/. Además, poniendo como ejemplo /aba/ y /aca/, podemos ver que, para las sonoras, Praat estima la frecuencia fundamental sin discontinuidades (sonido periódico, aunque para /aja/ y /aga/ sí se aprecia discontinuidad), mientras que, para las sordas, aparecen líneas verticales en el espectrograma.



Para esta segunda parte, se han analizado los espectrogramas desde /ama/ hasta /aza/, y observamos que las sonoras son /m/, /n/, /ñ/, /r/, mientras que las sordas son /p/, /s/, /t/, /x/ y /z/. Aunque se ven algunos casos como la /r/ y la /z/ que pueden generar cierta confusión.

Ejercicio 3: Anotación de Audio en Praat. Formato TextGrid.

En este ejercicio se ha generado un TextGrid para el sonido aeiou-Ramirez.wav, seleccionando una capa de intervalos y otra de puntos de la siguiente manera:



Se ha editado el TextGrid para familiarizarnos con el uso de los puntos y los intervalos y el etiquetado de los mismos, seleccionando los intervalos de las vocales y los centros de las correspondientes ondas para los puntos:

