

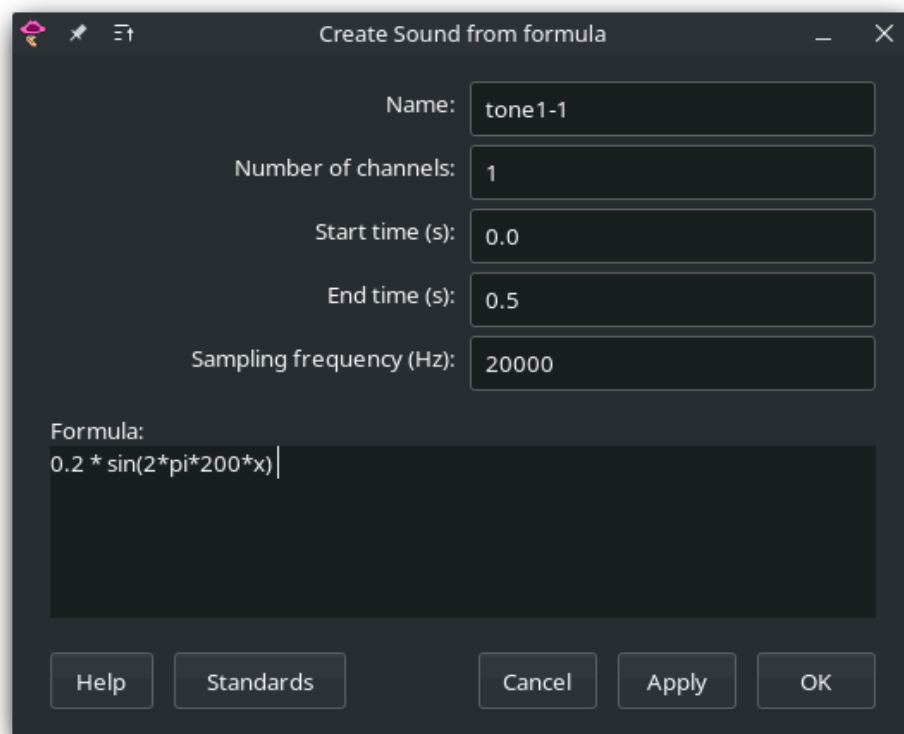
Entrega 1 - Tecnologías del Habla

Ejercicio 1

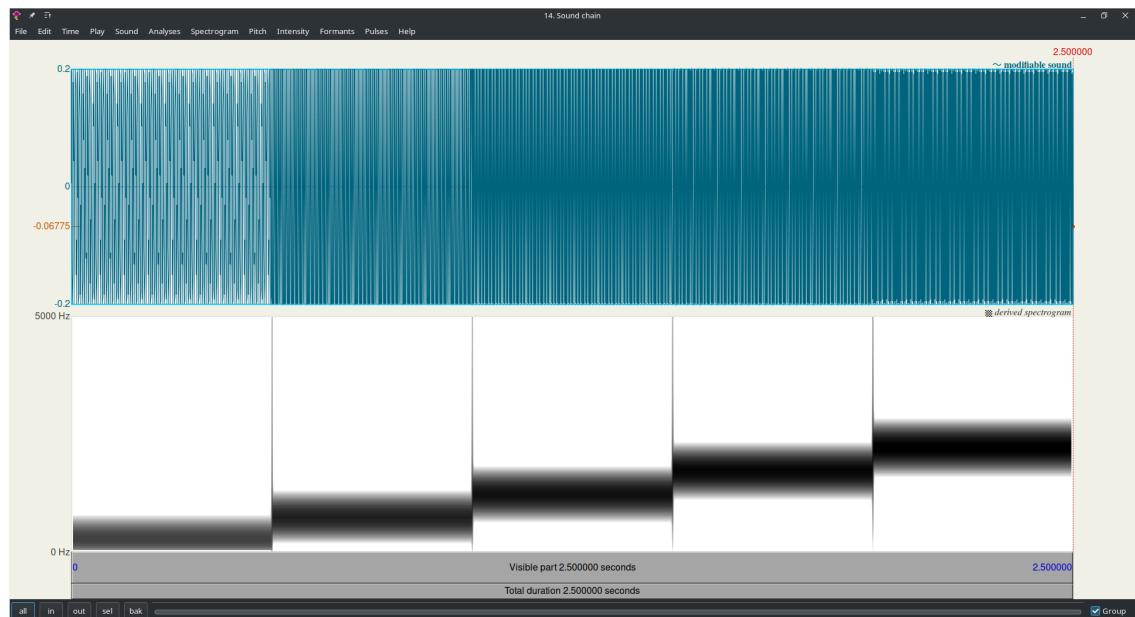
Para este ejercicio se han de generar cuatro secuencias de tonos puros con Praat, cada una se almacenará en formato WAV, mono, con una tasa de muestreo de 20000 Hz:

1. Esta primera secuencia está formada por cinco tonos de 0.5 segundos cada uno, de igual amplitud (0.2 Pa) y frecuencias separadas de forma equidistante entre 200 y 2200 Hz (200, 700, 1200, 1700, 2200).

Para crear los tonos se ha empleado la fórmula correspondiente al modelo de una onda simple: $p(t) = A \cdot \sin(2\pi f t)$. Donde A es la amplitud de la onda, en este caso será 0.2 para todos los tonos, y f es la frecuencia de la onda, que dependerá del tono en concreto. El tiempo inicial y final para todos los casos es 0.0 y 0.5 segundos respectivamente. Los tonos se han creado a partir de sus respectivas fórmulas empleando el diálogo correspondiente de Praat como se puede ver en la siguiente figura:

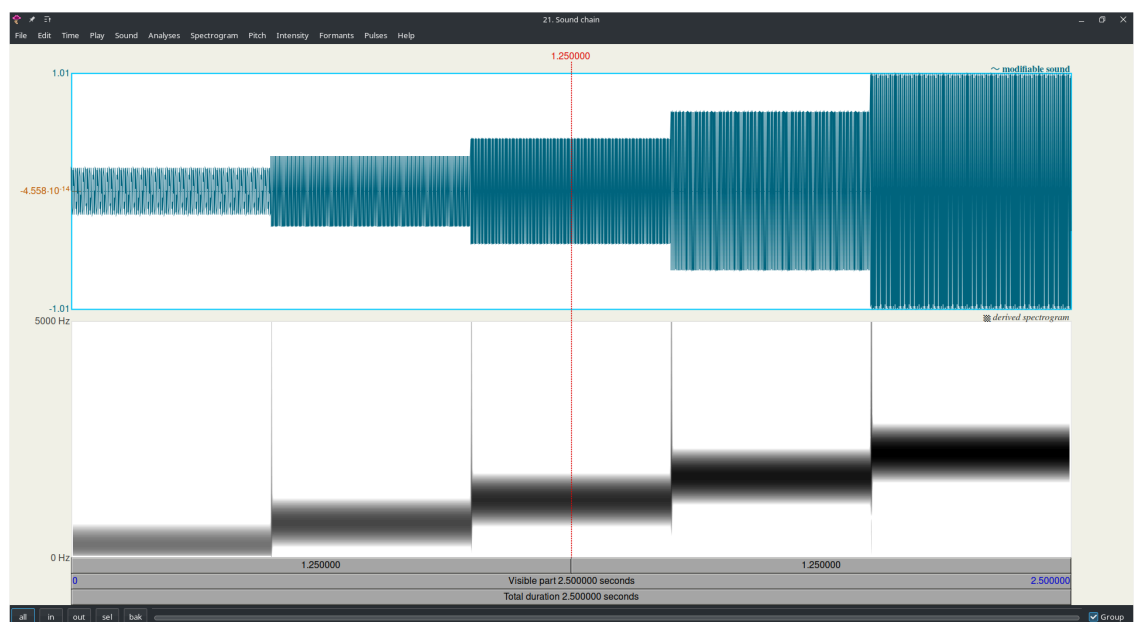


Una vez generados los cinco tonos se han combinado en una sola señal, representada en la siguiente figura, donde se representa la señal sobre el tiempo y su correspondiente espectrograma:



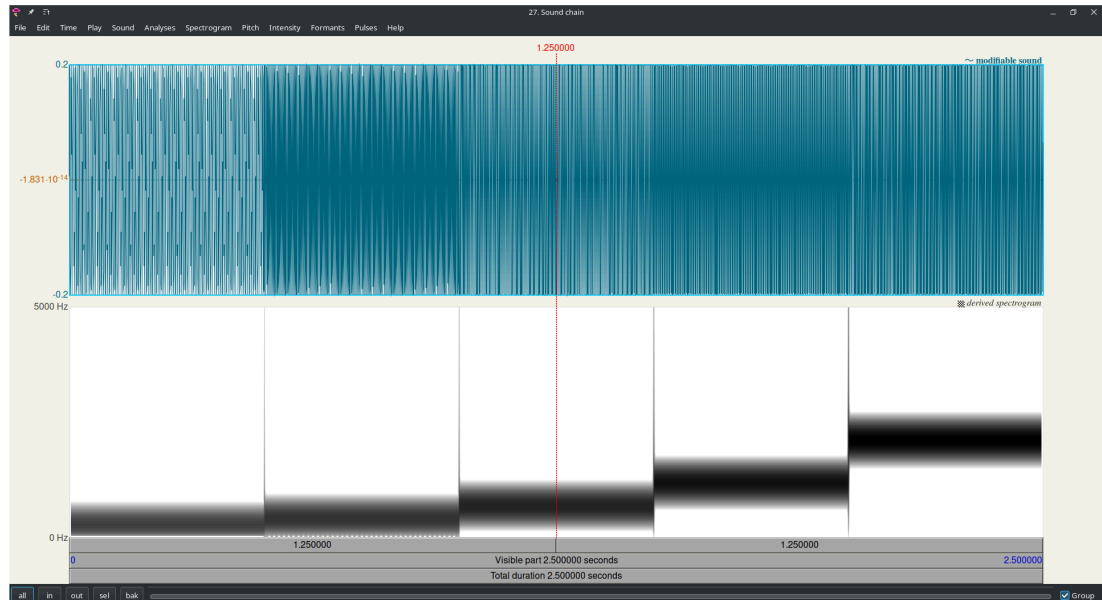
Esta señal se ha guardado en el archivo `secuencia1.wav`.

2. Para la segunda secuencia se ha seguido el mismo procedimiento, pero cambiando la amplitud de las fórmulas de forma creciente (0.2, 0.3, 0.45, 0.68, 1.01). En la figura de la señal generada se puede ver como la amplitud de la señal se incrementa con el tiempo:



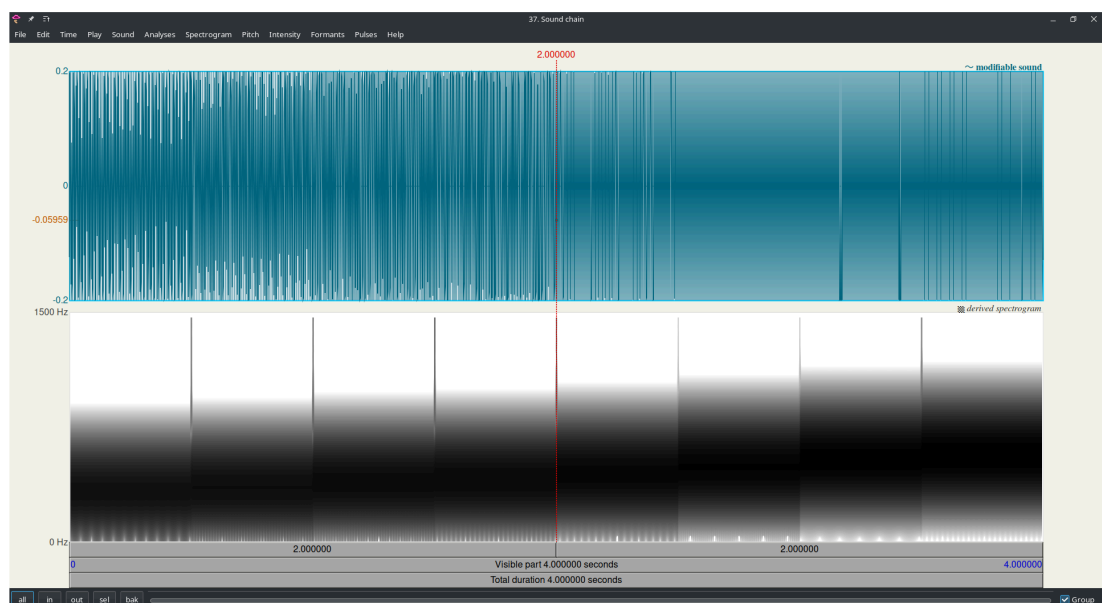
Esta señal se ha guardado en el archivo `secuencia2.wav`, al guardarlo avisa que la amplitud de algunas frecuencias han sido recortadas al superar el límite (correspondiente a la amplitud 1.01 de la frecuencia 2200).

- En la tercera secuencia los tonos mantendrán la misma amplitud, pero su frecuencia se incrementará en un ratio de 180% (200, 360, 648, 1166, 2099). La señal obtenida es la siguiente, se puede ver en el espectrograma como la frecuencia aumenta de forma más exponencial, en vez de lineal:



Esta señal se ha guardado en el archivo secuencia3.wav.

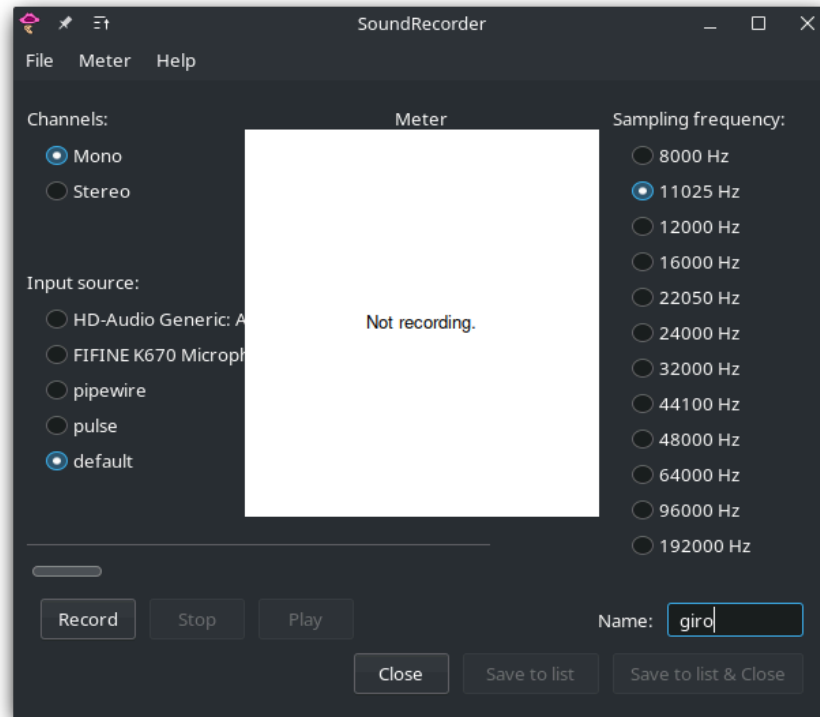
- La cuarta secuencia estará formada por los tonos correspondientes a las notas del solfeo en la cuarta octava (Do = 261.13, Re = 293.66, Mi = 329.63, Fa = 349.23, Sol = 392, La = 440, Si = 493.88, Do' = 523.25). En la siguiente figura se puede observar la señal resultante:



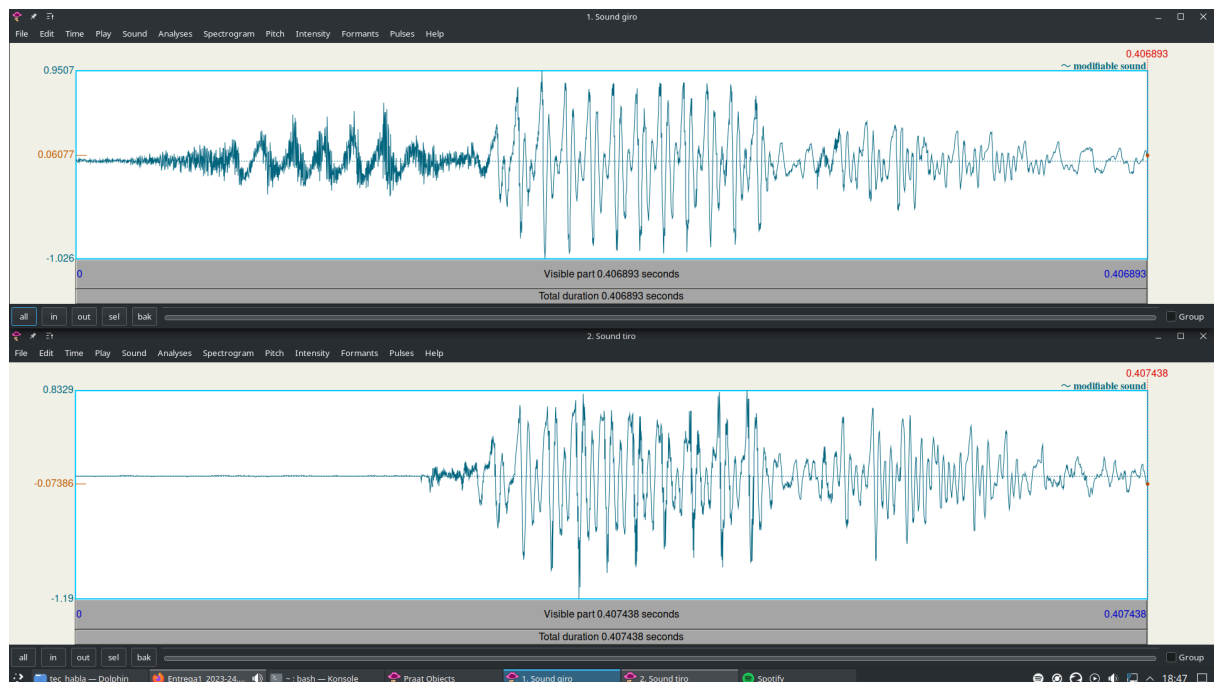
Esta señal se ha guardado en el archivo secuencia4.wav.

Ejercicio 2

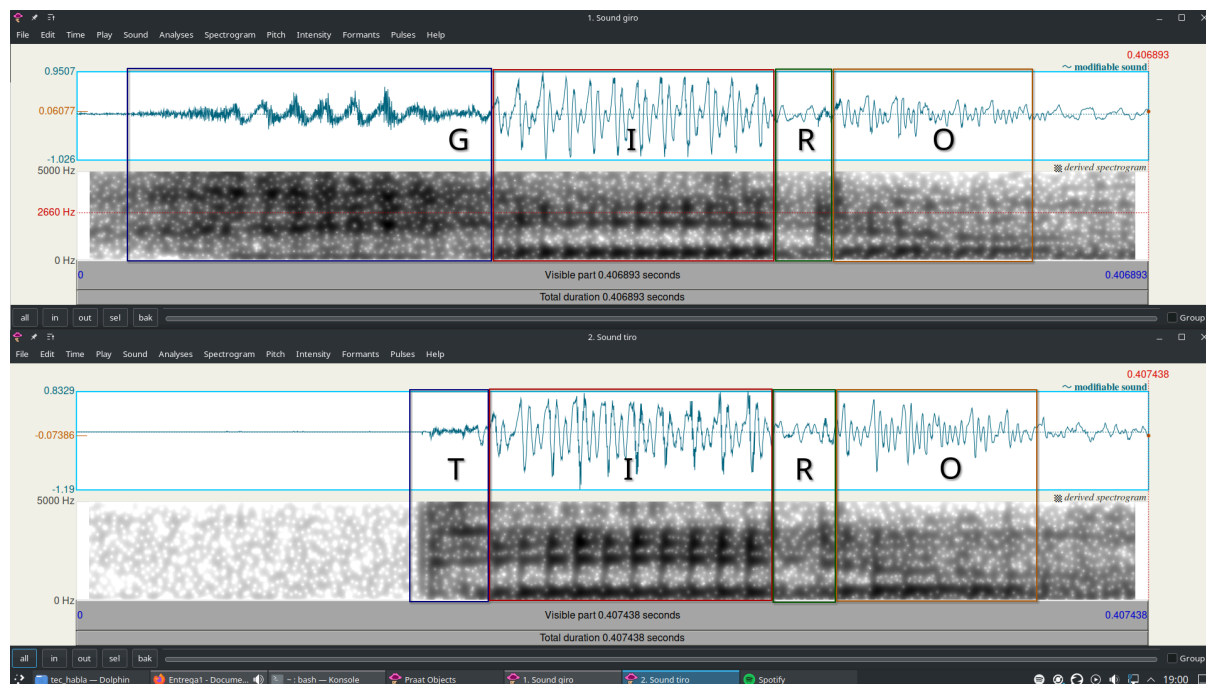
- A. Primero se han grabado, en mono y con una tasa de muestreo de 11025 Hz, las palabras “giro” y “tiro”, que difieren en un fonema. Estas grabaciones se han guardado en formato WAV en los archivos giro.wav y tiro.wav respectivamente.



- B. La representación gráfica de las ondas de cada una de las grabaciones es la siguiente:



- C. En el siguiente gráfico se han identificado las regiones que corresponden a cada fonema de cada palabra. Se puede ver claramente la diferencia entre el fonema correspondiente a la “g” de giro con el fonema correspondiente a la “t”, el primero tiene una mayor duración que el segundo. En el resto de fonemas se puede ver también la similitud en ambas grabaciones, la duración de cada fonema es prácticamente el mismo, por lo que la velocidad de pronunciación es muy similar para ambas palabras.



Ejercicio 3

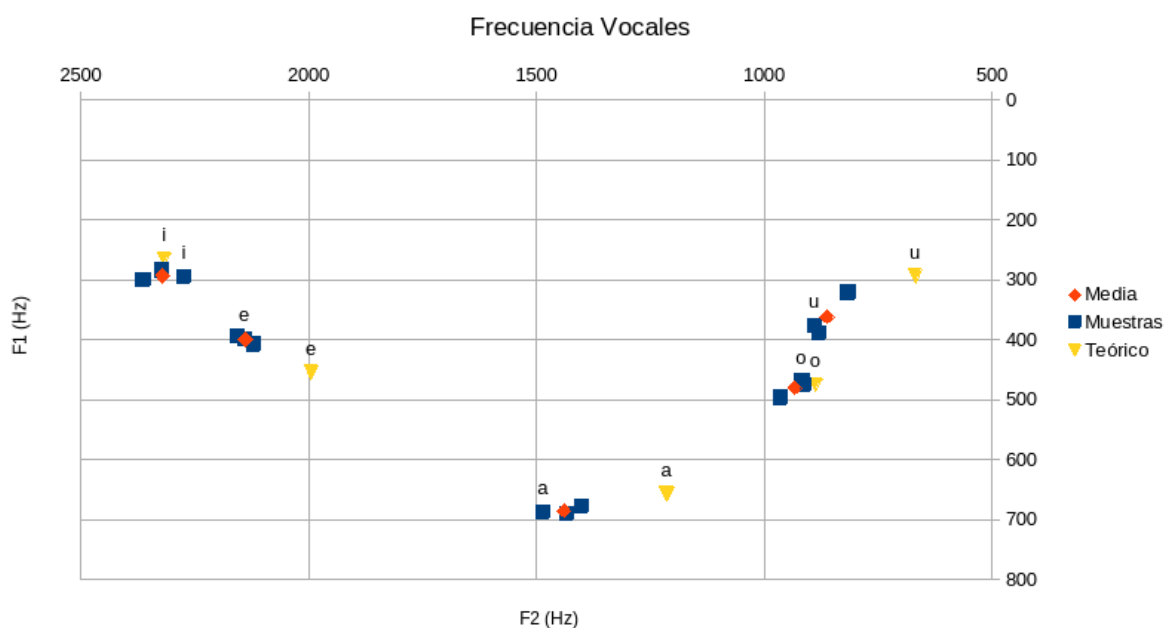
Se han realizado tres grabaciones distintas pronunciando la secuencia de vocales “aeiou”, las medidas del primer y segundo formante para cada vocal son las siguientes:

Grabación	1		2		3	
Formante	F1	F2	F1	F2	F1	F2
a	688	1486	690	1434	678	1402
e	399	2141	393	2158	407	2122
i	295	2275	299	2364	284	2323
o	469	919	496	966	475	915
u	376	891	388	880	321	818

La siguiente tabla muestra la media de los valores recogidos comparados con los valores teóricos para cada vocal en el caso de los hombres:

Formante	Valores Medios		Valores Teóricos Hombres	
	F1	F2	F1	F2
a	685	1441	657	1215
e	400	2140	454	1995
i	293	2321	265	2318
o	480	933	475	888
u	362	863	294	669

El siguiente gráfico de dispersión muestra los valores de las muestras tomadas para cada vocal, junto con su media y su valor teórico, el eje X representa la frecuencia del segundo formante, y el eje Y la frecuencia del primer formante. La vocal “u” es la que más se desvía de su valor teórico, esto es especialmente cierto si se tiene en cuenta que su frecuencia es más baja que el resto, y el oído humano identifica mejor las diferencias de frecuencia en frecuencias bajas comparadas con las más altas.



Ejercicio 4

- A. Secuencia de seis palabras cuya distancia acústica en términos lingüísticos entre dos consecutivas es uno:

CHALET > BALLET > VALE > BAILE > BAILA > BAHÍA

- B. Secuencia de seis palabras cuya distancia en grafía entre dos consecutivas es uno:

HOLA > OLA > BOLA > BOCA > BACA > VACA

Ejercicio 5

Las transcripciones fonéticas de las palabras del ejercicio anterior empleando el Alfabeto Fonético Internacional y el Alfabeto fonético de la RFE son las siguientes:

Palabra	Transcripción AFI	Transcripción RFE
chalet	/tʃ a l e t/	/ç a l e t/
ballet	/b a l e t/	/b a l e t/
vale	/b a l e/	/b a l e/
baila	/b a i l a/	/b a i l a/
bahía	/b a i a/	/b a i a/
hola	/o l a/	/o l a/
ola	/o l a/	/o l a/
bola	/b o l a/	/b o l a/
boca	/b o k a/	/b o k a/
baca	/b a k a/	/b a k a/
vaca	/b a k a/	/b a k a/

Los rasgos articulatorios para los fonemas de “bahía” son los siguientes:

	Vocales				Consonantes
	Anterior	Central			Bilabial
Cerrada	i				Sonora
Abierta		a		Oclusiva	b