



**Universitat de les
Illes Balears**

Facultad de Ciencias

Memoria del Trabajo de Fin de Grado

*Psicoacústica comparativa: Percepción dicótica vs diótica

Carlos Ruiz Quirós

Grado de Física

Año académico 2018-19

Trabajo tutelado por Oreste Piro Perusín
Departamento de Física

Se autoriza a la Universidad a incluir este trabajo en el Repositorio Institucional para su consulta en acceso abierto y difusión en línea, con finalidades exclusivamente académicas y de investigación	Autor		Tutor	
	Sí	No	Sí	No
	X		X	

Palabras clave del trabajo:

Psicoacústica, percepción binaural, dicótico, diótico, fundamental faltante, altura, armónicos.

Resumen

La psicoacústica es una rama de la Acústica que estudia la percepción subjetiva de las diversas cualidades del sonido. Por ejemplo, la psicoacústica describe la manera en que el oído percibe las propiedades del sonido tales como son el *tono*, la *altura*, fenómenos perceptivos asociados a la audición *binaural*, etc.

El presente trabajo de fin de grado, se enfoca, precisamente, en el estudio de un fenómeno perceptivo particular en la audición binaural: la fusión de tonos de diferente frecuencia presentados a cada uno de los oídos. Dos ejemplos relevantes de estos fenómenos son los denominados “*Binaural Beats*” o pulsos binaurales en el que dos tonos puros presentados individualmente en cada oído batan de la misma manera que si se los presentara superpuestos en cada oído. Otro, de consecuencias muy importantes en la teoría de la percepción de altura de los sonidos, es el de la reconstrucción binaural de la altura de la fundamental ausente, en un sonido compuesto por armónicos presentados en combinaciones diferentes sobre cada oído. Aquí nos concentraremos en este último fenómeno.

El problema concreto que se investiga en este trabajo, es la percepción comparativa de la altura de un sonido compuesto por dos armónicos superiores contiguos de una frecuencia fundamental no incluida en el estímulo, cuando dichos armónicos se presentan en forma dicótica (uno en cada oído) en contraste de cuando se presentan su suma en forma diótica (el mismo estímulo en los dos oídos). Este problema se aborda experimentalmente, desarrollando un protocolo para interpelar a sujetos voluntarios sobre su percepción comparativa entre un estímulo dicótico compuesto por el tercer y cuarto armónico de una frecuencia fundamental fija, y una sucesión estímulos dióticos cuya frecuencia fundamental varía a elección del sujeto, que tiene que decidir para cual cuando la altura percibida de ambos estímulos coincide.

La conclusión preliminar de los experimentos realizados sobre un total de 36 sujetos voluntarios, con y sin educación musical, indica que hay una desviación sistemática de las alturas diótica y dicótica. Este hallazgo es una fuerte motivación para realizar ulteriores comprobaciones con técnicas experimentales más refinadas y mayores muestreos, y permite, asimismo, intuir la necesidad de una reformulación teórica de ciertos aspectos de la interpretación del problema “pitch” de la fundamental ausente.

Finalmente, quiero agradecer para la obtención de datos y el correspondiente análisis numérico al Dr. Oreste Piro Perusín y a los voluntarios que se han ofrecido para este estudio. Especialmente, destacar la ayuda de Carlos Rey Romero y de Enric Munar, alumno de Doctorado y profesor catedrático de Psicología respectivamente por la gran ayuda prestada para la realización de los mencionados experimentos.

Índice

1. Introducción	5
1.1 Nociones básicas de acústica.	5
1.2 Fisiología auditiva.	11
1.2.1 Oído externo.	11
1.2.2 Oído medio.	12
1.2.3 Oído interno.	12
1.3 Análisis de Fourier	13
1.4 El concepto de la psicoacústica.	14
1.5 La percepción del <i>pitch</i>	15
1.5.1 Fenómeno de la “ <i>missing fundamental</i> ”	17
1.5.2 Teoría del lugar.	18
1.5.2 Teoría temporal.	19
2. Percepción binaural de la altura virtual	19
2.1 Estudio de fenómenos binaurales.	20
2.2 Percepción dicótica vs diótica.	22
2.3 Experimento realizado	22
3. Resultados obtenidos	25
3.1 Datos generales.	26
3.1.1 Bloque 1	26
3.1.2 Bloque 2	27
3.2 Datos de los sujetos musicales	28
4. Discusión y conclusiones	31
5. Bibliografía	32

1.Introducción

Antes de pasar de lleno a la descripción del fenómeno psicoacústico bajo estudio en este trabajo, creo necesario hacer una introducción exhaustiva a varios fenómenos de la acústica y la anatomía y fisiología del aparato auditivo fundamentales para nuestro trabajo. Seguidamente se incluirán nociones del análisis de Fourier aplicado a señales sonoras. A continuación, se introducirán las nociones generales del campo de la psicoacústica con énfasis en la percepción del “pitch” que denominaremos, en castellano, “altura” para distinguirlo de “tono” que se referirá a las frecuencias físicas de los estímulos sonoros. En particular, se describirá el fenómeno de la percepción de la altura de estímulos con fundamental ausente y finalmente las distintas teorías desarrolladas para explicar este fenómeno.

1.1 Nociones básicas de acústica

La acústica es la rama de la Física que estudia el sonido, el cual, de una manera general podemos entender como la propagación de energía mecánica en un medio determinado (sólido, líquido o gas). Un modelo simple de esta propagación en un medio unidimensional podría consistir en una cadena de bloques representando, cada uno, una porción del medio en cuestión, unidos mediante muelles elásticos representando las fuerzas intermoleculares en tal medio. En la figura, se indica la configuración del modelo en equilibrio.

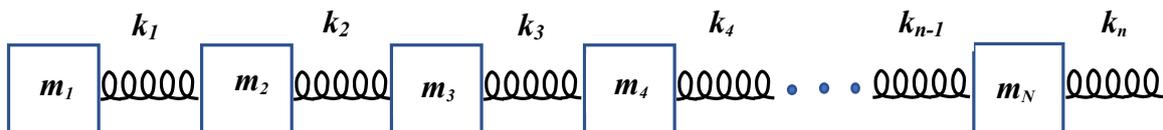


Figura 1: Medio de propagación como un sistema de resortes conectando sucesivas masas

Si el medio que se quiere modelar es homogéneo, podemos considerar que todas las m_i y k_i son idénticas a m y k respectivamente. Cuando uno deforma esa configuración de equilibrio, se produce una especie de efecto dominó: si se empuja un bloque hacia la derecha, el muelle de ese lado se comprime y empuja al consecutivo; y así sucesivamente como se indica en la figura 2.

La región donde los bloques son obligados a juntarse se denomina *compresión* mientras que la zona en que son empujados a separarse, recibe el nombre de *rarefacción*.

En medios reales, la propagación está basada en una *compresión* seguida de una correspondiente *rarefacción* para que vuelva a su estado normal y esto tiene lugar sucesivas veces.

Este modelo admite dos tipos de propagación de las perturbaciones del equilibrio:

- Si los bloques se mueven en la dirección de propagación, tendremos las llamadas *ondas longitudinales*, que son una buena representación del sonido en aire.
- Si se mueven en dirección perpendicular o lateral a la propagación, tendremos las *ondas transversales*, que, en acústica, solo ocurren cuando el medio es un sólido elástico.

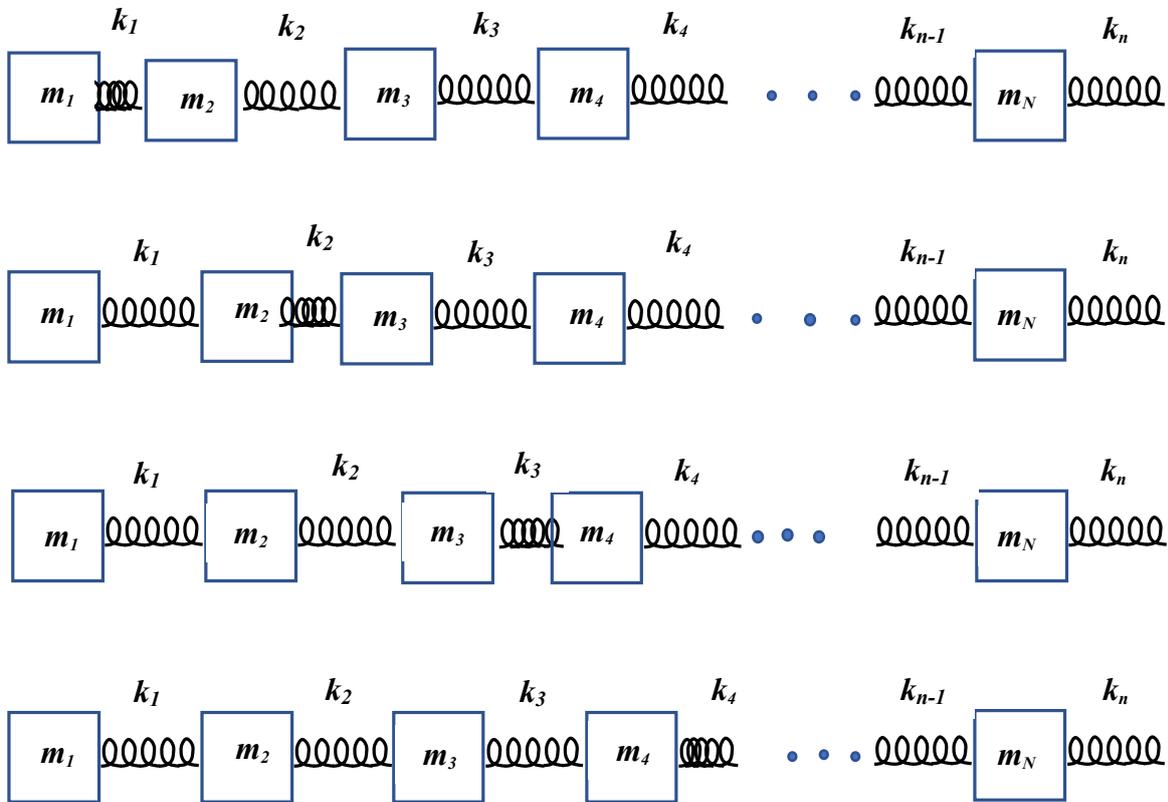


Figura 2: Propagación de una perturbación en el medio idealizado

La velocidad de propagación de estas ondas depende principalmente de dos factores:

- La masa de los bloques, que en materiales reales es proporcional a la densidad.
- La rigidez de los resortes, que en medios materiales reales y continuos es equivalente al módulo de Young E (módulo elástico).

En un medio elástico tridimensional, la velocidad de propagación, v , de las ondas longitudinales resulta ser:

$$v = \sqrt{\frac{E}{\rho}} \quad (1)$$

donde E es el módulo de Young o elástico ($\text{kg/m}\cdot\text{s}^2$) y ρ es la densidad del medio sólido (kg/m^3). La densidad es independiente a la dirección de propagación mientras que el módulo de Young puede serlo o no, dependiendo de si el medio es homogéneo e isótropo, o no lo es. Esta ecuación se puede demostrar fácilmente para el caso de una barra homogénea elástica como la indicada en la figura en la que consideramos deformaciones a lo largo del eje longitudinal.

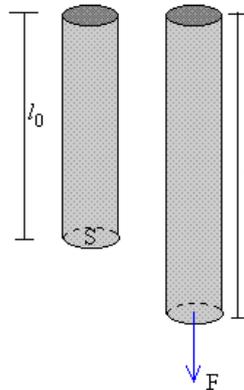


Figura 3: Barra homogénea elástica

El módulo de Young E mide la relación entre esfuerzo F/S aplicado y la deformación longitudinal que este produce en la barra:

$$\frac{F}{S} = E \frac{l - l_0}{l_0}$$

donde F es la fuerza aplicada, S es la sección de la barra y l_0 y l son su longitud inicial y final respectivamente.

Considerando el caso de la barra de sección S en la posición x , definimos como su anchura dx .

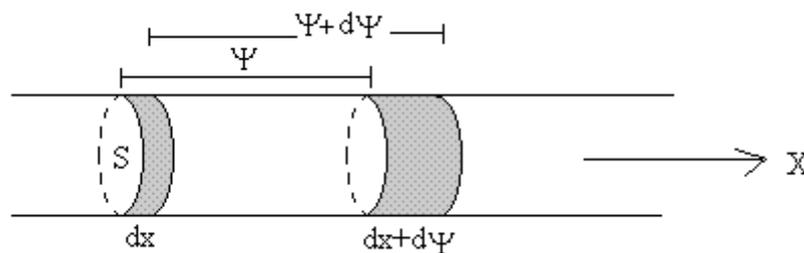


Figura 4: Deformación de la barra homogénea

A causa de la perturbación que experimenta, el elemento se desplaza ψ y se deforma $d\psi$, de modo que la nueva anchura del elemento es $dx + d\psi$. Entonces:

$$\frac{F}{S} = E \frac{dx + d\psi - dx}{dx} = E \frac{d\psi}{dx}$$

Por tanto, la fuerza neta es igual a:

$$F' - F = dF = SE \frac{\partial^2 \psi}{\partial x^2} dx = SE \frac{\partial^2 \psi}{\partial x^2} dx$$

Aplicando la segunda ley de Newton, igualamos esta fuerza a la masa por la aceleración que experimenta de manera que resulta lo siguiente:

$$dF = (\rho S dx) \frac{\partial^2 \psi}{\partial t^2}$$

Igualando ambas expresiones:

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{E}{\rho} \frac{\partial^2 \psi}{\partial x^2}$$

Resolviendo el sistema mediante separación de variables, obtenemos que la fórmula de velocidad de propagación para el movimiento ondulatorio longitudinal viene dada por la ecuación (1).

En el caso de la propagación de sonido en el aire, podemos considerar el medio como un gas sometido a compresión adiabática, de manera que de la ecuación de los gases en estas condiciones tenemos:

$$pV^\gamma = cte \quad (2)$$

En el caso del aire $\gamma=1.4$.

De aquí, puede determinarse el módulo de Young para un gas adiabático que viene dado como:

$$E = \gamma \cdot P \quad (3)$$

Y la densidad del mismo es

$$\rho_{gas} = \frac{m}{v} = \frac{PM}{RT}$$

Donde $R=8,31 \text{ J/mol}\cdot\text{K}$

Con estos datos, podemos afirmar que la velocidad del gas vendrá dada como:

$$v = \sqrt{\frac{E_{gas}}{\rho_{gas}}} = \sqrt{\frac{\gamma RT}{m}} \quad (4)$$

donde T es la temperatura del aire en K.

Para ilustrar esta dependencia de la temperatura, evaluamos esta ecuación para 0°C y 20°C :

$$v_{sonido \ 0^\circ\text{C}} = 332 \text{ m/s}$$

$$v_{sonido \ 20^\circ\text{C}} = 344 \text{ m/s}$$

El otro caso de movimiento en ondas es en aquella que la propagación es en dirección perpendicular al movimiento, y este tipo de ondas recibe el nombre de ondas transversales. Para determinar su velocidad, consideramos el ejemplo fácil de una cuerda en que su movimiento viene descrito mediante esta figura:

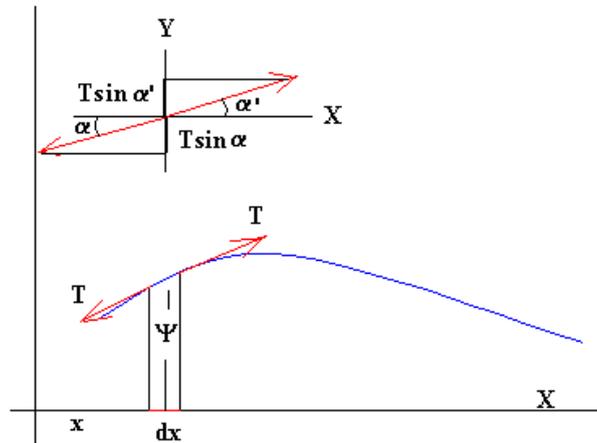


Figure 5: Movimiento descrito por una cuerda considerada como un movimiento ondulatorio transversal

Entonces, haciendo uso de la famosa segunda ley de Newton, aplicándola en dirección tangencial al movimiento, se llega a que:

$$T (\sin \alpha' - \sin \alpha) = m \cdot a$$

En caso de que estos respectivos ángulos no sean muy grandes, pueden considerarse como si fuesen funciones tangenciales.

$$T (\operatorname{tg} \alpha' - \operatorname{tg} \alpha) = T d(\operatorname{tg} \alpha) = T \frac{\partial(\operatorname{tg} \alpha)}{\partial x} dx = T \frac{\partial^2 \psi}{\partial x^2} dx = \mu dx \frac{\partial^2 \psi}{\partial t^2}$$

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{\mu}{T} \frac{\partial^2 \psi}{\partial t^2}$$

Este sistema puede resolverse de manera sencilla a partir del método de separación de variables, haciendo servir los conocimientos en ecuaciones diferenciales para resolver las dos ODEs que se forman en el proceso. De esta manera se llega a que la velocidad de propagación en ondas transversales es la presente:

$$v = \sqrt{\frac{T}{\mu}} \quad (5)$$

Un término de gran importancia en la acústica es la impedancia, que es debida al movimiento de las partículas (fuerzas de compresión y aceleración) y por la presión de las partículas. Esta impedancia acústica se puede expresar matemáticamente de la siguiente manera:

$$Z_{acust} = \frac{p}{U} \quad (6)$$

Es análogo al caso eléctrico en que $R=V/I$. Aquí esta impedancia acústica actúa como la resistencia, la presión es como una tensión y la velocidad por su parte se trata de una especie de intensidad que se va propagando.

$$Z_{acust} = \sqrt{\rho E} = \sqrt{\rho^2 \frac{E}{\rho}} = \sqrt{\rho^2 c^2} = \rho c \approx 416 \frac{kg}{m^2 s} \quad (7)$$

Otro parámetro importante en el estudio es la intensidad del sonido, que suele expresarse generalmente en escala logarítmica:

$$SIL = 10 \log \left(\frac{I_{actual}}{I_{ref}} \right) \quad (8)$$

cabe destacar que la $I_{ref} = 10^{-12} \text{ W/m}^2$. También debe destacarse la potencia del sonido, que suele venir dada en decibelios para que la escala sea más sencilla.

$$SWL = 10 \log \left(\frac{W_{actual}}{W_{ref}} \right) \quad (9)$$

donde $W_{ref} = 10^{-12} \text{ W/m}^2$, que es la potencia de referencia. Finalmente, el tercer factor que suele considerarse es la presión del sonido, que oscila entre $20 \mu\text{Pa}$ y 20 Pa para 1kHz

$$SPL = 10 \log \left(\frac{p_{actual}}{p_{ref}} \right) \quad (10)$$

$$I_{acust} = Up = \frac{p}{Z_{acust}} p = \frac{p^2}{Z_{acust}} \quad (11)$$

$$p_{ref} = \sqrt{Z_{acust} I_{ref}} = 20.4 \cdot 10^{-6} \text{ Pa} \quad (12)$$

Si se producen varios sonidos simultáneamente, ya sea por diferentes instrumentos o por reflexiones con las paredes, encontramos dos tipos de sonidos:

- Correlativos: diferentes fuentes relacionadas
- No Correlativos: diferentes fuentes no relacionadas (dos instrumentos distintos o efecto considerable de reflexiones). Estos dan lugar a ondas primarias, que son ondas adicionales.

La presión de correlación viene dada como:

$$p_{tot\ correl}(t) = p_1(t) + p_2(t) + p_3(t) + \dots + p_N(t) \quad (13)$$

La presión en un punto viene dada como:

$$p_{en\ un\ punto}(t) = p_{amplitud\ del\ sonido} \sin(2\pi ft)$$

Si hay ondas desfasadas, una respecto de otra con un cierto retraso τ , encontramos que la presión en ese caso es:

$$p_{en\ un\ punto}(t) = p_{amplitud\ del\ sonido} \sin(2\pi f(t+\tau))$$

De esta manera la presión total es la suma de las ondas desfasadas y normales, es decir:

$$p_{tot\ correl}(t) = p_{no\ retrassin}(2\pi ft) + p_{retrassin}(2\pi f(t+\tau))$$

$$p_{tot\ correl}(t) = p_1(t) + p_2(t) + p_3(t) + \dots + p_N(t) \quad (14)$$

Para el caso de la no correlación, la presión en este caso viene dada de esta manera:

$$p_{tot \text{ sin correl}}(t) = \sqrt{p_1^2(t) + p_2^2(t) + p_3^2(t) + \dots + p_N^2(t)} \quad (15)$$

$$p_{tot} = \sqrt{Np^2} = \sqrt{N}p \quad (16)$$

Las ondas de sonido pueden experimentar una serie de fenómenos ligados a su carácter ondulatorio:

- Superposición: Interferencia destructiva entre diferentes ondas
- Refracción: La velocidad de la luz varía según el material que sea
- Reflexión: El sonido cambia el sentido de propagación al encontrar una pared
- Absorción: Interacción con un objeto físico, en que hay una transferencia de energía de la onda al objeto por fricción.

Para nuestro propósito no necesitaremos abundar en estos fenómenos por el momento.

1.2 Fisiología auditiva

Antes de entrar en el tema que nos concierne, la psicoacústica, es conveniente introducir algunas nociones de la fisiología del oído o aparato auditivo humano, que es el conjunto de órganos encargado de transformar los sonidos en señales nerviosas que el cerebro pueda procesar para generar la percepción acústica.

Anatómica y funcionalmente, es conveniente separar al oído en tres secciones bastante definidas:

- El oído externo (oreja y conducto), formado por una concha de entre 25 y 35 mm. que dirige el sonido ambiente hasta la membrana timpánica.
- El oído medio, constituido por osículos que capturan los movimientos del tímpano y los transmite al órgano del oído interno conocido como cóclea.
- El oído interno, que es el responsable de transformar las vibraciones mecánicas en impulsos nerviosos.

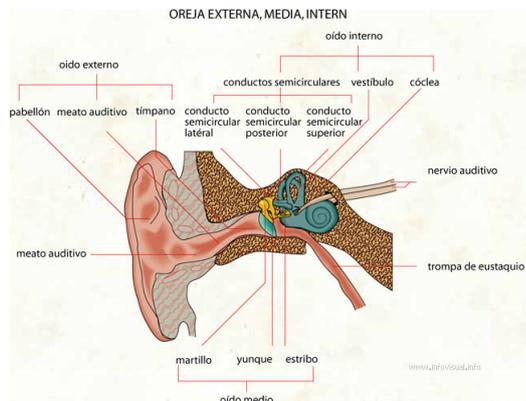


Figure 6: Representación del oído con sus tres partes

1.2.1 Oído externo

En la Figura 7 se muestra un esquema del oído externo. La función principal de esta sección del oído es modificar la respuesta de la frecuencia de sonidos, y su componente más importante es el tímpano que se compone de 3 capas: la externa (mucosas), la media (responsable de fuerza y elasticidad de la oreja) y la interna (pasa la presión como amplitud).



Figure 7: Muestra del oído externo discerniendo todas sus componentes

1.2.2 Oído medio

El oído medio contiene tres pequeños huesos interconectados mecánicamente, denominados respectivamente, martillo, yunque y estribo el primero apoyado sobre la membrana timpánica y el último pegado a la membrana oval de la cóclea. El oído tiene, esencialmente, dos funciones principales:

- Transmitir movimientos del tímpano, al oído interno.
- Proteger el aparato auditivo de los efectos de sonidos demasiado fuertes.

Su cadena de huesecillos funciona como un transformador de impedancia mecánica, entre la de la membrana del tímpano, relativamente de gran área y masa y la pequeña ventana (ventana oval) en el extremo la cóclea –el órgano principal del oído interno– de área mucho menor que la anterior. Su estructura se puede ver representada en la Figura 8:

La ecuación que rige la transformación de impedancias en el oído medio es:



Figure 8: Muestra del oído medio con cada una de sus partes.

$$F_1 x l_1 = F_2 x l_2 \quad (17)$$

donde F_1 es la fuerza de la membrana timpánica, l_1 es la longitud del martillo, F_2 es la fuerza del estribo de la plataforma y l_2 es la longitud del yunque. Esto puede ponerse en función a las presiones, de manera que la expresión anterior se ve reescrita como:

$$p_2 x A_2 = p_1 x A_1 x \frac{L_1}{L_2} \quad (18)$$

1.2.2 Oído interno

Finalmente, el oído interno, que es el que contiene la estructura de forma de caracol conocida como cóclea, tiene por función transformar vibraciones mecánicas en disparos nerviosos procesados posteriormente en el cerebro. Esta sección del oído es la más importante a nivel del estudio acústico que nos concierne, ya que como veremos posteriormente en detalle, es donde se encuentra la membrana basilar donde tiene lugar el filtro de Fourier de los tonos complejos y está como veremos muy relacionado con la altura percibida y varios fenómenos relacionados con ésta. La Figura 9 muestra un diagrama de la cóclea indicando sus diferentes partes anatómicas.



Figure 9: Muestra del oído interno con cada una de sus partes.

Se trata de una estructura tubular enrollada en forma de caracol a lo largo cuyo interior encontramos las denominadas membranas de Reissner y basilar que dividen al tubo principal, en dos canales externos, uno “de entrada” llamado *escala vestibular* (V) otro de salida, *escala timpánica* (T), y entre ambas membranas, un tercero llamado *escala media* (M). Todo el interior de la cóclea está lleno de un fluido incompresible que se denomina perilinfa. Así caracol o cóclea, es un tubo con paredes que lo separan en tres cavidades como son las siguientes. La Figura 9 muestra una sección de dicho tubo indicando en detalle dichas cavidades.

En uno de los extremos de la escala timpánica se encuentra *ventana oval* que recibe las vibraciones transmitidas por el oído medio. La escala vestibular, por su parte, termina en una pequeña membrana más flexible que el tejido circundante, llamada *ventana circular* cuya función es compensar la presión ejercida por la membrana oval sobre el fluido. Ambos canales se encuentran conectados en el extremo del tubo (centro del caracol) denominado, *ápex*. La *membrana basilar* que separa las escalas vestibulares y media contiene el denominado *órgano de Corti*. La membrana basilar disminuye progresivamente su rigidez hacia su extremo del final en el *ápex* la rigidez se hace menor, con lo cual disminuye la tensión y por consiguiente su frecuencia resonante mecánica.

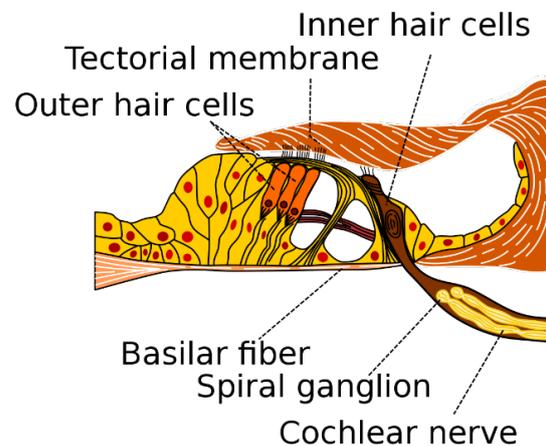


Figure 10: Estructura de la membrana basilar.

Cada sección de la membrana basilar responde mecánicamente a una frecuencia específica en un rango que varía entre 20 y 20000 Hz en el que se sitúan las frecuencias audibles dependiendo de la edad que presente el individuo. La localización espacial de esta respuesta mecánica no es, sin embargo, tan aguda como la resolución con la que el oído es capaz de distinguir tonos de frecuencias diferentes. Hay un detalle llamativo relativo a esto: algunos experimentos psicoacústicos, muestran fenómenos diferentes si los estímulos recibidos en el oído son de frecuencias próximas o no. Si presentan frecuencias cercanas, excitan casi el mismo punto de la membrana basilar, mientras que, si esa diferencia es mayor, se estimulan dos regiones independientes de esa membrana. El rango de frecuencias en que sus curvas de respuesta en la membrana basilar se solapan, recibe el nombre de *banda crítica*.

1.4 Análisis de Fourier

En la membrana basilar, tiene lugar la transducción de los sonidos complejos mediante un proceso que recuerda al análisis de Fourier. Tal proceso permite transformar dichos sonidos complejos en tonos simples o “puros”, y su implementación coclear recibe el nombre de *filtro coclear* de Fourier o más descriptivamente *banco de filtros coclear*. Los componentes de una única frecuencia en un sonido complejo se denominan parciales. Estos aparecen como picos

agudos (idealmente deltas de Dirac) en el espectro de Fourier del sonido complejo como veremos más adelante. La psicología de la percepción del sonido parece corresponder en varios aspectos al análisis Fourier. El estudio de los fenómenos de percepción del sonido conforma la disciplina llamada *psicoacústica*. Esta incluye, por ejemplo, la capacidad auditiva para comparar tonos, intensidades y timbres de los sonidos, entre otros fenómenos y cualidades.

La percepción del sonido y de la música es tal que, en ciertas ocasiones, los humanos pueden distinguir los armónicos de una onda de sonido periódica compleja. Esto se fundamenta en el sentido de la audición, que proporciona a los seres humanos una serie de datos que son relativos al entorno.

De manera más específica, cada frecuencia está vinculada a un filtro, y de este modo transforma tonos complejos en puros. Por tono puro entendemos, es una onda sinusoidal de frecuencia o periodicidad, amplitud y fase definidas:

$$x(t) = A \sin(2\pi f t - \phi) = A \sin\left(\frac{2\pi t}{T} - \phi\right) = A \sin(\omega t - \phi) \quad (19)$$

donde A es la amplitud, f su frecuencia y ϕ su fase. El período de esta señal simple $T = 1/f$; o a y su denominada frecuencia angular es $\omega = 2\pi/T = 2\pi f$.

En grado creciente de complejidad tenemos las señales sonoras periódicas, cuya magnitud satisface la relación $x(t + nT) = x(t)$ con $n = 1, 2, 3, 4, \dots$ para todo tiempo t .

Estas señales pueden representarse mediante una serie de Fourier:

$$x(t) = x_0 + \sum_{i=1}^N (a_n \cos n\omega t + b_n \sin n\omega t) \quad (20)$$

que podemos interpretar como la superposición de tonos puros de frecuencias múltiplos de una fundamental $f_0 = 2\pi/\omega$. a_n y b_n son los correspondientes coeficientes de Fourier. El espectro de onda periódica es, entonces, una sucesión de líneas o picos espectrales que, usualmente, en la frecuencia fundamental f_0 y sus múltiplos sucesivos $2f_0, 3f_0, \dots$ denominados *armónicos*. Los sonidos periódicos cuyas componentes de Fourier son solamente la fundamental y sus armónicos se denominan complejos armónicos y serán las piezas claves del experimento de percepción que llevaremos adelante en este trabajo. En particular, consideraremos complejos armónicos en los cuales habrá presentes solo un número finito de armónicos superiores de una frecuencia fundamental f_0 pero no la frecuencia fundamental misma.

1.5 Psicoacústica

En el pabellón auditivo se lleva a cabo una gran cantidad de procesos relacionados con el sonido pero que acaban en un conjunto de impulsos nerviosos que son interpretados por el cerebro en términos de ciertas categorías. La parte de la acústica que recibe el nombre de psicoacústica se encarga del estudio del sonido en término de estas categorías, tales como el *volumen*, la altura tonal –o *altura*, a secas–, el *timbre*, etc. De estas, la que concierne directamente a nuestro experimento es la altura, cuyas características veremos en detalle más adelante.

Una característica importante de los sonidos es su **volumen**, esencialmente proporcional a la potencia sonora que llega al oído en relación umbral de sensibilidad de dicho oído el cual varía según el individuo e incluso la edad. Este parámetro cambia también con la frecuencia del tono,

si se trata de un tono puro. La apreciación del volumen depende también de si el sonido es monoaural (la señal sonora llega a un solo oído) o binaural (la señal llega a ambos). También influyen de manera considerable los componentes psicológicos o del ambiente; por ejemplo, si el individuo está en una conversación o en solitario, pues en ciertos casos ignora o rebaja el volumen de algunos sonidos, para oír mejor aquello que desea escuchar.

Otras dos cualidades perceptivas fundamentales del sonido son, como dijimos, la altura y el timbre. Ambas características dependen del contenido espectral de los sonidos, pero es necesario distinguirlas cuidadosamente:

La **altura** tonal, es esencialmente el lugar que le asignamos a un determinado sonido en la escala musical, es decir la *nota* musical que asociamos a ese sonido. Para tonos puros, dicha altura coincide rigurosamente con la frecuencia física del tono que se trata. Esta, por tanto, nos indica la tonalidad del sonido: frecuencia baja corresponde a un sonido grave mientras que una frecuencia alta a los agudos. Si se trata de un sonido armónico (una señal periódica de forma arbitraria), la altura percibida también corresponde rigurosamente a la frecuencia fundamental de la señal.

El **timbre** es la otra propiedad destacable del sonido que está asociada de alguna manera al espectro de frecuencias del sonido escuchado. El timbre es la característica que nos permite distinguir sonidos de dos instrumentos diferentes, aunque correspondan a la misma altura o nota musical. Aunque la percepción del timbre es un fenómeno muy complejo de analizar, en primera aproximación y para complejos armónicos podemos decir que el timbre es una representación de las amplitudes relativas de los diferentes armónicos que los componen. En nuestro estudio particular sólo encontraremos muy circunstancialmente fenómenos relacionados con el timbre

Un ejemplo simple aclaratorio de la diferencia entre tono y timbre es el siguiente: si se tocan dos notas de un mismo instrumento como pueden ser un “do” y un “re”, estas presentan el mismo timbre pero distinto tono, mientras que si se compone la misma nota en dos instrumentos diferentes, presentarán el mismo tono pero distinto timbre.

Otra propiedad característica es la **duración**, es otro factor que afecta de manera considerable al sonido, ya que no interpretamos igual tonos puros (sonidos de más de 200 ms) a tonos complejos derivados del carácter finito de la ventana temporal durante la que ese sonido es emitido.

1.5 Percepción del pitch

Para el presente trabajo de fin de grado, un concepto que es de gran importancia es el **pitch** o la **altura**, que, como dijimos, es uno de los tres conceptos esenciales de la psicoacústica junto al timbre y a la intensidad.

El pitch es un atributo de las ondas sonoras que está relacionado con las frecuencias de los sonidos. Frecuencias altas implican una altura “alta” o aguda, mientras que las bajas corresponden a una altura “baja” o grave. La percepción de esta altura tiene lugar en la membrana basilar del oído interno y en varias etapas de procesamiento de componentes de la frecuencia los centros superiores del procesamiento neuronal. Las funciones “matemáticas” básicas de ese procesamiento pueden considerarse como alguna combinación de las siguientes:

- Localizar la componente de frecuencia fundamental f_0

- Encontrar la diferencia de frecuencia entre los armónicos que viene dada entre el armónico f_n y el subsiguiente f_{n+1} , que a partir de la definición podría darnos también la frecuencia fundamental: $(n+1)f_0 - nf_0 = f_0$
- Determinar el máximo común divisor entre todas las frecuencias del complejo al cual asignaríamos el carácter de frecuencia fundamental.

En sonidos armónicos la percepción de la altura está ligada principalmente a f_0 . Luego, la estimación de f_0 se lleva a cabo principalmente por la posición y la teoría temporal de la percepción del sonido.

Cabe destacar que hay dos factores que afectan a la altura percibida: la intensidad y la duración del sonido. Un aumento de la intensidad de un tono puro lleva implícita una disminución de la altura percibida a frecuencias bajas o un aumento de dicha altura a frecuencias altas. Y la duración tiene también una importancia notable ya que tonos cuya duración va de 10 y 60 ms dependiendo de la frecuencia y la intensidad, se perciben con alturas diferentes a las del tono continuo.

Dos magnitudes a tener en cuenta en el estudio de la percepción de la altura son la JND (*just noticeable difference* o diferencia apenas notable) y DL (diferencial limen o umbral de diferencia) que se refieren respectivamente al cambio más pequeño perceptible de una variable física y a los límites en los valores de una variable física más allá de los cuales la variable pierde su identidad perceptiva.

La JND en el caso del pitch se encuentra aproximadamente entre 0.3 y el 1% de la frecuencia, según el registro (es decir, en la región de frecuencia, que observamos en la Figura 11). Dicho de otro modo, corresponde a aproximadamente 1/30 de la banda crítica o 1/12 de un semitono de la escala equitemperada:

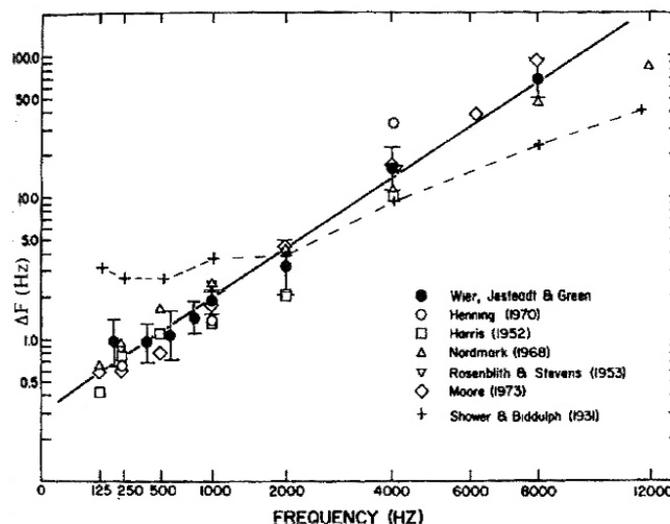


Figura 11: Dependencia del JND con la altura.

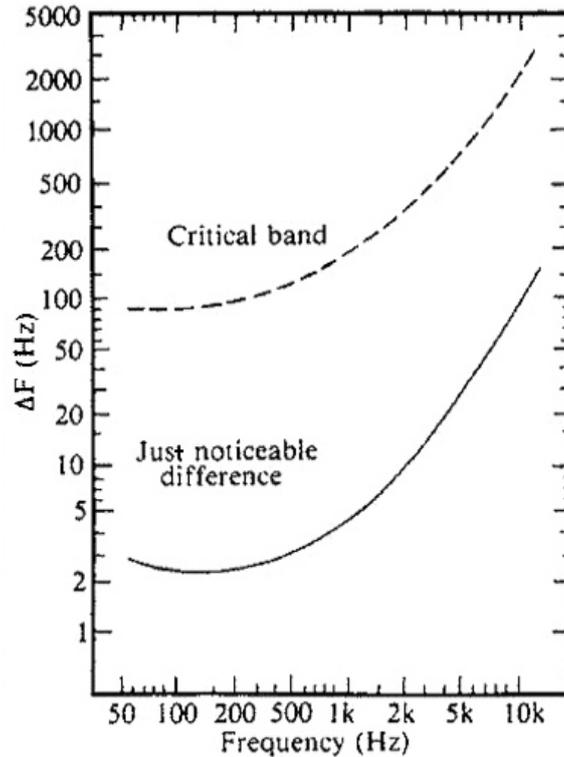


Figura 12: Comparativa del JND con la banda crítica.

Las Figuras 11 y 12 se han extraído del capítulo de Pitch of complex tones. Journal of the Acoustical Society of America 41(6), 1526–1533.

La cosa se complica al hablar de altura en tonos complejos ya que su dependencia con la frecuencia, la intensidad y la duración si bien algo similar al de los tonos puros aparecen fenómenos nuevos de explicación mas elaborada. Por ejemplo, la JND de la altura puede ser mayor o menor que en los puros dependiendo del contenido espectral del complejo.

1.5.1 Fenómeno de la “missing fundamental”

Un efecto no intuitivo en la percepción de sonidos complejos, que es de particular importancia para el estudio que realizamos en este trabajo, es el de la altura percibida de complejos armónicos en los que la fundamental y algunos de sus sucesivos armónicos han sido suprimidos. En general, para tonos complejos con espectro armónicos (múltiplos de la frecuencia fundamental f_0), la altura, en general, suele corresponder a la de dicha frecuencia fundamental. Lo curioso es que si esa frecuencia se suprime del espectro la altura percibida sigue siendo la misma, aunque dicha frecuencia ya no esté presente. Este fenómeno es lo que se conoce como *altura de la fundamental ausente (missing fundamental pitch)*.

En general, el fenómeno del pitch de la fundamental ausente (también referido como *altura virtual*) es el origen de muchas controversias entre las diversas teorías de la percepción de la altura y proporciona evidencias de que la altura no sólo depende la frecuencia, la intensidad y la duración, sino que también de la distribución espectral. La Figura 13 indica esquemáticamente un espectro típico de un sonido armónico con fundamental y primer armónico suprimido. En este caso, la altura percibida continúa siendo la correspondiente a un tono puro de 200 Hz a pesar de que a dicha frecuencia la señal no tiene ninguna energía.

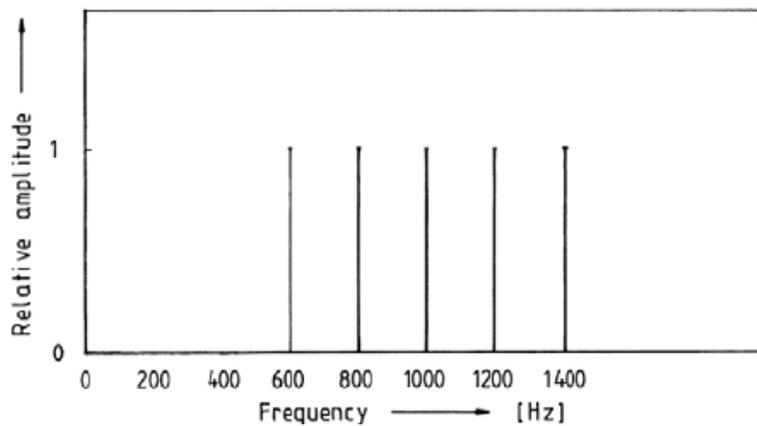


Figure 13: Esquema de un espectro armónico con la fundamental y primer armónico faltantes.

Las señales complejas armónicas se perciben como una unidad en lugar de un conjunto de múltiples tonos puros. Sin embargo, los armónicos individuales se pueden “escuchar”, si interrumpimos su presencia a intervalos y los volvamos a introducir seguidamente. Más complicado aún es el caso en que los tonos complejos son con espectros inarmónicos, es decir, cuyas componentes de frecuencias puras no son múltiplos enteros de una fundamental. En estos casos puede ocurrir que provoquen una sensación de altura clara y única. Pueden provocar más de una sensación de altura o pitch en mutua competencia, que aparentemente podrían parecer acordes, o pueden sonar como ruido, dependiendo de su distribución espectral.

Históricamente se desarrollaron dos marcos teóricos para la explicación de varios de los fenómenos de la percepción de altura: la teoría del lugar (place-theory) y la teoría temporal (time-theory) del que se mantuvieron en mutua competencia hasta recientemente cuando ambos han sido parcialmente superadas por la tendencia a asignar la detección del pitch a circuitos de procesamiento a nivel cortical central.

1.5.2 Teoría del lugar

La teoría del lugar tiene una larga historia que abarca desde de los días de Helmholtz (1954) a ideas que aún siguen considerándose en la actualidad (una de las últimas de Cheveigne 2004).

Según la teoría del lugar de resonancia del oído, el interior el oído actúa como un órgano analizador de frecuencias, y el estímulo que llega a nuestro oído se descompone en muchos componentes sinusoidales, cada uno de los cuales excita diferentes lugares a lo largo de la membrana basilar, donde las células ciliadas con distintas frecuencias características se encuentran vinculadas con las neuronas. También sugirió que la altura de un estímulo está relacionada con el patrón de la excitación producida por el estímulo a lo largo de la membrana basilar. Para un tono puro, la altura generalmente corresponde a la posición de máximo de excitación; mientras que, para un tono complejo con muchos picos espectrales, es más complicado.

Entre los partidarios de la teoría del lugar, Goldstein, Terhardt y Wightman son los más famosos. El mayor problema de la teoría del lugar es que falla la identificación del tono de un estímulo con fundamental faltante. Según la teoría de Helmholtz, es imposible percibir el pitch cuando no hay un pico espectral en la posición a lo largo de la membrana basilar que corresponde a la frecuencia de un tono. No obstante, hay un tono pronunciado en la fundamental faltante. Schouten (Schouten 1938) hizo experimentos extensivos sobre este problema usando

un pulso periódico donde descubrió una altura baja asociada con componentes armónicos altos de un estímulo y llamó a esta percepción un "tono de residuo".

1.5.3 Teoría temporal

Mientras que la teoría del lugar trata de explicar la sensación de la altura al encontrar lugares en la membrana basilar donde se encuentra la excitación del estímulo en que es máximo, la teoría temporal es un mecanismo de dominio del tiempo que se basa en eventos; es decir, intenta detectar el intervalo de tiempos entre eventos, que pueden ser picos o envolvente general de las formas de onda de entrada. Estos eventos determinan la periodicidad de la forma de onda y el recíproco de la periodicidad que es la misma que la frecuencia fundamental. En su lanzamiento de residuos de la teoría, Schouten (Schouten 1938; Schouten, Ritsma y Cardozo 1962) mencionó el importante papel de los altos armónicos no resueltos. Según él, se produce sensación de tono de un tono complejo cuando una interacción de esos varios armónicos no resueltos, resulta en un patrón de tiempo periódico de la forma de onda, y la inclinación del residuo está determinada por la periodicidad.

La teoría temporal como la de Schouten puede explicar los fenómenos que la teoría del lugar no logra interpretar. El primero de ellos es el problema de la altura de la *fundamental ausente* o de la "missing fundamental". Debido a la echada del residuo no corresponde a ninguno de los componentes físicos sinusoidales del estímulo, pero está determinado por el patrón de tiempo total causado por una interacción de varios armónicos, el fundamental puede estar físicamente ausente y aún así invocar la sensación de tono.

Además, la teoría temporal también explica la sensación de tono de ruido interrumpido sin picos espectrales. Sin embargo, el papel de los armónicos no resueltos, que era esencial en la teoría del residuo, resultó errónea por los hallazgos de Plomp (Plomp 1967) y Ritsma (Ritsma 1967). En sus estudios, encontraron una llamada región dominante cubierta por las componentes de frecuencia de los armónicos tercero, cuarto y quinto, y propusieron que altura de un tono complejo esté determinada por la región espectral dominante, donde los armónicos son, obviamente, resolubles.

2. Nuestro Tema: Percepción binaural de la altura virtual.

Una de las dificultades mayores que encuentran las teorías físicas de la percepción de la altura virtual o residuo, descritas en la introducción, es el hecho sorprendente de que esta se percibe aún si los armónicos superiores que componen al estímulo de portadora suprimida, o los parciales, en el caso de estímulos no armónicos, se presentan en forma dicótica, es decir, algunos en un oído y otros en el otro. Cualquier fenómeno puramente mecánico o físico que tuviera lugar en cada uno de los oídos y diera cuenta de la aparición de la fundamental ausente, no podría explicar su percepción en la presentación diótica. Esto da vida a hipótesis de que esta percepción ocurre gracias a un procesamiento central posterior a la detección física del sonido. Sin embargo, hasta el momento, no se ha realizado un estudio exhaustivo de la comparación directa de la altura de la fundamental ausente percibida de sonidos similares en ambos tipos de presentación (dicótica vs. diótica). El objetivo del presente Trabajo de Fin de Grado es investigar experimentalmente la percepción de la altura de complejos armónicos presentados en forma binaural comparando esos dos tipos de estimulación de ambos oídos: diótica y dicótica. Tal investigación intenta validar la hipótesis preliminar, concluida a partir de ensayos individuales aislados, de que hay una diferencia sistemática en las alturas percibidas en ambas situaciones.

Antes de describir el diseño experimental desarrollado y los resultados obtenidos, introduzcamos algunos conceptos, definiciones y efectos conocidos de la percepción binaural en general.

2.1 Fenómenos binaurales:

El término *binaural*, hace referencia a los fenómenos de la audición que conciernen a la estimulación diferenciada en ambos oídos, por oposición *monoaural* que se refiere a la estimulación de los oídos a partir de una única fuente. Una de las funciones fundamentales de la binauralidad es la detección direccional de la fuente de un sonido. Dado que las orejas están separadas unos 18 cm se produce una diferencia de camino considerable entre el sonido que llega a una y a la otra, en función de la dirección desde donde proviene el sonido en cuestión. Esta diferencia se ve reflejada en diferencias de fase (o tiempo) e intensidad entre las vibraciones detectadas por cada oído. Se pueden definir estas diferencias como ITD (diferencia de tiempo binaural) y IID (diferencia de intensidad binaural).

ITD: debido a la distancia entre las orejas esta diferencia tiempo depende de la distancia d entre las orejas y del ángulo θ hacia la fuente sonora con la siguiente relación:

$$\Delta t = \frac{d \sin \theta}{c} \quad (20)$$

donde c es la velocidad del sonido. Pero esta ecuación no tiene en cuenta que, en realidad, el sonido no se propaga en línea recta, sino que esquiva la cabeza como un obstáculo. Considerando este efecto en una cabeza “esférica”, se llega a que:

$$ITD = \frac{r(\theta + \sin \theta)}{c} \quad (21)$$

y la diferencia de fase correspondiente a este retardo es:

$$\phi_{ITD} = 2\pi f r(\theta + \sin \theta)/c$$

IID: El aparato auditivo percibe junto a esta diferencia de fase diferencias de intensidad entre los dos oídos y las procesa para determinar de qué lado viene el sonido.

Los efectos binaurales son de capital importancia en las técnicas de reproducción de sonido. El estéreo binaural crea una ilusión direccional mediante el uso de dos micrófonos omnidireccionales situados en los oídos de un modelo de cabeza. La distancia entre micrófonos es, de manera aproximada, la de las orejas y se colocan en un objeto que sombrea el sonido, de manera que un cambio en la posición del músico modifica la intensidad y el retraso entre las señales registradas por cada micrófono.

Como adelantamos mas arriba, dentro de la estimulación binaural, conviene distinguir dos versiones, la estimulación diótica y la dicótica.

- La estimulación **diótica** es aquella en que se hace escuchar la misma combinación de una, dos o más fuentes de sonido con ambos oídos simultáneamente.

- La estimulación **dicótica**, por el contrario, se refiere al estado generado artificialmente al escuchar un sonido o combinación de sonidos diferente en cada oído, como es el caso por ejemplo enviar sonidos de diferente frecuencia a cada oído mediante auriculares.

Dentro de los fenómenos relacionados directamente con la estimulación con tonos puros de diversas frecuencias, está el sorprendente efecto de los batidos binaurales y más propiamente de los batidos dicóticos. Cuando dos tonos puros de frecuencia muy parecidas, $f + \Delta$ y $f - \Delta$, se superponen físicamente, por ejemplo cuando son producidos simultáneamente en la misma fuente, o cuando llegan al oído simultáneamente desde dos fuentes diferentes, tiene lugar el efecto de “batido” físico que podemos entender fácilmente desde la trigonometría:

$$\sin(f + \Delta)t + \sin(f - \Delta)t = \frac{1}{2} \cos\Delta t \sin ft$$

Físicamente el estímulo es una senoide de frecuencia f cuya amplitud está modulada por otra de frecuencia Δ mucho menor. Si Δ es del orden de una fracción de Hz, se puede escuchar este fenómeno como un tono “puro” pero pulsante. Lo curioso, sin embargo, es que, si se presenta una de las frecuencias en un solo oído y la otra en el otro, el batido, ahora virtual y no físico, se percibe igual, de ahí el nombre de *batido dicótico*.

El fenómeno del batido dicótico se extiende a las consonancias desordenadas, BMCs que son el resultado de la presentación de dos senoideas en frecuencias ligeramente diferentes de las que corresponderían a una relación racional con enteros pequeños entre ellas. Algunos estudios han sugerido que la fuente de BMCs dicóticas es una interacción presente en la banda crítica binaural. En un caso, el mecanismo se ha explicado como consecuencia de que un armónico auditivo del tono de baja frecuencia, que llamaremos f_1 crea batidos binaurales con el tono de alta frecuencia denotado como f_2 . La otra explicación implica una correlación cruzada binaural entre el patrón de excitación de f_1 y el contralateral f_2 que ocurre dentro de la banda crítica binaural centrada en f_2 . Un estudio examinó la detección de BMCs dicóticas para la octava y la quinta. En un experimento con la octava, el ruido de banda estrecha centrado en f_2 se presentó junto a f_1 y la otra oreja se presentó con f_2 . El ruido se usó para prevenir interacciones en la banda crítica binaural centrada en f_2 . Las BMC dicóticas aún se detectaron en estas condiciones, lo que sugiere que la interacción binaural en una banda crítica no explica el efecto. También se observaron efectos de localización bajo esta condición de enmascaramiento para las inversiones de fase de los estímulos de octava dicóticos sintonizados. Estos hallazgos sugieren una nueva teoría de BMC dicóticos como efecto de fase entre canales. En cualquier caso, la existencia de este tipo de batidos da cuenta de la complejidad neurológica que subyace a la percepción binaural.

Otra categoría interesante de fenómenos binaurales, aunque algo alejados de nuestro interés particular en este trabajo, se relaciona con la percepción mensajes hablados. Entre ellos se cuenta el del llamado “cocktail party effect” estudiado clásicamente desde 1953 por Colin Cherry. Se trata de la habilidad del aparato auditivo de reconocer mensajes hablados en medio de un entorno ruidoso donde se cruzan multitud de otros mensajes descorrelacionados. Por ejemplo, durante una reunión multitudinaria en la que cada huésped se involucra en su conversación individual, y a pesar de que a su oído llegan las voces de todo el salón, es capaz de reconocer el mensaje particular sobre el que presta atención. En conexión con nuestro tema, Cherry realizó un conjunto de experimentos en los que se cuantifica la habilidad de discernir dicóticamente dos conversaciones diferentes presentadas a sendos oídos de un mismo individuo.

2.2 Altura virtual dicótica vs diótica

El fenómeno de la percepción del residuo, altura virtual o altura de la fundamental ausente, se ha investigado con cierta profundidad en el contexto del problema de la percepción de altura de sonidos complejos, particularmente como método de validar o refutar las teorías físicas del “pitch”. En particular, en contra de especulación teórica sobre la influencia de fenómenos no lineales en el interior del oído, se ha presentado el hecho de que el fenómeno de la percepción de la fundamental faltante tiene lugar también cuando los armónicos que sí están presentes físicamente en el complejo se presentan en forma dicótica a los sujetos interrogados. También se sabe que el “pitch shift” (desplazamiento de altura) que ocurre cuando el conjunto de armónicos presentes se desplaza rígidamente en el eje de las frecuencias, dando lugar a complejos no armónicos, también ocurre si parte de esos parciales se presentan a un oído y la otra parte al otro.

Sin embargo, a nuestro conocimiento, no existen experimentos directos reportando cuan precisa es la coincidencia de la altura virtual (la percepción de la fundamental ausente) entre ambos tipos de presentaciones. En experimentos preliminares realizados con pocos individuos y en condiciones psicoacústicas no controladas, poco óptimas y por lo tanto de resultados no concluyentes, parecen mostrar una desviación sistemática “hacia arriba” de algunos Hz de la altura percibida en forma dicótica de complejos relativamente simples formados por dos parciales armónicos de tercer y cuarto orden de una frecuencia virtual dada.

La motivación de este TFG fue intentar proveer evidencia estadísticamente más significativa a favor o en contra de esta hipótesis. En efecto, la parte principal de este proyecto consistió en el desarrollo de un diseño experimental psicoacústico que permita interrogar en forma sistemática a un conjunto suficiente de sujetos sobre la percepción relativa de la altura apreciada en ambos estímulos y la realización de experimentos preliminares con este entorno.

2.3 Descripción del experimento.

Como hemos dicho más arriba, el objetivo principal del presente trabajo es comparar la altura virtual de sonidos armónicos con fundamental suprimida entre la presentación dicótica y la diótica. En particular, se trata de interrogar a los sujetos sobre la coincidencia o no en la altura de ambas presentaciones, de un sonido “complejo” compuesto por los armónicos tercero y cuarto de una fundamental virtual (ausente) f_0 de 210 Hz dentro del rango donde el fenómeno de la altura virtual es más pronunciado y repetible.

Los sonidos correspondientes a la presentación dicótica, son, pues la representación acústica de las funciones $X_1(t) = \sin 3 \cdot 2\pi f_0 t$ y $X_2(t) = \sin 4 \cdot 2\pi f_0 t$, respectivamente, sintetizadas cada una en el correspondiente canal estéreo.

Las señales dióticas, por el contrario, son la representación acústica de

$$X(t) = \sin(2\pi \cdot 3f \cdot t) + (2\pi \cdot 4f \cdot t)$$

en la que la frecuencia f varía a voluntad del sujeto, en pasos de uno o dos Hz y en un intervalo que contiene a f_0 , hasta ajustar las alturas percibidas por el. Las muestras utilizadas como estímulos tienen una duración de un segundo y se presentan la diótica a continuación de la dicótica, separadas por una décima de segundo de silencio.

El experimento se realizó cautelosamente en una cámara insonorizada, con tal de obtener el mínimo ruido posible durante la realización de la prueba. Los sujetos fueron advertidos previamente de que el timbre percibido en ambas presentaciones no tiene por qué coincidir, y se les presenta un ejemplo de muestras de diferente timbre, pero idéntica altura y viceversa, de distinta altura y mismo timbre. Los estímulos de sonido fueron sintetizados a 48000 muestras digitales por segundo a 16 bits de resolución utilizando el software Mathematica.

A fin de verificar la ausencia de errores sistemáticos en la generación de los ficheros de audio, se realizaron comprobaciones físicas directas en el laboratorio de Física Experimental. En la figura 14 se muestra una comparación de la traza de osciloscopio de uno de los sonidos dióticos generados con la gráfica de la función utilizada para generarla.

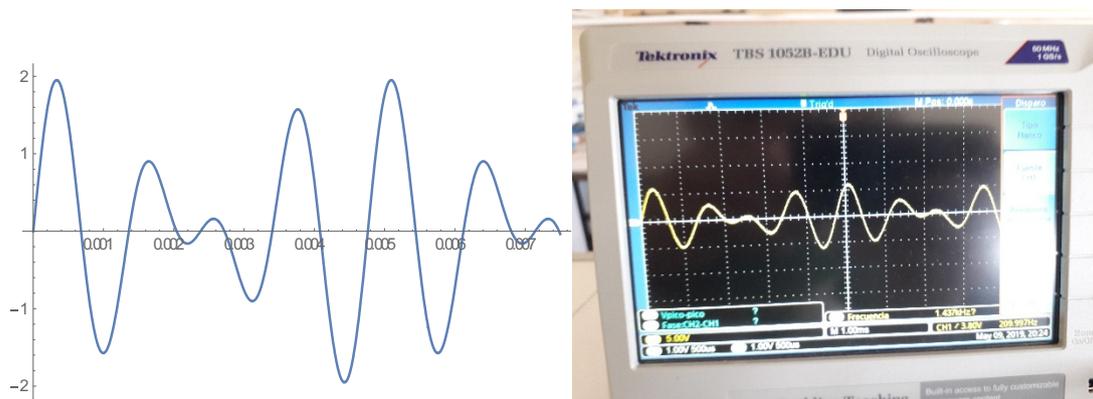


Figura 14: Gráfica de la superposición del tercer y cuarto armónico de una fundamental f (sonido diótico) junto a la correspondiente traza de osciloscopio en ambos canales.

En la figura 15, en cambio mostramos en cada canal del osciloscopio, la señal sinusoidal correspondiente a cada canal estéreo del fichero de audio generado para la muestra dicótica, indicando que no existe comunicación cruzada en el proceso de digitalización.

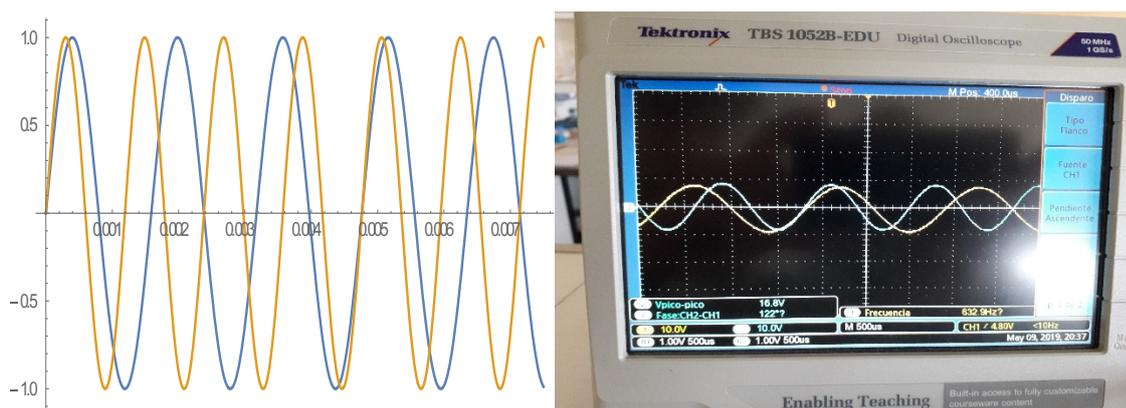


Figure 15: Gráfica sinusoidal del tercer y cuarto armónico en cada canal (sonido dicótico)

Los experimentos preliminares se realizaron enteramente en Mathematica con un programa que crea una tabla de frecuencias f que nos permite emitir sonidos de entre 180 y 240 Hz a intervalos de 1 Hz de frecuencia. De esta manera, tenemos un rango de frecuencias posibles, de las cuales el objetivo es que cada sujeto de estudio seleccione a ciegas el valor de f para el que le la altura virtual dicótica resulte más parecida a la correspondiente altura virtual diótica. Es decir,

el valor “físico” de la coincidencia sería 210 Hz, pero el objetivo del experimento es validar la hipótesis de que entre ambas presentaciones hay una discrepancia sistemática.

El experimento definitivo, por el contrario, se realizó utilizando el software Open Sesame especialmente diseñado para experimentos en diversas ramas de la psicología. Dado el carácter preliminar del experimento a realizar, se eligió esta opción por la facilidad que presenta a la hora de realizar prácticamente el experimento en las cabinas insonorizadas del grupo de Psicología. El propósito es contar con suficiente evidencia del fenómeno que justifique la ulterior realización de una nueva versión más pulida del experimento con software y equipamiento más específico para psicoacústica.

En el programa de Mathematica, generamos una rejilla de botones correspondientes a las distintas frecuencias f y se hace servir la función Partition con tal de dividirla para el conjunto de cada una de ellas, a continuación, se utiliza comando MapIndex que nos da en cada uno su valor en las rejillas de cada uno y también un botón que emite un sonido a través de la función EmitSound que nos genera audios dicóticos y dióticos con un intervalo de silencio en medio de ambos.

Para la realización del experimento definitivo, se exportaron todas las muestras triples a una carpeta a través de la función Export utilizando SetDirectory para localizar la carpeta donde se quiere enviar esos archivos de audio con el valor numérico de las frecuencias convertidos a strings como filename haciendo servir la función “ToString” y los guardamos en formato .wav para su fácil inclusión en el programa de Open Sesame que se utilizó para generar la interface de usuario en el experimento definitivo. Open Sesame es un software usado por los experimentadores de la Facultad de Psicología de la UIB, quienes nos han ayudado muchísimo en la realización de las pruebas psicoacústicas, ofreciéndonos una cabina para llevar a cabo las medidas y enseñándonos como utilizar el software. El notebook de Mathematica con el código utilizado en este trabajo programa de Open Sesame con el que se realizaron los experimentos definitivos, se adjuntan como parte de este informe.

El protocolo del experimento definitivo consistió en las siguientes etapas.

1. Explicación oral del experimento a realizar
2. Breve entrenamiento del sujeto sobre la tarea a ejecutar
3. Comparación de la altura virtual de las muestras presentadas, comenzando por una frecuencia de la fundamental virtual diótica muy por debajo de la dicótica (experimento ascendente)
4. Comparación de la altura virtual de las muestras presentadas, comenzando por una frecuencia de la fundamental virtual diótica muy por encima de la dicótica (experimento descendente)
5. Comparación de la altura virtual de las muestras presentadas, comenzando por una frecuencia de la fundamental virtual diótica elegida al azar (experimento aleatorio)
6. Estimación de la altura relativa percibida cuando ambas frecuencias fundamentales (ausentes) coinciden ($f = 210$ Hz)

Estas etapas se repiten dos veces, y otras dos veces invirtiendo la lateralidad de la excitación dicótica. En cada caso, excepto el último, el sujeto puede explorar a su voluntad los sonidos alrededor de aquellos en los que percibe coincidencia de altura. Para las seis etapas, utilizó una interface de comunicación entre la cabina y el ordenador realizada a través de Open Sesame mediante el siguiente esquema de ejecución:

Primero, está configurado con un par de sets que ponen en ajuste la inicialización del programa y que configuran la pantalla que se va a mostrar a cada sujeto, estableciendo las diferentes escalas.

A continuación, creamos dos definiciones, una para establecer el teclado como método para que se active el comando respectivo al pulsar un botón del ordenador y otro para que salga en pantalla las instrucciones que el sujeto debe seguir en cada momento.

Después, vienen definidos un conjunto de bloques secuenciales que nos permite estructurar el experimento en las diversas partes que éste consta.

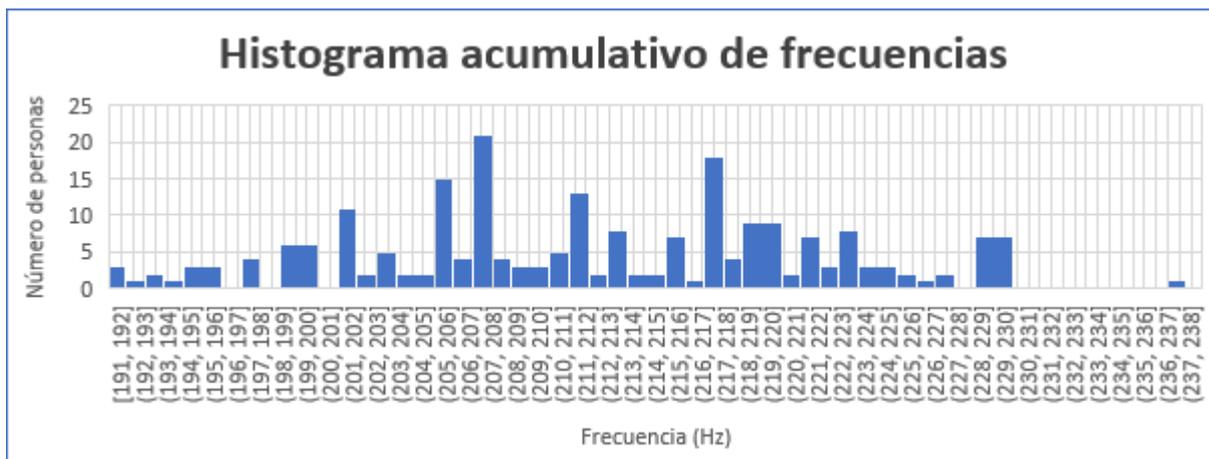
Se establece un “block gen” consistente en conectar las flechas arriba y abajo con subir y bajar la frecuencia en pasos de 2 Hz para que cada individuo pueda decidir aumentar o disminuir la frecuencia del segundo sonido diótico de la muestra según lo considere necesario hasta emparejar la altura que percibe con la del primero, dicótico, que es siempre el mismo en todos los bloques del experimento.

Cuando el sujeto está conforme con la coincidencia, pulsa la tecla “m”, interrumpiendo el loop y registrando automáticamente la frecuencia a la cual ha escuchado coincidencia de alturas. Seguidamente pasa automáticamente al siguiente bloque hasta continuar con todas las secuencias. Cabe señalar que, para una mayor precisión de los resultados, establecimos una doble vuelta en cada uno de los bloques, es decir, se repite cada uno dos veces y se promedia entre ambos datos para tener una precisión mayor.

Hacia el final del experimento se interroga al sujeto sobre sus datos de edad, conocimiento musical y su apreciación subjetiva sobre la complejidad de la tarea que se le ha encomendado. Esto permitirá, en un análisis más detallado, separar los resultados de cada uno de los bloques entre los correspondientes a sujetos con y sin formación musical.

3. Resultados.

Durante la experimentación se entrevistaron 36 sujetos, de los cuales 5 tenían educación musical avanzada, 12 educación musical media y el resto no tenían educación musical alguna. Los datos recogidos se resumen en el siguiente histograma, en el que se acumulan todas las estimaciones de coincidencia en altura para el complejo diótico de la correspondiente frecuencia fundamental ausente, con la del complejo dicótico con fundamental (ausente) de 210 Hz obtenidas en el total de bloques del experimento:



El histograma muestra un claro vacío alrededor de la coincidencia física (210 Hz).

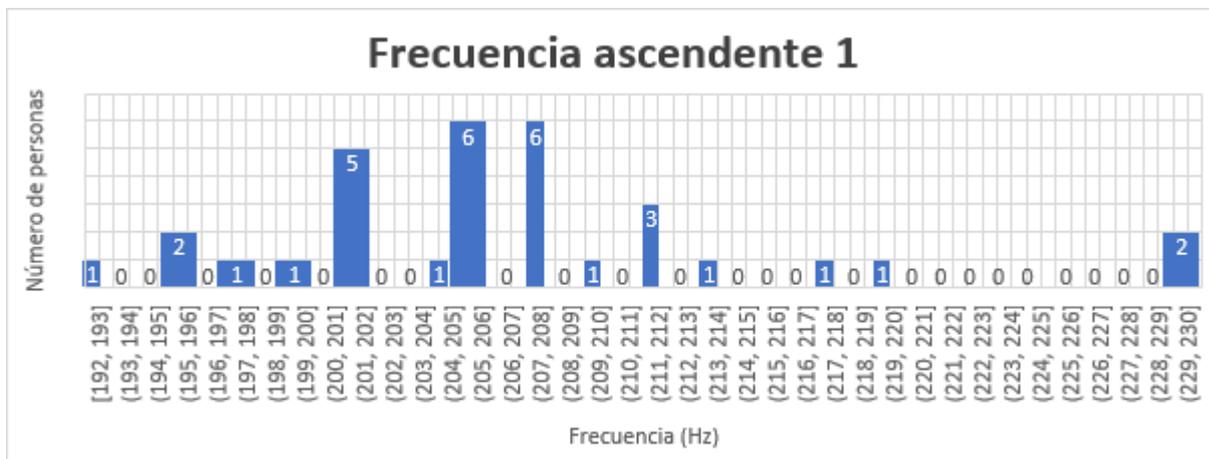
Está claro, sin embargo, que la dispersión en los datos es enorme, lo cual revela muy probablemente que una buena parte de los entrevistados o bien no ha comprendido la tarea que se les ha asignado, o bien carece de la capacidad o el entrenamiento necesarios para comparar alturas lo que a su vez sugiere la necesidad de re-diseñar el experimento para incluir un período de entrenamiento con tonos musicales antes de recoger sus respuestas relativas a esta cuestión.

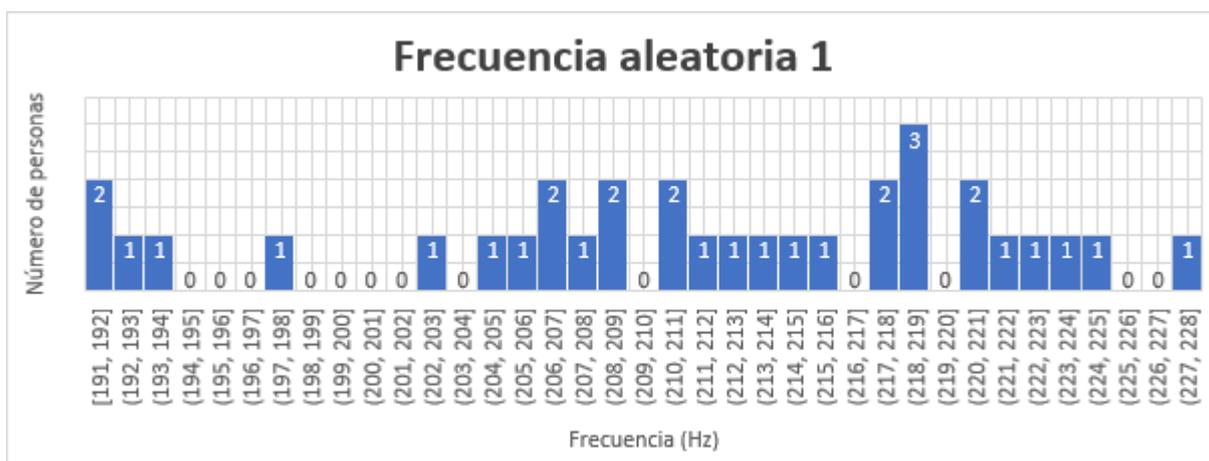
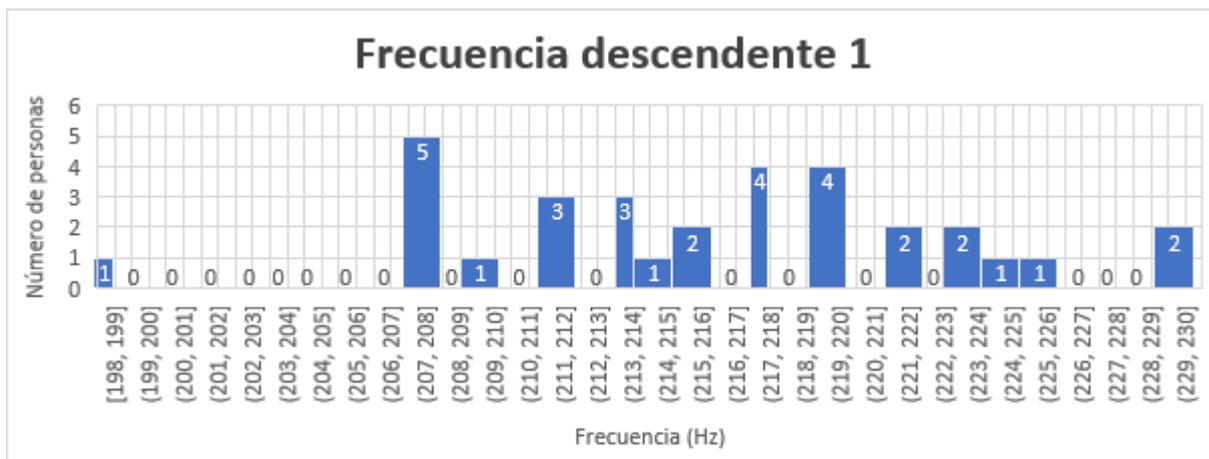
Los siguientes histogramas corresponden a cada uno de los bloques repetidos dos veces. Por “Frecuencia ascendente” nos referimos al bloque del experimento que se inicia con la frecuencia diótica bien por debajo de la dicótica (el sujeto debe modificar en aumento para alcanzar la coincidencia). Por “Frecuencia descendente” debe entenderse el bloque en el cual ocurre lo opuesto (se comienza por una frecuencia diótica muy superior a la dicótica y el sujeto debe explorar hacia abajo). “Frecuencia aleatoria” denota el bloque en el que la frecuencia de partida del estímulo diótico se elige al azar.

Nótese que en cada caso la ausencia de coincidencias correspondientes a la de las frecuencias físicas (210 Hz), es evidente.

3.1 Datos Generales.

3.1.1 Bloque 1.





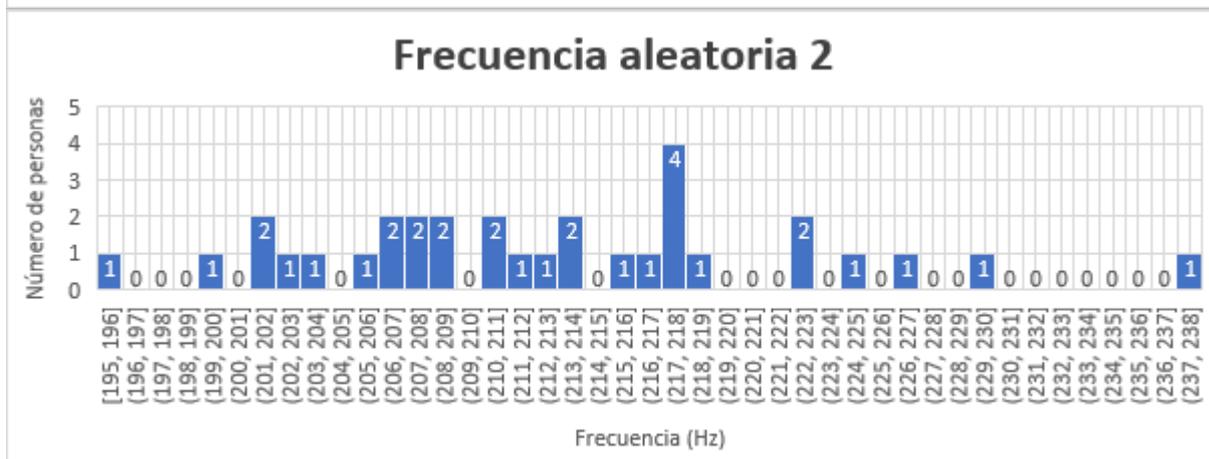
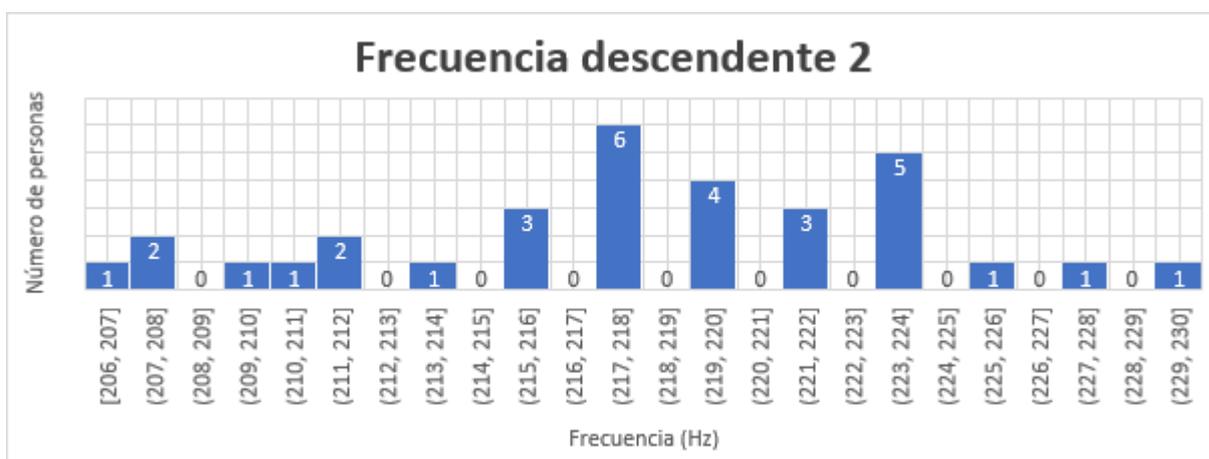
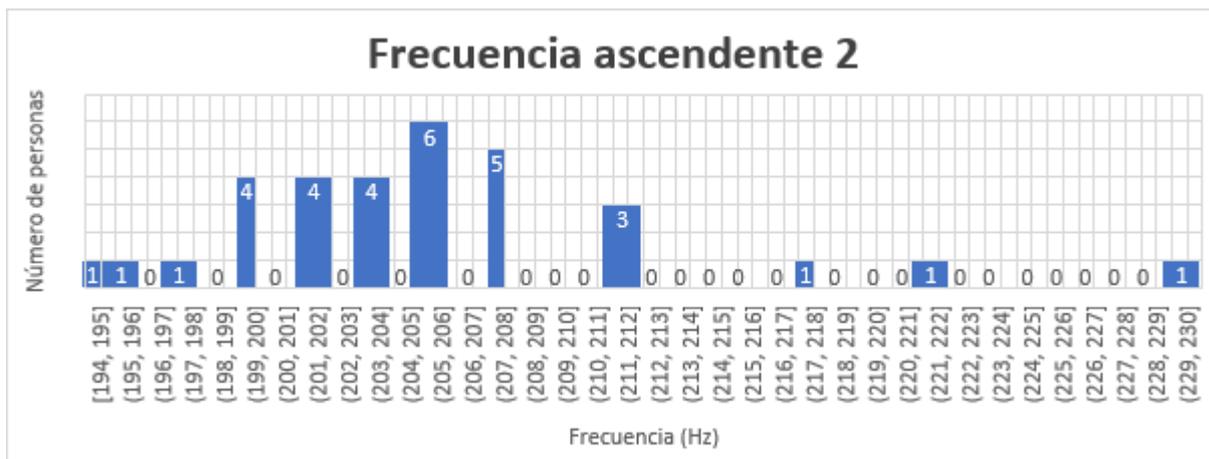
En la primera etapa, donde se presentaban sonidos dicóticos correspondientes al tercer armónico en el izquierdo y el cuarto en el derecho, podemos apreciar como la gran mayoría de los resultados que se han obtenido se encuentran por debajo de la frecuencia fundamental, discerniendo la mayoría en tan sólo un par de Hertzios. Como puede apreciarse, 14 de los 32 sujetos considerados, se encuentran en el rango de entre 205 y 209 Hertzios lo que representa casi un 50% de los individuos estudiados.

Con el descendiente ocurre una cosa parecida, pero por arriba de la frecuencia fundamental, encontrándose casi todos ellos entre 211 y 220 Hz. Lo evidente que se quería comprobar en este estudio es que ninguno de los resultados o prácticamente ninguno se detuviese en la propia fundamental ($f_0 = 210\text{Hz}$), y como puede notarse sólo uno de los sujetos se detuvo en esa zona en cada una de las partes de las que constaba el experimento.

En el caso del aleatorio o random, ninguno se detuvo en la fundamental y hay una especie de gaussiana en las medidas, lo cual es una buena señal ya que implica que casi todos se detuvieron en frecuencias similares.

3.1.2 Bloque 2.

Invertiendo los canales auditivos del dicótico, es decir los respectivos armónicos presentados a cada oído, no se produce una diferencia sustantiva en la apreciación de las alturas relativa:

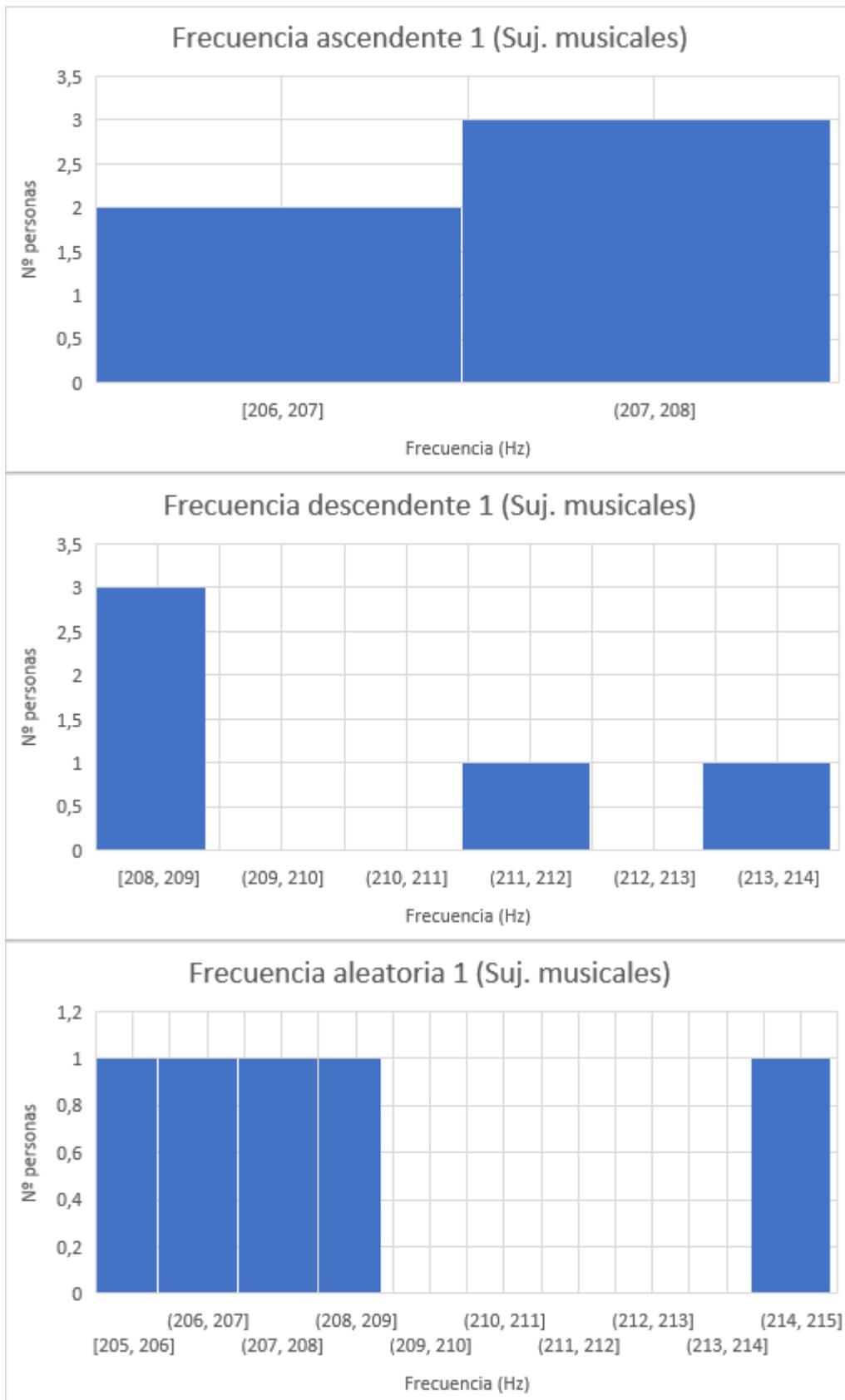


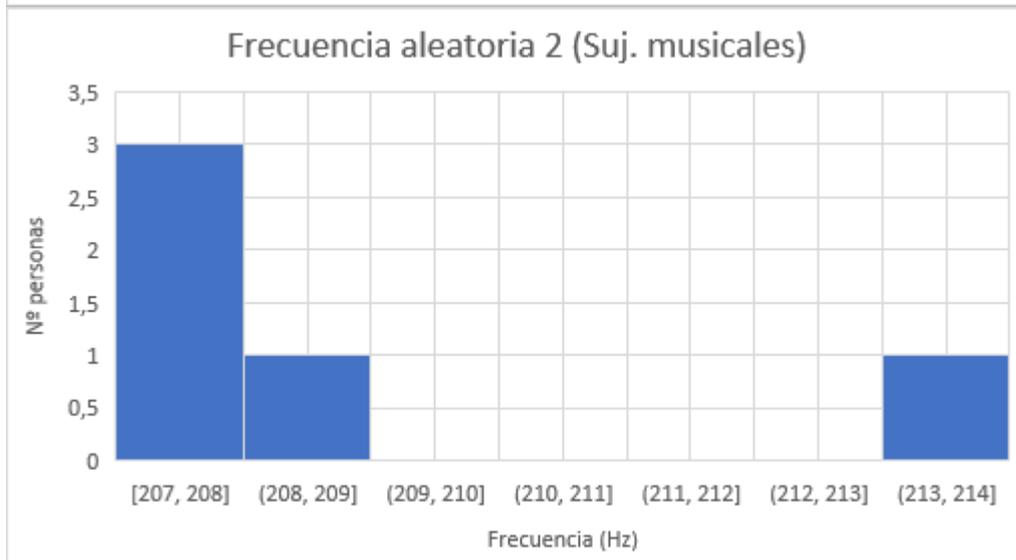
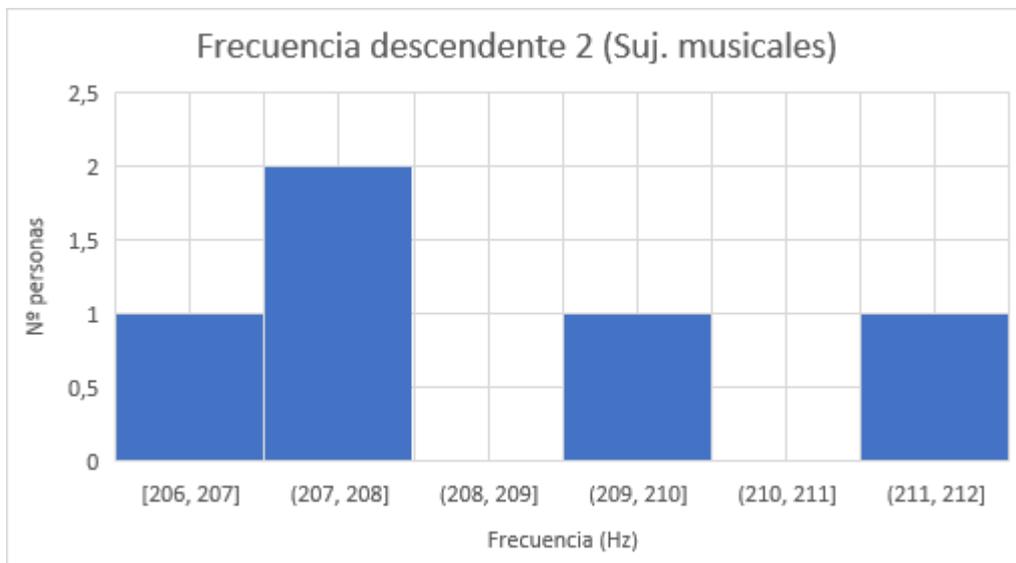
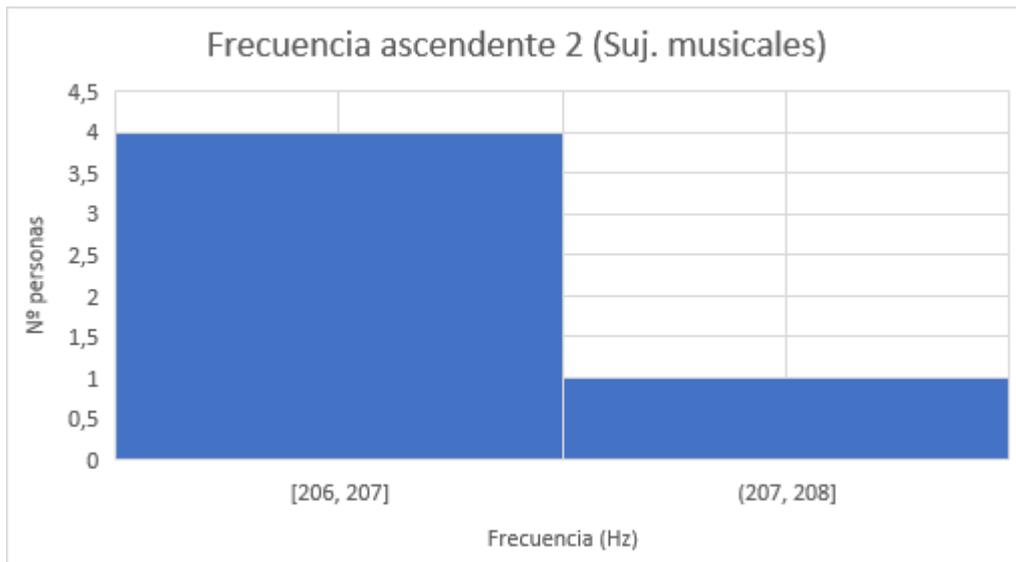
En el último bloque del experimento, la gran mayoría han mencionado que oyen más agudo el primer sonido, es decir, el sonido dicótico. Los porcentajes son los siguientes: 59%: mencionan el primero 32%: mencionan el segundo 9% difieren de la primera a la segunda escucha en que se intercambia la lateralidad del dicótico.

3.2 Datos de los sujetos musicales.

Restringiéndonos a las respuestas de los 5 con estudios superiores de música los resultados son más precisos y repetitivos, como lo muestran las gráficas que se presentan a continuación. Las respuestas de estos sujetos parecen convalidar la observación, que dio lugar originalmente a

este estudio, que la altura percibida del tono dicótico está unos cuatro Hz por debajo de la altura de la fundamental faltante del tono armónico combinado.





En el bloque final, los sujetos con oído entrenado y conocimiento musical sólido coinciden en un 66% contra 17% en escuchar más agudo el diótico que el dicótico de las mismas frecuencias. Mientras que el 17% restante cambian de opinión al invertir el canal dicótico.

4. Discusión y conclusiones

El experimento realizado valida preliminarmente, aunque con bastante solidez, la hipótesis de que las alturas percibidas de un complejo armónico de fundamental ausente presentado en forma dicótica y en forma diótica no coinciden.

Si bien la enorme dispersión de los resultados obtenidos impide dar por el momento una valoración numérica y estadística significativa de cuanto y en que dirección es esta discrepancia, es sugestivo que los 6 sujetos con educación musical avanzada coincidan mayoritariamente en evaluarla en el entorno de los 4 Hz, y percibiendo el tono dicótico sistemáticamente como el más grave entre ambos. Este resultado coincide con las valoraciones preliminares mencionadas antes y que fueron motivo de esta experimentación más extensa.

Un fuerte soporte adicional para esta validación viene provisto por los resultados del último bloque del experimento en personas con oído educado, en el cual los sujetos tienen que evaluar la altura relativa del estímulo dicótico frente al diótico de exactamente la misma fundamental de 210 Hz. En este caso el individuo solo puede escoger entre dos opciones: más agudo o más grave (excluyendo la coincidencia). Si la apreciación mayoritaria fuese de coincidencia, las fluctuaciones se deberían distribuir simétricamente, sin embargo, los datos muestran un sesgo significativo hacia considerar el estímulo diótico como más agudo que el dicótico, lo cual coincide con la hipótesis inicial.

Quedaría pendiente un análisis estadístico más riguroso de los datos obtenidos, que debido al tiempo invertido en el diseño y realización del experimento no se ha podido completar al momento de la presentación de esta memoria. Además, buena parte de los sujetos interpelados, nos ayudaron en el proceso de ajuste del diseño experimental, por lo que los datos obtenidos a partir de sus respuestas no son de la misma calidad que los obtenidos con el resto de los participantes una vez que el protocolo experimental fue ajustado.

Por esta razón sería conveniente extender la experimentación en dos direcciones: primero continuar con las entrevistas a un conjunto de sujetos indiscriminados, pero incluyendo un proceso previo de entrenamiento en la comparación de alturas percibidas; y segundo, incrementar el número de sujetos con formación musical elevada.

5. Bibliografía

- [1]- J. Angus, D. Howard, “Acoustics and Psychoacoustics”, 4th Edition, 2017.
- [2]- Alonso M., Finn E. J. Addison, “Física”, Wesley Iberoamericana, 1995.
- [3]- Stefanos A. Iakovides, Vassiliki TH Iliadou “Psychophysiology and psychoacoustics of music: Perception of complex sound in normal subjects and psychiatric patients”, 2004
- [4]- Plomp R., Terhardt E. “Contemporary hybrid theories of pitch incorporate aspects of both”, 1964.
- [5]- Moore B.C.J, “An introduction to the Psychology of Hearing”, Academy Press, 1st Edition, 1977.
- [6]- Cartwright J.H.E , González D.L., Piro O. “A New Nonlinear Model for Pitch Perception”, Wesley Iberoamericana, 1999.
- [7]- Schouten, J. F., “The perception of subjective tones in Psychological Acoustics”. Edited by E.D. Schubert, 1979.
- [8]- Cherry, E.C., “Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*”, 1953.
- [9]-Trimble, O.C., “Concerning the meaning of the terms diotic and dichotic. *American Journal of Psychology*”, 43:144, 1931.
- [10]- Zwicker E., Fastl H., “Psychoacoustics: Facts and models”, 1990
- [11]- Nakamura R., Aikawa K., “Difference between diotic and dichotic presentations on audiovisual synchronization”, *The Journal of the Acoustical Society of America* **117**, 2423, 2005.
- [12]- Demany, L., and Semal, C., “Dichotic fusion of two tones one octave apart: Evidence for internal octave templates”, 1998.