



## Portafolio personal

Nicolás A. Peña Escarpentier

**Nota:** La materia del curso y anexos están en letra normal, mientras que las reflexiones se escribieron en letra cursiva.

### 1. Sonido

El sonido lo podemos descomponer en dos partes: una física y otra perceptual. Para aclarar esta dicotomía, podemos usar una analogía con la visión. La parte física se refiere a la luz, la vibración de las partículas, mientras que la parte perceptual se refiere a lo que interpretamos de todas las señales que recibimos. Es lo mismo en el sonido, las ondas de presión son la componente física, pueden ser grabadas, procesadas o analizadas; mientras que la interpretación que le damos, el como suena en nuestra cabeza, es la componente perceptual. En esta clase nos enfocamos a la parte del fenómeno físico.

Físicamente, el sonido es una onda mecánica, por lo que necesita un medio que se perturbe para poder transmitirse. Al ser una onda mecánica, el sonido puede rebotar, ser absorbido, cambiar de medio, entre otras cosas. Como se dijo anteriormente, estas ondas se pueden almacenar, reproducir, modificar, desviar, etc. Si consideramos una fuente de sonido puntual, podemos decir que el sonido se propaga de forma esférica a cierta velocidad, la cual depende del medio de propagación. Podemos ver distintos ejemplos de la velocidad de propagación del sonido en distintos medios en la tabla 1.1<sup>1</sup>.

Medio	Velocidad de propagación
Aire (20°C)	343 $\frac{m}{s}$
Agua (25°C)	1593 $\frac{m}{s}$
Madera	3700 $\frac{m}{s}$
Hormigón	4000 $\frac{m}{s}$
Acero	6100 $\frac{m}{s}$

Cuadro 1.1: Velocidades de propagación del aire en distintos medios.

*Se dijo que en clases que la base de los instrumentos musicales tradicionales es la resonancia. En general, los instrumentos tradicionales, no eléctricos, poseen un sistema de caja, usado para amplificar la intensidad de las ondas de sonido, puesto que la sola vibración de la cuerda no es suficiente para hacer que el sonido sea escuchado a una gran distancia. Además, se aprovecharon de la capacidad acustica de la madera y de las propiedades de las ondas mecánicas. La madera puede transportar el sonido a la caja, en donde las ondas de sonido rebotan y se amplifican.*

¿Qué necesitamos para producir un sonido?

- Energía cinética (fuente de energía). Necesitamos energía para poder generar una onda, lo cual no es lo mismo que la oscilación. Podemos definir vibración como un movimiento el cual vuelve al equilibrio. Y desde este punto definimos una oscilación como un ciclo del movimiento en torno al punto de equilibrio.
- Fuente en vibración. Si no hay nada que vibre, no se produce la diferencia en la presión, la cual provoca el sonido.

<sup>1</sup>Fuente: Wikipedia.org

- Medio de propagación. El sonido no se puede propagar en el vacío, puesto que es una onda mecánica.
- Receptor. Este no es necesario desde el punto de vista físico, aunque es necesario para medir el sonido.  
*Desde este punto de vista, la pregunta “Si un árbol cae en medio de un bosque y no hay nadie que lo escuche, hace ruido?” es bastante paradójica y nada fácil de resolver, aún desde un punto de vista físico. Se puede decir que se van a producir ondas sonoras debido al movimiento de partículas y que se generan cambios de presiones, pero es necesario un receptor para poder medirlas y tener certeza de que existieron. Pero la pregunta niega la posibilidad de tener el receptor, por lo que esa respuesta no la satisface a cabalidad.*

Podemos ver como las vibraciones producen sonido con ciertas características es el diapasón. Este es un instrumento fabricado de tal manera que vibra a una frecuencia determinada, lo cual lo interpretamos como cierta nota. Al golpear el diapasón, sus extremos vibran, generando una variación en el campo de presiones en las moléculas presentes en el aire. Como se vio en los videos en clases, las moléculas no experimentan movimiento, sino que un cambio de presiones unidireccional. Éstas no viajan junto con la onda, sino que oscilan en torno a su estado de equilibrio (cuando no se produce sonido). Es posible distinguir dos zonas en este campo de presiones: una zona de compresión y otra zona de tracción. Así se conforma la onda longitudinal que es el sonido.



Figura 1.1: Diapasón

## Propiedades de los sistemas vibratorios

- **Ciclo:** Unidad de repetición.
- **Período:** Tiempo que toma un ciclo en completarse. Su unidad de medida en el SI es el segundo.
- **Frecuencia:** Número de ciclos por unidad de tiempo ( $= 1/\text{período}$ ). Su unidad de medida en el SI es el Herz (Hz).
- **Fase:** Parte del ciclo que una onda ha avanzado en relación a un punto de referencia arbitrario.  
 Es imposible distinguir la fase de un sonido puro, pero para efectos perceptuales y cuando se van “agregando” sonidos importa mucho.  
 Para sonidos de distinta frecuencia, no tiene sentido comparar fases.

La superposición de las notas depende de la diferencia de frecuencias. Esto depende, eso sí, de que frecuencias son! Si son muy parecidas, se van acoplando y desacoplando, pero si son muy diferentes, se escuchan dos notas por separado.

## Ondas sonoras

- **Onda simple o sinusoidal:** Ondas simples, completamente definidas. Representa la solución matemática de un movimiento armónico simple. Necesita de tres parámetros para quedar definida: amplitud ( $A$ ), frecuencia ( $\omega$ ) y fase( $\phi$ ). Se puede escribir como:

$$A \sin(\omega t + \phi)$$

- **Ondas complejas:** Son los sonidos “reales”. No existe instrumento alguno que haga sinusoides puros. Algunos tipos de ondas complejas son:
  - Onda diente de sierra
  - Onda cuadrada
  - Sonido vocal

Se puede hablar de una frecuencia *fundamental*, pero no de una frecuencia única, puesto que estas ondas son la unión de una cantidad enorme de otras.

## Análisis de Fourier

Método desarrollado por el matemático francés Jean-Baptiste Joseph Fourier. Cuando inventó este método, le interesaba ver 1 problema termodinámico en el cual tenía que trabajar con ondas, pero no podía analizarlas. Debido a esto, demostró que cualquier onda se puede analizar como la suma de ondas sinusoidales (armónicas) con  $A$ ,  $\omega$  y  $\phi$  particulares. De este modo, podemos analizar los sonidos como la suma de sonidos más básicos (*lo cual es claramente mucho más simple!*) Todo tipo de ondas puede ser descompuesto como sinusoides, incluso ondas cuadradas.

Es importante mencionar que el cerebro también es capaz de descomponer las ondas según sus frecuencias, lo cual se verá más adelante en el curso.

*Este tipo de análisis es muy importante. Los cuerpos al vibrar lo hacen en varios modos de vibración y sin saber exactamente las propiedades del cuerpo, no es posible saber en que frecuencias vibra y con qué intensidad.*

## Representación visual del sonido

- **Amplitud v/s tiempo:** Es muy difícil saber de que se trata el sonido. No se puede recuperar la frecuencia.
- **Amplitud v/s frecuencia:** No tenemos información de la duración de cada frecuencia, ni cuando suena cada una.
- **Frecuencia v/s tiempo:** No podemos representar amplitud, es decir cuán fuerte suena cada sonido.
- **Espectrograma:** Los espectrogramas son una buena solución, puesto que pueden mostrar la frecuencia v/s tiempo, indicando la amplitud (intensidad) con una escala de colores.

## 2. Sistema auditivo

El sistema auditivo es el conjunto de órganos responsables de dotarnos del sentido del oído. Ya vimos anteriormente la composición física del sonido. El primer paso es el saber físicamente como llegan las ondas a nuestra cabeza para poder procesarlas, para lo cual estudiaremos este sistema. El sistema auditivo está compuesto por tres partes: oído externo, medio e interno, las cuales detallaremos a continuación y se pueden apreciar en la figura ?? <sup>2</sup>.

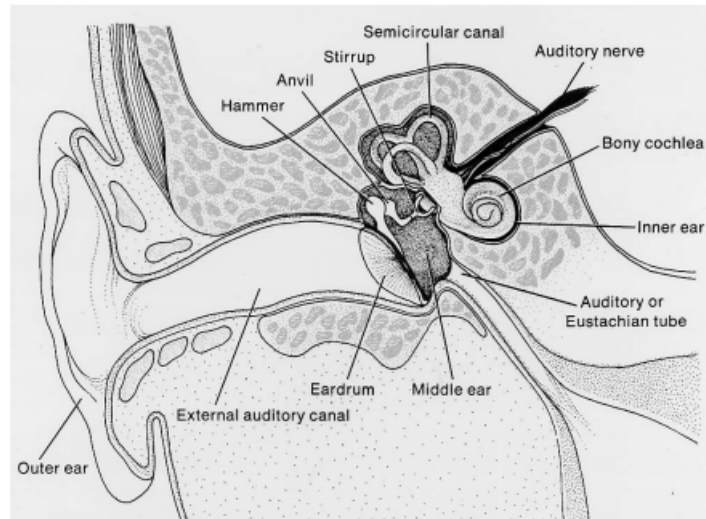


Figura 2.1: Corte transversal del sistema auditivo humano, mostrando cada una de sus partes

### Oído externo

El oído externo es el encargado de captar las ondas sonoras y transportarlas al oído medio. Está compuesto por la pinna (oreja), el canal auditivo y el tímpano. La oreja, pinna o pabellón auricular se encarga de captar las vibraciones externas y llevarlas hacia el interior del oído. Además, aprovecha de amplificar ciertas frecuencias y atenuar otras, con el fin de enfatizar las frecuencias que más nos interesan. Sin el aporte de la pinna, sería mucho más difícil el proceso auditivo, puesto que sería bastante complicado que las ondas sonoras entraran al canal auditivo. Por otra parte, la pinna es esencial para la localización de los sonidos. El canal auditivo es una cavidad de aproximadamente 2.5 cm de largo, la cual conduce las vibraciones recibidas desde la pinna hacia el tímpano. Este canal también amplifica cierto contenido de frecuencias, entre 1000 Hz y 4000 Hz, siendo el máximo alrededor de los 3000 Hz. Finalmente, el tímpano es una membrana elástica de aproximadamente 9 mm de diámetro. Éste se encarga de transmitir las vibraciones recibidas desde el canal auditivo hacia el oído medio. El tímpano también tiene ciertas frecuencias predilectas, las cuales se ven amplificadas. Por otra parte, el tímpano aprovecha de sellar el conducto auditivo, aislando al oído medio del exterior. La respuesta final, en términos de frecuencia, del conjunto del oído externo se puede ver en la figura ?? <sup>3</sup>. *El hecho de que ciertas frecuencias se vean amplificadas es lógico debido al concepto de modos de vibración de medios elásticos. Aún así, es interesante ver como el cuerpo ha evolucionado para que se amplifique el contenido de frecuencias que más nos interesa.*

### Oído medio

El oído medio es un sistema mecánico compuesto por tres huesecillos que actúa como un transductor de vibración. Esto se debe a que el oído interno está lleno de líquidos, por lo que necesitamos una forma de transmitir estas vibraciones sin perder una gran cantidad de energía de dichas ondas. Los huesecillos que forman este sistema son el martillo, el yunque y el estribo. Además, tenemos un músculo llamado trapedius, el cual protege contra sonidos muy intensos, amortiguando la vibración de dichos huesos. Lamentablemente, dicho músculo necesita cierto tiempo de reacción, no es de reacción inmediata, por lo cual no alcanza a responder contra sonidos súbitos como un disparo.

El funcionamiento de este sistema mecánico es el siguiente: El martillo está conectado a la membrana timpánica en su base. Al vibrar la membrana, también lo hace el martillo que está conectado en su extremo opuesto al yunque. Al moverse el

<sup>2</sup>Fuente: Clases

<sup>3</sup>Fuente: Clases

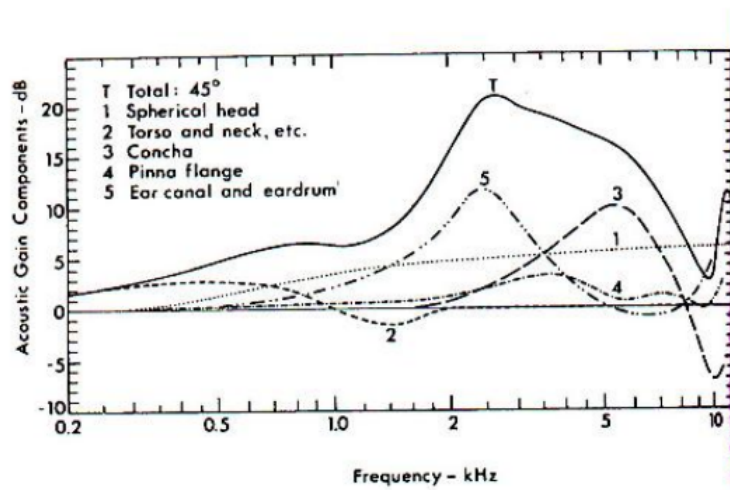


Figura 2.2: Respuesta de frecuencias del oído externo

martillo también lo hace el yunque al unísono. Este último está conectado en su extremo opuesto al estribo, traspasándole las vibraciones. A su vez, el estribo se apoya en la ventana oval, la entrada del oído interno, transmitiendo las señales sonoras a esta última parte. Podemos ver la conexión de los tres huesecillos en la figura ?? <sup>4</sup>.

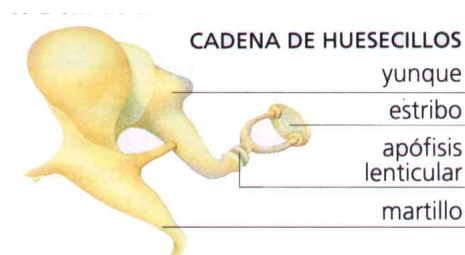


Figura 2.3: Huesecillos del oído medio

## Oído interno

El oído interno es la última parte del sistema auditivo, en donde se transforman las vibraciones en señales nerviosas que se dirigen al cerebro. Es una cámara hueca ubicada dentro del hueso temporal, llena de líquido incompresible y compuesto por las ventanas oval y circular, los canales semicirculares y la cóclea junto a la membrana basilar. No nos enfocaremos mucho en los canales semicirculares, puesto que son los encargados de la parte espacial y del equilibrio. Nos enfocaremos en las partes que tienen que ver con la audición: la cóclea y la membrana basilar. Podemos ver el oído interno en la figura 2.4 <sup>5</sup>.

La cóclea es un conducto de tres tubos enrollados, con forma de caracol enrollado en si mismo 3.5 veces y que tiene un diámetro de 2 mm y un largo de 35 mm. Los tubos de la cóclea se llaman “scala tympani”, “scala vestibula” y “scala media”, las cuales se unen por el helicotrema y se separan por la membrana de Reissner y la membrana basilar. Esta membrana comienza en la base, conectada a la ventana oval y termina en el apex en el helicotrema. Esta membrana varía tanto su densidad como su rigidez en su largo, lo que hace variar su frecuencia de resonancia en cada punto. En la base es más rígida, por lo que su frecuencia fundamental es mayor (20 kHz), mientras que a medida que se avanza hacia el apex, se vuelve menos rígida, disminuyendo su frecuencia fundamental (20 Hz). Esta distribución de frecuencias se puede ver en la figura 2.5 <sup>6</sup>.

El funcionamiento del oído interno es el siguiente. Las vibraciones recorren la cóclea desde la ventana oval. A medida que avanzan estas vibraciones van rebotando contra la membrana basilar hasta que encuentran el lugar correspondiente a su frecuencia predominante. En dicho lugar, las vibraciones son capaces de traspasar la membrana basilar y volver por la

<sup>4</sup>Fuente: [http://usuarios.discapnet.es/ojo\\_oido/esquemas\\_cuerpo\\_humano/estribo\\_martillo\\_yunque.jpg](http://usuarios.discapnet.es/ojo_oido/esquemas_cuerpo_humano/estribo_martillo_yunque.jpg)

<sup>5</sup>Fuente: Clases

<sup>6</sup>Fuente: Clases

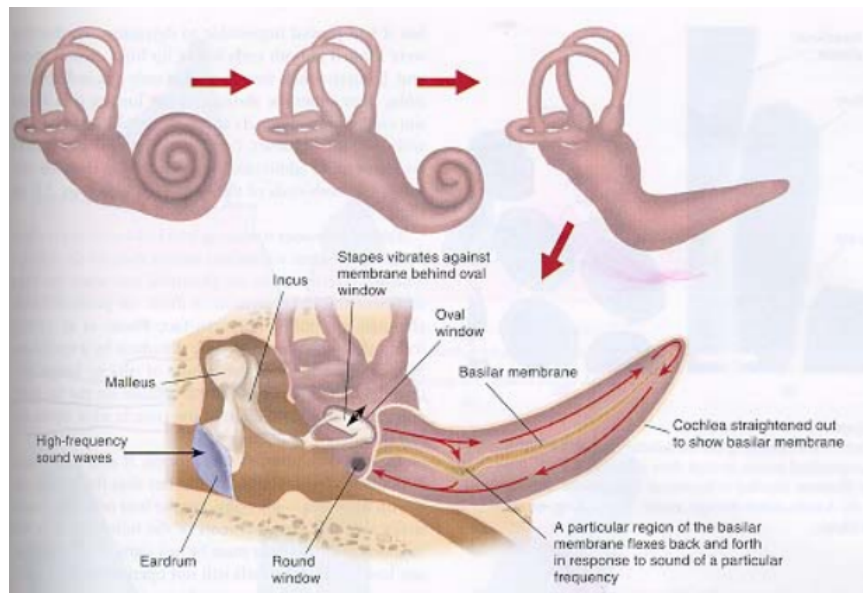


Figura 2.4: Vista y detalle del oído interno

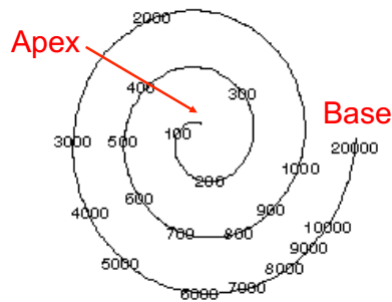


Figura 2.5: Distribución de frecuencias en la membrana basilar

cóclea hasta la ventana circular. Al atravesar la membrana basilar, se estimula el órgano de Corti, causando vibración en algunas de sus células ciliares. Finalmente, estas vibraciones son traducidas en impulsos eléctricos los cuales se transmiten por el nervio auditivo al cerebro para ser procesados.

## Daños al sistema auditivo y pérdida de audición

Hay dos tipos de pérdidas auditivas, las conductivas y las neurosensoriales. Las pérdidas conductivas se dan antes de la ventana oval, son daños al sistema físico o mecánico de transmisión de vibraciones. Estas pérdidas se pueden reemplazar por distintos aparatos que simulan el órgano dañado. Incluso se puede reparar la funcionalidad de los huesecillos del oído medio, mediante el implante de un dispositivo vibratorio que ayuda a mover dichos huesos. Por otra parte, las pérdidas neurosensoriales ocurren en la cóclea o en la membrana basilar y se pueden arreglar mediante un trasplante de cóclea o con un implante de electrodos que mandan señales directamente al nervio auditivo.

### 3. Metodos psicofísicos

En el estudio de la percepción de los sonidos hay varios campos de interés. Uno de estos es el poder cuantificar y medir nuestra percepción. Entonces, ¿cómo hacer un experimento que haga sentido y nos permita medir nuestra sensibilidad? Y luego, ¿cómo interpretamos tales resultados? Por otra parte, no podemos olvidarnos que el cambio de intensidad y los cambios de frecuencia afectan de forma distinta la percepción y se afectan entre si también (*como ya se vió en temas anteriores y en la tarea encargada esta clase*). Por último, debemos ser capaces de incluir el sesgo de la persona en cuestión, puesto que no todos tenemos criterios iguales. Así, se han desarrollado varios experimentos con el fin de medir nuestra percepción, corrigiendo los resultados de forma estadística o anulando el error en el mismo experimento.

#### Psicología matemática/experimental/estadística

También conocido como psicología experimental o estadística o psicofísica. Es una rama de la psicología que se enfoca en medir la percepción. Su precursor fue Gustav Fechner, filósofo y psicólogo alemán del siglo XIX. Él señaló que el incremento de intensidad mental frente a un incremento de intensidad físico puede ser medido, logrando establecer la siguiente ecuación que refleja dicha relación:  $S = K \log(I)$ . En donde  $S$  es la intensidad de la percepción,  $I$  es la intensidad física del estímulo y  $K$  es una constante que depende del tipo de estímulo.

Se define el **Umbral** (threshold) como la intensidad mínima de un estímulo para poder inducir una reacción. Se definen dos tipos de umbrales, según lo que se quiere medir:

- **Umbral absoluto:** Se define como la menor intensidad a la cual un estímulo puede ser detectado.
- **Umbral de discriminación:** Se define como la diferencia mínima en intensidad física, como para percibir un cambio a nivel perceptual.

#### Métodos clásicos

Dentro de los métodos clásicos de detección de señales tenemos:

- **Método de límite:** Consiste en determinar la intensidad a la cual ya no se puede detectar el estímulo. Con tal fin, el observador va cambiando la intensidad y preguntándole al sujeto si es capaz de detectar el estímulo o no.
- **Método de ajustes:** Experimento similar al anterior, con la diferencia que es el mismo sujeto quien maneja el cambio de intensidad, ajustando el nivel de mínima audición.
- **Método de estímulo constante:** Se presentan estímulos a intensidades aleatorias y se mide el porcentaje de respuestas positivas del sujeto.

Todos estos métodos se basan en análisis estadísticos, por lo que se genera un promedio para la población de ciertos lugares. Pero la verdad es que el umbral puede cambiar o el criterio de las personas afecta de gran forma este tipo de experimentos. Por ejemplo, después de un concierto somos menos sensibles a intensidades bajas, o ciertas personas son más propensas a decir que sí a que no. Debido a esto, las mediciones quedan afectadas por una gran cantidad de ruido que necesita ser eliminado para tener la percepción “pura” de las personas.

#### Teoría de detección de señales

La detección de señales está controlada por dos procesos internos básicos:

- Una **impresión sensorial** que en promedio es dependiente de la intensidad de la señal.
- Una **respuesta interna**, la cual depende de nuestro proceso subjetivo interno. Este proceso es dependiente de cada persona.

Además, se trabaja sobre el supuesto de que hay ruido. Este ruido es una perturbación aleatoria que puede ser confundida con la señal a medir. Es fundamental considerar el ruido a la hora de realizar las mediciones, ya que debido a este podemos tener distintos tipos de resultados. Se realizan cuatro tipos de mediciones distintas, de acuerdo a las combinaciones entre las respuestas y si se emitió señal o no, las que se pueden observar en la tabla ?? . En general, se busca maximizar los “Hit” y “Correct rejections” para ver la percepción real del humano. Para correrir el registro, se ve la respuesta interna, comparando la curva de respuesta entre el ruido (N) y ruido + señal (N+S), como se ve en la figura ??.

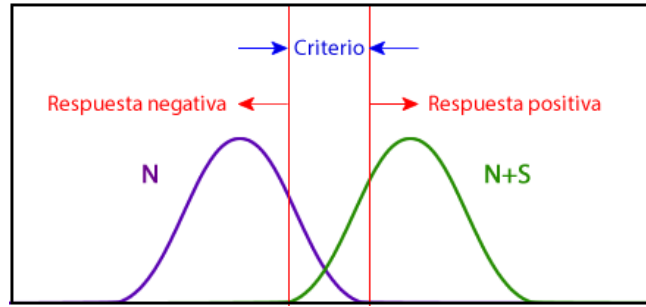


Figura 3.1: Las curvas no son arbitrarias, representan la percepción de la persona. Lo que puede cambiar es el criterio. Este criterio es mediante el cual escogemos una respuesta positiva o negativa cuando tenemos tanto ruido como ruido+señal.

		Señal emitida	
		Sí	No
Respuesta	Sí	Hit	False alarm
	No	Miss	Correct rejection

Cuadro 3.1: Resultados de mediciones experimentales

Es esencial lograr independizar el criterio, para adquirir mediciones correctas, puesto que este está condicionado por el ambiente. Ahora, como logramos eliminar el sesgo en las preguntas? Este sesgo es especialmente influyente cerca del umbral, donde los niveles de ruido y ruido+señal son muy parecidos. Es más, incluso un observador sin sesgo podría dar una respuesta positiva aun cuando no hubiese ruido, sólo debido al ruido existente.

Se define el sesgo como la tasa de acierto (“hit”) v/s falsas alarmas. Así, tomando en consideración el sesgo de la persona, podemos medir la sensibilidad ( $d'$ ). La sensibilidad mide la distancia (en desviaciones estandar) de las curvas N y N+S, como se muestra en la figura ???. Cabe mencionar que ambas curvas son curvas normales de igual varianza. Además, con la misma sensibilidad igual puede cambiar el criterio de la persona. De la misma forma, puede cambiar la sensibilidad sin cambiar el criterio utilizado. Por ejemplo: el cambio antes y después de un concierto de rock no tiene que ver con el criterio.

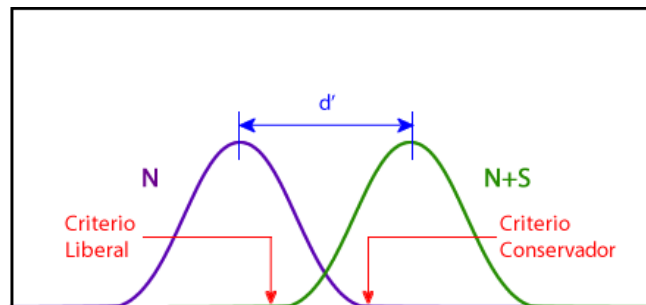


Figura 3.2: Sensibilidad auditiva no cambia, pero el criterio puede cambiar. Con un criterio liberal, decimos sí la mayoría de las veces, en cambio con un criterio conservador, sucede lo contrario.

Con los resultados del experimento anterior se puede graficar en una **curva ROC** para ver el desempeño obtenido. Las curvas ROC (Receiver Operating Characteristic o Característica Operativa del Receptor) se obtienen de graficar el porcentaje de aciertos v/s porcentaje de falsas alarmas como se puede apreciar en la figura ???. La recta central ( $d' = 0$ ) representa a alguien que no logra distinguir ruido de estímulo. La curva obtenida representa la sensibilidad de la persona, mientras que los puntos de la curva son los posibles cambios de criterio para el mismo  $d'$ .

Otro experimento que se puede realizar es el de **opción forzada de dos alternativas**. Este consiste en que se le presentan al receptor dos sonidos, un ruido y un ruido + señal y se le hace escoger cuál de los dos contiene la señal. De esta manera, el sesgo se le aplica a ambas elecciones y se cancelan.

*Esto me recuerda a algunos métodos en eléctrica, en donde ambos canales de una señal se pasa por el mismo ruido, con el motivo que se cancelen.*



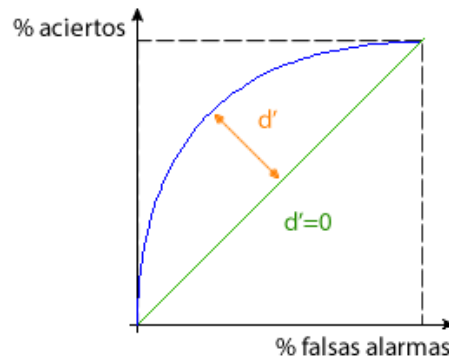


Figura 3.3: Curva de ROC

## 4. Resolución de frecuencia

El concepto de resolución tiene relación con la calidad de algo. Si se habla de resolución de una imagen digital, se refiere a la densidad de píxeles por área, es decir, frecuencia espacial. Así podemos decir que resolución se puede referir a la precisión a la cual un estímulo se puede distinguir de otro con el que son similares en alguna dimensión. Por lo tanto, resolución de frecuencias es la habilidad de separar una componente de un sonido de una masa compleja.

Claramente, el humano tiene un límite de percepción de frecuencia y tiempo. Además, como ya se vio anteriormente, nuestra percepción de intensidad depende de la frecuencia de la señal. Por otra parte, nuestro sistema auditivo ecualiza los sonidos para poder compensar otros, es decir, le damos más importancia a ciertas frecuencias.

Para poder graficar esto se realiza un **audiograma**. Este mide la pérdida auditiva respecto a un estándar. Cada persona escucha distinto, lo que se traduce en un audiograma diferente para cada persona. El audiograma posee variadas microvariaciones, lo que se puede apreciar si se realiza uno con una gran resolución, lo cual implica que el sistema auditivo también posee una gran resolución.

La resolución de frecuencias de nuestro sistema auditivo nos permite realizar varias tareas, como detectar un tono en un ruido (detectar una intensidad de una frecuencia predominante) o detectar un tono entre otros (algo complejo en otra cosa más compleja). En clases realizamos un experimento, el cual consistió en una señal compuesta de armónicos de 200 Hz, a los cuales se le iban quitando ciertos tonos. Al quitar y volver a colocar dichos tonos, el sistema auditivo empieza a analizar lo que escucha. Es decir, somos capaces de analizar e incluso descomponer los sonidos en sus distintos tonos.

*Es interesante ver la diferencia en el escuchar la mezcla de tonos al comienzo y la suma de distintos tonos al final, siendo que es la misma señal auditiva. Todo depende de como entrenamos al oído o con que disposición estamos escuchando lo que se nos presenta. Además, esto nos demuestra la capacidad de entrenamiento del sistema auditivo. Me pregunto hasta que punto se podrá llegar entrenando para descomponer los tonos.*

### Enmascaramiento y bandas críticas

Queremos estudiar el umbral de detección de la señal frente a una cierta máscara que se le presenta. Una máscara es un ruido que se presenta detrás de una señal para intentar confundirlo (enmascararlo). Cabe recordar que estos umbrales son definidos de forma estadística y siempre de acuerdo a cierta referencia. Para esto realizamos un experimento, el cual consiste en que se presenta una señal de 2000 Hz más un ruido de cierto ancho centrado en la señal presentada. La intensidad de la señal se va disminuyendo de 5 dB en 10 repeticiones. Debemos contar cuantas somos capaces de contar. Luego se modifica el ancho del ruido para ver las diferencias. Mis resultados se pueden ver en la tabla 4.1. Con este experimento, nos podemos dar cuenta del efecto del **enmascaramiento**: bajo cierta intensidad el ruido se “come” a la señal, por lo que ya no se escucha. Además, entre un ruido ancho de banda y el de 250 Hz no debiera haber mucha diferencia, ambos estorban casi lo mismo. En cambio el ruido de 10 Hz apenas molesta (enmascara). Resultados experimentales muestran que a partir de los 400 Hz de ancho, el enmascaramiento es el mismo, a pesar de que se aumente el ancho del ruido. El concepto de **banda crítica** muestra el umbral de intensidad según el ancho de ruido para todas las frecuencias.

El filtro auditivo (o banda crítica) de una señal nos muestra la porción de frecuencias a distintas intensidades que enmas-

Solo	Broadband	1000 Hz	250 Hz	10 Hz
10	6	5	5	9

Cuadro 4.1: Resultados experimentales

cara, como se puede ver en la figura 4.1. Este fenómeno se puede explicar con el mecanismo de la membrana basilar. Los sonidos viajan por la cóclea y buscan el lugar ideal para atravesar la membrana basilar. Pero para llegar a dicho lugar deben viajar a lo largo de un tramo de la membrana, estimulando ciertas porciones de ésta a medida que avanzan. Al estimular estas secciones, la membrana se vuelve menos receptiva a otros estímulos, se satura. Esto explica que no sea capaz de detectar tonos de muy baja intensidad que están bajo el filtro auditivo. Además, logra explicar la asimetría de estas bandas críticas. Los tonos más graves deben viajar a un lugar más cercano al ápex, por lo que estimulan los tonos más agudos cercanos a éste. En cambio, los tonos agudos se quedan más cercanos a la base, por lo que su banda crítica no estimula a tonos más graves. Esto se puede ver claramente en la imagen 4.2 <sup>7</sup>. Esta asimetría del enmascaramiento se probó en un experimento. Se tienen dos tonos, de 1200 y 2000 Hz, los cuales se reproducen simultáneamente, mientras uno de los dos baja progresivamente con un paso de 5 dB. El tono de 2000 Hz se enmascara antes que el de 1200 Hz, debido a la asimetría de las bandas críticas.

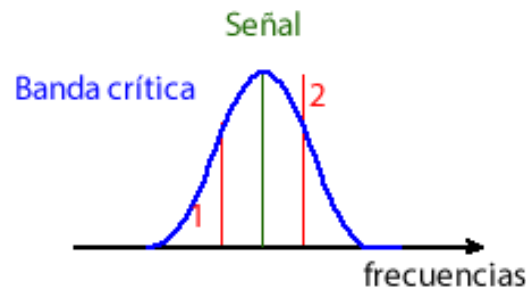


Figura 4.1: Una señal produce una banda crítica. En este caso, la señal 1 se ve enmascarada por la banda crítica. En cambio, la señal 2 no se enmascara al ser de mayor intensidad.

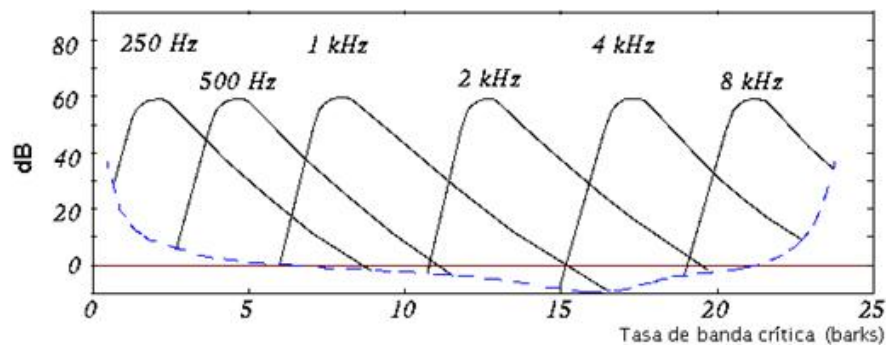


Figura 4.2: Bandas críticas para señales de distintas frecuencias. Cada señal afecta un área distinta, la cual enmascara a distintas intensidades.

Por otra parte, al comportarse como un filtro, cualquier ruido que estorbe la banda crítica, no importa mucho el lugar en el que lo haga, el “output” del filtro será el mismo. Esto se puede ver en la figura 4.3. Así, si utilizamos un ruido de ancho de banda que tenga un ancho libre, centrado en la señal (llamado ruido mellado o notch), podemos ver hasta donde llega la banda crítica para cada frecuencia, al ajustar el ancho libre del notch. A medida que envejecemos se ensanchan las bandas críticas. Esto quiere decir que los ruidos nos afectan más, ya que interrumpen una mayor área de nuestras bandas críticas.

*Se dijo en clases que el concepto de enmascaramiento y bandas críticas fue fundamental en el desarrollo del MP3. Esto se debe a que el MP3 busca comprimir los datos de tal forma de lograr un menor tamaño. Esto lo logra aprovechándose*

<sup>7</sup>[http://www.lpi.tel.uva.es/~nacho/docencia/ing\\_ond\\_1/trabajos\\_04\\_05/io1/public\\_html/MP3.htm](http://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_04_05/io1/public_html/MP3.htm)

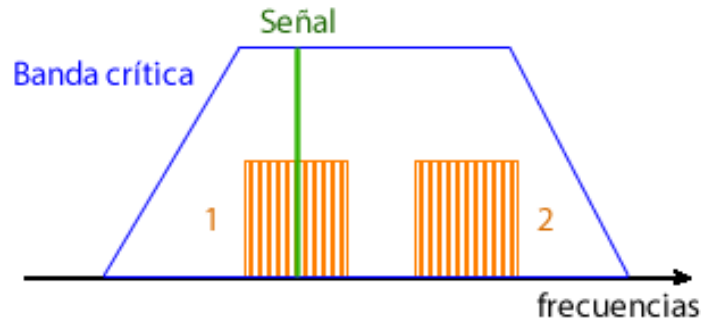


Figura 4.3: Una señal produce una banda crítica. En este caso, los ruidos 1 y 2 producen el mismo output del filtro, puesto que tienen la misma área sobre la banda crítica.

del enmascaramiento, puesto que elimina los datos de las frecuencias que se ven enmascaradas por el resto del sonido<sup>8</sup>. Encuentro que esta es una aplicación impresionante de los procesos psicofísicos, puesto que utiliza de tal forma los procesos de percepción humana que logra producir un formato de codificación altamente eficiente, con una mínima pérdida de calidad. Quizás, con una mayor comprensión de otros procesos, seamos capaces de lograr en tiempo real efectos que parecieran ser muy complejos, al aprovecharnos de lo que realiza el cerebro naturalmente.

Otra forma de medir las bandas críticas es mediante las **curvas de afinamiento psicofísico**. Se tiene una señal y una máscara y se sube progresivamente la intensidad de la máscara, hasta que esta logre enmascarar por completo a la señal. Con esto se puede obtener un espectro completo para una persona en particular. Además, nos permite ver los daños en nuestro sistema auditivo. Tanto en estas curvas, como en las bandas críticas vistas anteriormente, el daño auditivo se puede apreciar de distintas formas. Sí la forma del filtro queda igual, pero se corre hacia arriba, disminuye su eficiencia. Es decir, se necesita una mayor intensidad para poder captar los sonidos. El cambio de forma también supone un cambio en la eficiencia. Si la banda se vuelve más ancha, entonces es más probable que la señal se vea estorbada por algún ruido.

Dado que las bandas críticas y el enmascaramiento dependen de la intensidad del ruido y frecuencias que abarca, cabe preguntarse cómo varía el enmascaramiento con un aumento en la intensidad del ruido. Se ha demostrado experimentalmente que con un aumento en 10 dB de un ruido de banda ancha, se aumenta en 10 dB el enmascaramiento. Esto nos da una curva prácticamente recta, de pendiente 1. *Esta relación lineal creo que se podría explicar gracias que la pérdida de eficiencia en la audición se ve reflejada en mover la banda crítica hacia arriba, sin cambiar su forma.*

Otro fenómeno interesante es el **ajuste de la banda crítica** (off frequency listening), el cual se obtiene cuando el ruido sólo toma un lado de la banda crítica. En este caso, se produce la ilusión de que estamos escuchando una señal desplazada, que no corresponde a la señal verdadera, como se ve en la figura 4.4<sup>9</sup>. Pero si incorporamos otro ruido, esta vez por el otro lado, se pierde el efecto y escuchamos la señal a la frecuencia que corresponde. Esto se debe a una optimización de recursos por parte del cerebro para una mejor resolución e información de lo que nos rodea. El cerebro también juega un papel importante a la hora de estar esperando un sonido. Con la expectación, nos condicionamos a escuchar más fácilmente un sonido. Esto también se puede realizar artificialmente, reproduciendo un tono repetitivamente, para que el sujeto espere que suene nuevamente. Eso sí, con señales muy cortas no nos acostumbremos, puesto que el cerebro no las alcanza a procesar de la misma forma.

<sup>8</sup>[http://www.hpca.ual.es/~vruiiz/docencia/laboratorio\\_arquitectura/proyectos/01-02/E1%20formato%20MP3/fundamentos.html](http://www.hpca.ual.es/~vruiiz/docencia/laboratorio_arquitectura/proyectos/01-02/E1%20formato%20MP3/fundamentos.html)

<sup>9</sup> [http://en.wikipedia.org/wiki/File:Off\\_F\\_listening.svg](http://en.wikipedia.org/wiki/File:Off_F_listening.svg)

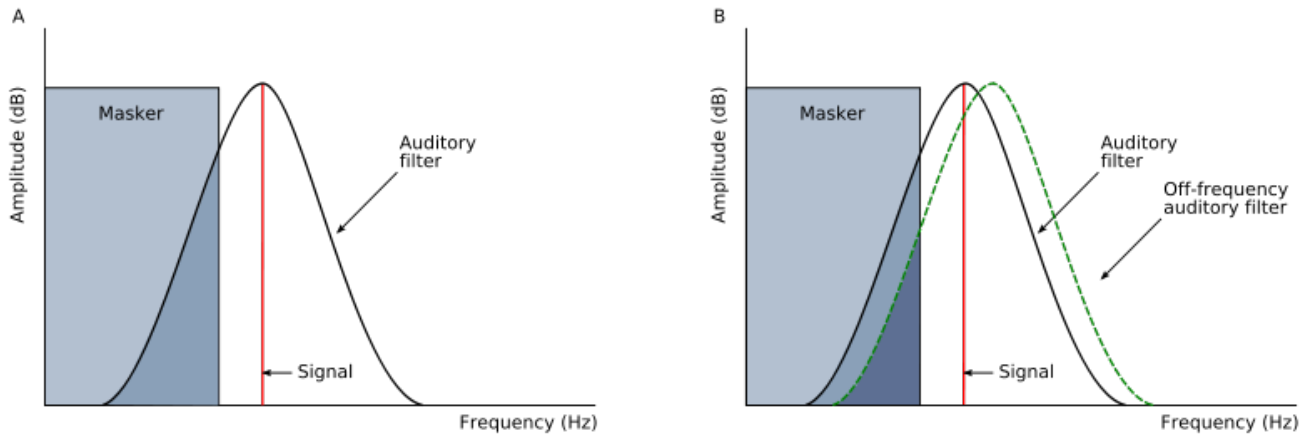


Figura 4.4: Una señal produce una banda crítica. En este caso, los ruidos 1 y 2 producen el mismo output del filtro, puesto que tienen la misma área sobre la banda crítica.

## 5. Percepción de alturas

El sonido tiene una gran cantidad de características, tanto físicas como perceptuales. Estas características pueden enlazarse entre sí. Es más, una característica física puede influenciar más de una característica perceptual, como se ve en la figura 5.1. Cabe mencionar que en nuestro idioma español, somos muy ambiguos para los términos perceptuales. Los términos utilizados son reciclados de otros contextos, muchas veces con más de una acepción para la misma palabra en el mismo contexto perceptual sonoro. En inglés los términos son bastante más claros. Para la “altura” del sonido se usa **pitch** y para el volumen o intensidad se usa **loudness**. La única excepción es el concepto de timbre, que no tiene un símil nativo anglosajón, ante lo cual adoptaron dicha palabra.

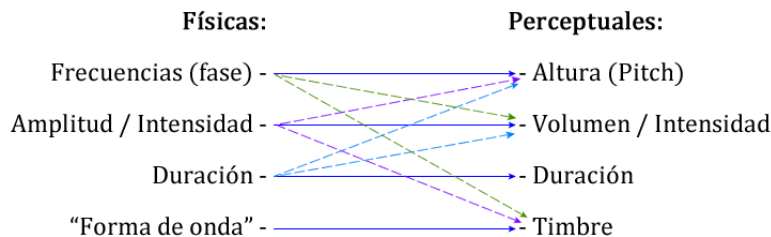


Figura 5.1: Relación entre las características físicas y perceptuales. La línea sólida representa la relación primaria, y las líneas discontinuas representan efectos secundarios.

Para ver la percepción de alturas realizamos un experimento, el cual consiste en diez grupos de cuatro pares de tonos con pequeñas diferencias entre ellos. Hay que establecer si los tonos suben o bajan con respecto a una referencia (1000 Hz) a medida que la diferencia entre estos baja. En un comienzo, fue bastante fácil detectar los cambios, pero ya en el sexto grupo se complica en sobremanera. Este experimento nos sirvió para confirmar la **Ley de Weber**, la cual dicta que el umbral de detección para una diferencia de intensidad, es proporcional a la señal original, es decir:

$$\frac{\Delta I}{I} = K$$

Matemáticamente, la función logarítmica cumple con dicha relación. Esto se puede explicar, en parte, debido a la distribución de frecuencias en la membrana basilar, aunque dicha teoría no lo logra explicar a cabalidad. Cabe mencionar que dicha ley se cumple en todo aspecto perceptual, no solo en la intensidad de sonido.

### Escala de pitch

¿En qué nos basamos para crear una escala de “alturas”? ¿Cómo podemos decir que una nota es más alta que otra? Existen dos teorías principales: Phase-locking y teoría del lugar.

1. **Phase-locking.** También llamada teoría temporal. Hay dos posibles explicaciones de dicha teoría, que difieren muy poco. Básicamente, se trata de que según la frecuencia de las velocidades del envío de señales al cerebro desde la

cóclea, éste es capaz de asignarle cierta sensación de altura. La otra explicación dice que es según el período de las señales. *Ambas explicaciones son lo mismo, teniendo en cuenta que periodo =  $\frac{1}{frecuencia}$* . El problema de esta teoría es que sobre los 5000 Hz no es posible tener sincronía entre la velocidad de las señales y la frecuencia del tono. Por lo tanto, esta teoría no logra explicar lo que sucede en este tramo.

2. **Teoría del lugar.** Esta teoría se fija en el lugar desde donde se disparan las señales en la membrana basilar. Según esto, el cerebro es capaz de interpretar a que altura se encuentran los tonos . Debido a las propiedades de la membrana basilar (como la rigidez) no funciona bien con frecuencias bajo los 5000 Hz.

Es posible darse cuenta que ambas teorías presentan problemas a partir de cierta frecuencia. Es por esto que el científico alemán August Seebeck desarrolló un experimento para poder dilucidar ciertas dudas. La sirena de Seebeck, figura 5.2<sup>10</sup>, deja pasar aire por alguno de sus agujeros, produciendo un tono a cierta frecuencia. Al abrir otros agujeros, podemos generar distintos tonos y según la ubicación en relación al primero producimos armónicos o tonos que no tienen relación con este. Así, a la hora de tener posibles resultados contradictorios con ambas teorías, es la teoría temporal la cual predomina.

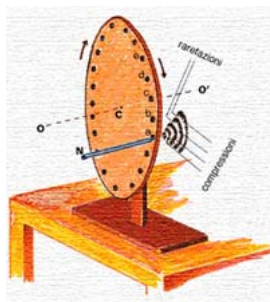


Figura 5.2: Sirena de Seebeck.

## Pitch virtual

El pitch virtual consiste en la apreciación de un cierto pitch, a pesar de que la onda fundamental no esté presente. Para este propósito realizamos un experimento, el cual consiste en una señal compuesta por 10 armónicos, a la cual se le van quitando progresivamente cada uno de estos componentes. Como resultado, se tiene que a pesar de que cambia el timbre del sonido, claramente el pitch no cambia, aún sin la fundamental. Otra particularidad es que los armónicos pueden no ser “perfectos”, pero el efecto es el mismo. Es decir, sin ser múltiplos enteros de la fundamental, el cerebro es capaz de detectar un cierto patrón y generar un pitch virtual de todas formas. Esto lo vimos en el experimento siguiente, en el cual, a pesar de que los armónicos no eran perfectos en la segunda ocasión, daban la ilusión de un pitch de todas formas, como se ilustra en la tabla 5.1.

	1200	1250 $\approx$ 1260
Frecuencia escuchada	1000	1050 $\approx$ 1050
	800	850 $\approx$ 840
Fundamental	200	?? $\approx$ 210

Cuadro 5.1: En el segundo conjunto, el cerebro interpretó lo escuchado y lo adaptó a algo con más sentido, obteniendo un pitch virtual de todas formas.

El pitch virtual también logra explicar que ciertos instrumentos, que no poseen tanta energía en la fundamental, de todas formas se logre percibir este pitch. Por ejemplo, el oboe tiene mucha más energía en el tercer armónico que en el primero, pero la nota se sigue escuchando con el pitch de la fundamental.

Otros efectos de pitch virtual se dan debido a la **no linealidad auditiva**. El sistema auditivo es no-lineal, esto quiere decir que se producen distintas distorsiones en el sistema auditivo, que no se pueden explicar con un análisis lineal y que son bastante difíciles de percibir. Un ejemplo de esto son los aural tones. Para ejemplificarlo, se realizó un experimento en clases. Se reproducen dos tonos de distinta frecuencia de 1000 hz ( $f_1$ ) y de 1200 Hz ( $f_2$ ), con lo cual se produce un aural

<sup>10</sup><http://digilander.libero.it/ricciardi/www.na.infn.it/Museum/schede/jpg/176b.jpg>

tone de 800 Hz ( $2f_1 - f_2$ ), como se puede apreciar en la figura 5.3. Este tono no es fácil de distinguir, por lo que se agrega un tono de 804 Hz, ilustrado en la figura 5.4. Al estar tan cercano al tono aural, se produce un efecto de abatimiento, en el cual se siente como si la intensidad del sonido vibrara. Normalmente, este efecto se da entre tonos con una frecuencia parecida, lo que comprueba la existencia del aural tone, puesto que no se está tocando realmente ningún tono de 800 Hz. Este efecto también se puede lograr mediante el **cross-ear pitch**, es decir, poniendo un tono en cada canal auditivo. Esto sucede, ya que el efecto sucede a nivel del sistema auditivo perceptual, en el cerebro, no de forma física.

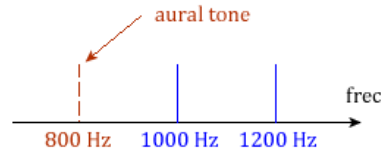


Figura 5.3: Los tonos de 1000 Hz y 1200 Hz, producen un tono aural de 800 Hz.

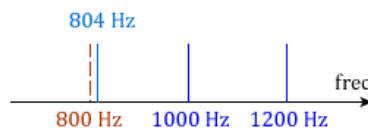


Figura 5.4: Se incluye un tono de 804 Hz, para causar un aural beat o abatimiento aural, el cual hace que la intensidad del sonido vibre debido a un acoplamiento con el tono aural.

También se puede experimentar **enmascaramiento en tonos virtuales**. Al crear un pitch virtual con ciertos armónicos y comparándolo con el tono real, en ambos casos se siente el mismo pitch. Ahora, si un ruido enmascara la fundamental, pero no los armónicos, se sigue sintiendo el pitch gracias a estos. Así mismo, si se enmascaran los armónicos, pero no la fundamental, se siente esta última, pero el pitch virtual desaparece.

*Es muy interesante este efecto, puesto que nos muestra que somos capaces de seguir escuchando cierto tono, a pesar de que tengamos ruido que enmascare la fundamental. Esto mismo se podría utilizar en la codificación de música, tal como se vio anteriormente en la codificación MP3, para no perder información aun sin tener que codificarla. Sería interesante ver el efecto de hacer música solo con pitch virtuales.*

Estos efectos del pitch se pueden complementar con otros resultados. Es interesante notar que el cerebro puede escuchar tanto **analítica** como **sintéticamente**. La forma analítica descompone el sonido y lo ve por partes, mientras que la forma sintética lo ve de una forma global. Esto se puede ver en un experimento en el cual se presentan dos sonidos complejos, compuestos por dos tonos puros. El primero está compuesto por tonos de 1000 y 800 Hz, mientras que el segundo por tonos de 1000 y 750 Hz. Al escuchar los tonos de forma analítica, percibimos que el segundo sonido es más grave que el primero, puesto que el tono de 1000 Hz se mantiene, y el otro tono baja en 50 Hz. En cambio, si escuchamos los tonos de forma sintética, percibimos que el segundo tono es más agudo, gracias al efecto del pitch virtual. Al verlo de forma global, el primer sonido genera un pitch virtual de 200 Hz, mientras que el segundo produce un pitch virtual de 250 Hz, claramente más agudo.

## Representación y circularidad

Del pitch podemos rescatar dos variables muy importantes: La altura (height) y el chroma. La altura es la que nos permite saber que nota es más aguda o más grave que otra. Pero al mismo tiempo, tenemos el efecto de que un do, suena como do sin importar en la escala en que se encuentre, tiene el mismo “color”. A esta propiedad se le denomina chroma. Como el chroma se repite en cada escala de forma periódica, podríamos representarlo de forma circular. De esta forma, una buena representación del pitch es una helicoide, en la cual varía la altura a lo largo del eje, y el chroma de forma radial, tal como se muestra en la figura 5.5<sup>11</sup>.

Analizando esta representación es posible explicar fácilmente otro fenómeno denominado **circularidad**. Se tiene una serie de 10 tonos cíclicos, los cuales varían de forma ascendente o decreciente en un semitono y se pasan por un cierto filtro como el que se ve en la figura 5.6<sup>12</sup>, se obtiene la ilusión de que se está subiendo o bajando eternamente, sin lograr llegar

<sup>11</sup>[http://nonsite.org/wp-content/uploads/2011/06/Kane\\_Figure-5.jpg](http://nonsite.org/wp-content/uploads/2011/06/Kane_Figure-5.jpg)

<sup>12</sup><http://asa.aip.org/gif/demo27a.gif>

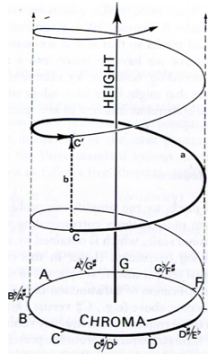


Figura 5.5: Representación gráfica del pitch.

a ninguna parte. Esto se debe a que cada 12 pasos, una nota llega a su siguiente octava, pero al mismo tiempo, hay tonos nuevos que ocuparon su lugar anterior. Por esta razón, la sensación de subir es real, pero tampoco será posible avanzar realmente. Usualmente, este efecto se compara con las escaleras infinitas del pintor M. C. Escher.

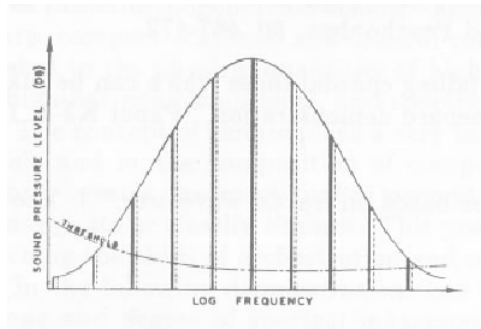


Figura 5.6: Filtro usado en una escala de Shepard.

## Comentarios finales y teoría moderna

De todos los efectos mencionados acerca del pitch, es importante volver a las teorías que intentan explicar el como se detecta el pitch. Cada una tiene evidencia en contra fundamental, como se puede ver en la tabla 5.2. Es por esto que se ha desarrollado una **teoría moderna del pitch**. Esta se basa principalmente en un procesador del pitch, el cual funciona a nivel cerebral y en dos etapas. En primer lugar analiza las componentes individuales de la señal de audio para luego sacar una conclusión de acuerdo a los datos obtenidos. En este último punto se generan los pitch virtuales y efectos no lineales. *A mi no me convence del todo esta teoría. Claramente refleja todo lo que se sabe acerca del pitch, pero dista de explicarlo. Se ve más como un intento de asignarle a una parte misteriosa del cerebro atribuciones y funciones que sabemos que suceden. Para mi falta descubrir donde se ubica dicho procesador y realizar experimentos que avalen dichas funciones. En caso contrario, es casi como “cocinar” un resultado, dado que sabemos la respuesta.*

Evidencia en contra	
Teoría temporal	Teoría del lugar
Región dominante	Escucha cerca del umbral
Cross-ear pitch	Enmascaramiento
	Pitch shift

Cuadro 5.2: Evidencia en contra de cada teoría clásica.

## 6. Intensidad perceptual

Al igual que sucede con el concepto de “altura” del sonido, no poseemos un concepto claro para la intensidad del sonido, por lo cual se prefiere usar el nombre del concepto en inglés: Loudness. Lo primero que queremos saber es el como somos capaces de discriminar distintas intensidades.

### Nivel del Loudness

En primer lugar es útil definir una forma de medir la intensidad de sonido o loudness a nivel perceptual. Vamos a presentar dos medidas de loudness parecidas entre sí: fonos (phon) y sonos (sone).

1. **Fono (phon):** Un fono es el nivel de loudness de un sonido que equivale a un tono de 1000 Hz en dB SPL. La intensidad de los tonos se compara con uno de 1000 Hz para poder compensar el efecto que tiene la frecuencia en la intensidad percibida.

Para medir en fonos, se debe ajustar la intensidad de un tono de 1000 Hz, hasta lograr la intensidad del tono reproducido. O también se puede ajustar el tono reproducido para lograr una igual loudness que un tono de 1000 Hz. De ambas formas podemos generar contornos de igual loudness, como se puede apreciar en la figura 6.1<sup>13</sup>.

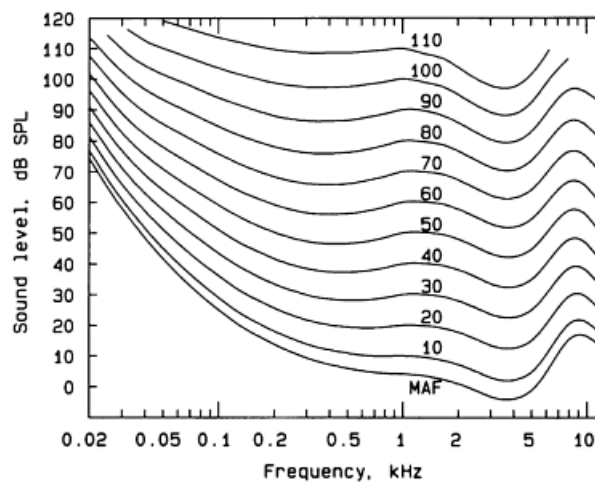


Figura 6.1: Contornos de igual loudness. Cada curva posee la misma intensidad perceptual.

2. **Sono (sone):** Un sono se define como la loudness de un tono de 1000 Hz a 40 dB SPL. Esta medida se usa para tener un punto de referencia, el cual nos permite tener cierta propiedad en la medida, la cual se discutirá más adelante.

### Efectos sobre el loudness

Uno de los efectos que se tiene en la percepción de loudness es la **sumación de loudness**. Es un efecto que se relaciona con las bandas críticas y muestra que con un ruido a igual intensidad física, aumentando el ancho de banda se percibe más fuerte. Esto sucede, ya que al aumentar el ancho de banda se estimulan más bandas críticas, lo que da la sensación de estar frente a un sonido más fuerte.

Otro efecto que tiene que ver con la percepción de loudness es el **corrimiento de umbral**. Cuando nos vemos expuestos a sonidos particularmente intensos, nos volvemos más sensibles. Esto causa un corrimiento del umbral de intensidad. *Este es uno de los fenómenos que más fácil se entiende, puesto que nos vemos expuestos a el todos los días. La diferencia entre el umbral cuando recién nos despertamos, luego de una noche calmada y al final del día, especialmente después de haber pasado por el metro, es abismante. Es muy valioso poder darle cierto significado a eventos tan cotidianos como este.*

Así como se puede correr el umbral de intensidad, también podemos sufrir un **corrimiento de pitch**, según la loudness. Dependiendo de la intensidad con la que presentemos un tono, este lo podemos llegar a percibir hasta media octava más alto. *Este es un efecto completamente perceptual, puesto que lógicamente, la frecuencia física de un tono no cambia, por más que cambie su amplitud.*

<sup>13</sup>[http://replaygain.hydrogenaudio.org/proposal/pics/equal\\_loudness.gif](http://replaygain.hydrogenaudio.org/proposal/pics/equal_loudness.gif)



## Medición y escalamiento de Loudness

La medida de loudness en sonos nos da una forma de hacer fácilmente una escala de loudness. Experimentalmente se ha descubierto que para sonidos sobre 40 dB SPL, si se aumenta la intensidad en 10 dB, se dobla la loudness. De esta forma se llegó a la siguiente escala:

$$L = k I^{0,3}$$

En donde  $L$  es la intensidad en loudness (sonos),  $I$  es la intensidad física y  $k$  es una constante. Para ver que tan ajustado es esta expresión, se hizo un experimento en el curso, en el cual se presentan sonidos para ser comparados con una referencia y nosotros debemos determinar cuántas veces más fuerte es dicho sonido. El resultado fue una curva exponencial, la cual ratifica la expresión antes señalada.

## 7. Audición espacial

Nuestro sistema auditivo nos permite hacer sentido espacial de los sonidos que escuchamos, ya sea de fuentes cercanas o muy lejanas. La localización lograda por nuestro sistema auditivo es en tres dimensiones. Nos permite tener una idea de la elevación, distancia y azimuth de la fuente, como se muestra en la figura 7.1<sup>14</sup>.

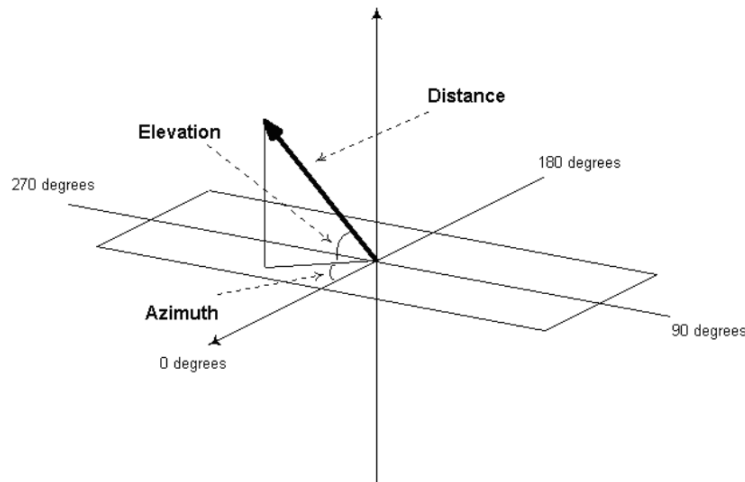


Figura 7.1: Espacialización del sonido.

Para estudiar la audición espacial requerimos de ciertas condiciones especiales. Necesitamos prevenir cualquier interrupción del sonido y aislarnos completamente de cualquier rebote posible, es decir, necesitamos una aislación acústica completa, un efecto de campo libre. Con este fin se utilizan salas especialmente acondicionadas para tal fin, llamadas salas anecoicas. Estos experimentos buscan entender que sucede con un tono que se mueve alrededor de una persona, como se distorsionan con nuestro cuerpo y las diferencias que se tienen en cada canal auditivo. Debido a esto, se hace un modelo de una cabeza, a la cual se le colocan micrófonos en el lugar de los tímpanos, cosa que se pueda grabar tal como si estuviésemos ahí. Esta técnica se llama **grabación binaural**.

Gracias a estos experimentos se ha podido determinar y caracterizar el proceso de espacialización de los sonidos, junto a todos los problemas que posee. Según el lugar de donde provenga un sonido, a cada tímpano le llega un sonido con un espectro distinto, es por esto que es de sumo interés saber claramente que sucede. Una de las cosas que se desea determinar es el mínimo ángulo audible, es decir, para qué diferencia angular sentimos una variación en la localización. También se han podido identificar varios planos auditivos, como se ve en la figura 7.2<sup>15</sup>. Por otra parte, uno de los problemas de la localización es que a ciertas frecuencias este proceso empeora. Cometemos errores en las frecuencias para las que somos más sensibles.

### Teoría dual de la localización de sonidos

La teoría dual de la localización de sonidos dice que los sonidos se ven afectados de dos formas distintas y simultáneas al cambiar de lugar. Al moverse un sonido a la derecha, la señal que llega al oído izquierdo sufre una disminución de intensidad y un desfase lo cual nos da las claves necesarias para detectar el lugar del cual proviene. Esta diferencia temporal se llama **Interaural Time Difference (ITD)** y la diferencia de loudness se llama **Interaural Level Difference (ILD)**. Una representación de lo mismo se puede ver en la figura 7.3<sup>16</sup> o más específicamente para la ITD en la figura 7.4<sup>17</sup>.

Cabe mencionar que dependiendo de la frecuencia de los sonidos, cuánto cambia la intensidad y el desfase con el ángulo de precedencia. Las frecuencias bajas tienen una menor diferencia en intensidad (ILD) debido a su longitud de onda.

Un modelo del sistema auditivo realizado para intentar explicar la localización fue realizado por Lloyd Jeffress en 1948. Este consiste en que la señal entra a cada canal auditivo y a la hora de procesarla, busca coincidencias entre la señal

<sup>14</sup>Fuente: Clases

<sup>15</sup>Fuente: Clases

<sup>16</sup>Fuente: <http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/localization/localization-slides/Slide16.jpg>

<sup>17</sup>Fuente: <http://upload.wikimedia.org/wikipedia/commons/5/5c/ITD.jpg>

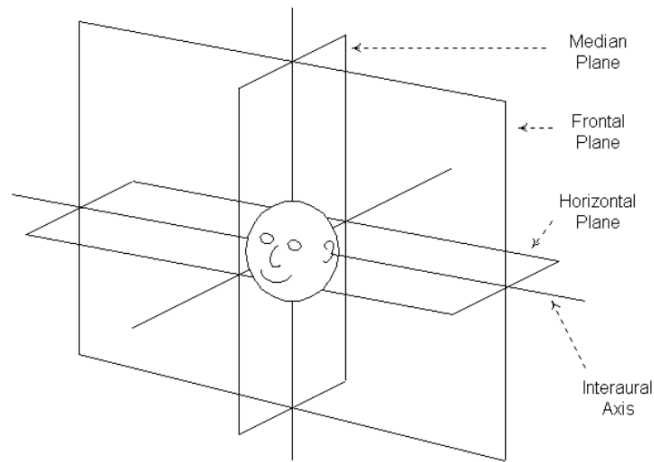


Figura 7.2: Planos de audición.

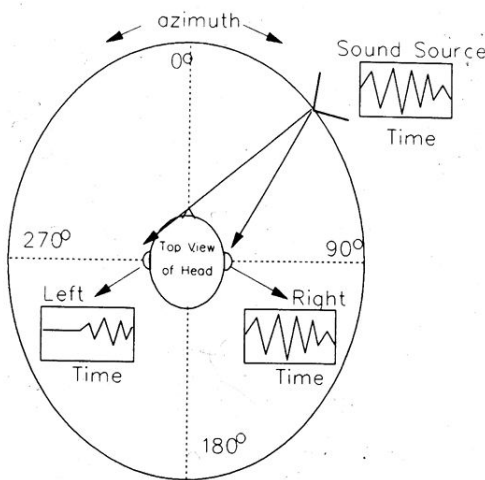


Figura 7.3: Representación de ITD e ILD.

de cada oído. Estas coincidencias las evalúa en diversos puntos, según el tiempo de llegada de la señal, como se puede modelar en la figura 7.5<sup>18</sup>. Así, una señal que llega por el lado izquierdo primero, recorre una mayor cantidad de puntos de coincidencia que la señal del lado derecho. Cuando ambas señales se encuentran en uno de estos puntos, el cerebro es capaz de concluir de qué lado provino la señal.

Uno de los problemas a los cuales nos enfrentamos a la hora de localizar sonidos se da por el llamado **cono de confusión**. Este consiste en que desde el oído se proyecta una superficie cónica, en la cual todos los sonidos nos llegan de forma indistinguible. El eje de estos conos calza con el eje interaural mostrado en la figura 7.2. Una de las particularidades de este cono es que no nos permite diferenciar sonidos sobre el plano medio, puesto que los conos de ambos oídos se intersectan ahí. Esto causa que sea difícil saber si un sonido está adelante o detrás de uno (*esto es asumiendo que estamos en una sala anecoica, puesto que los rebotes pueden dar información adicional*). Sin embargo, existe un mecanismo muy fácil para contrarrestar este efecto. Al mover la cabeza, el espectro de los sonidos que nos llegan cambia, lo cual es interpretado por el cerebro y nos puede dar la información precisa del lugar de origen de la fuente.

### Head Related Transfer Function (HRTF)

Esta función trata de describir cómo nos llegan los sonidos a cada tímpano. Para esto toma en consideración la ubicación de la fuente sonora y el efecto de la misma cabeza sobre las ondas de audio. En general, estas curvas se han obtenido de forma experimental (las más precisas al menos), realizando una gran cantidad de mediciones. Un ejemplo se puede ver en

<sup>18</sup>Fuente: [http://www.scholarpedia.org/article/Jeffress\\_model](http://www.scholarpedia.org/article/Jeffress_model)

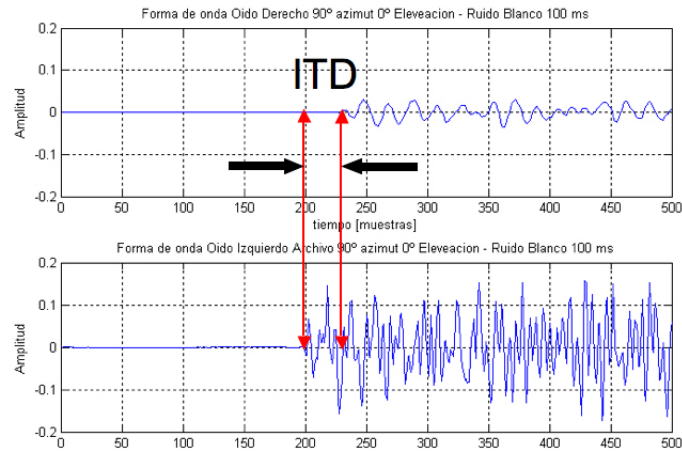


Figura 7.4: Formas de onda en cada oído, mostrando la ITD.

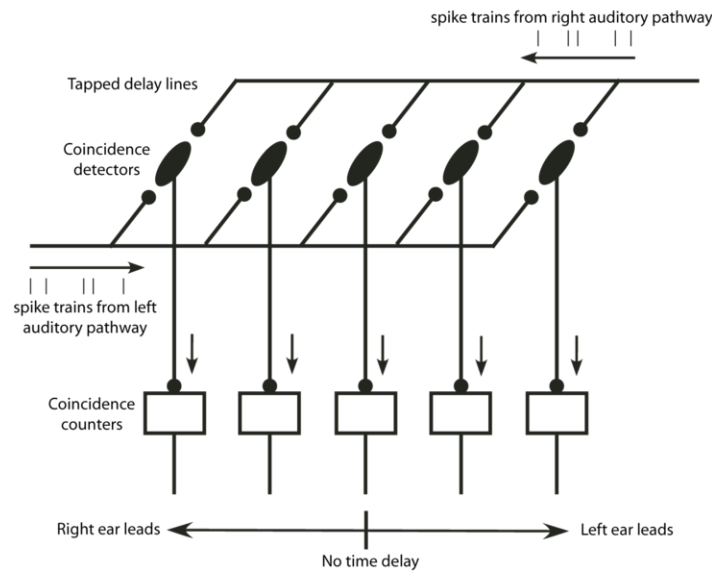


Figura 7.5: Esquema del modelo de líneas de retraso de Jeffress.

la figura 7.6<sup>19</sup>.

Gracias a estas curvas ha sido posible el generar audio binaural sin la necesidad de realizar la tarea de grabar el sonido con la técnica descrita anteriormente. Con este método, basta con tener las curvas HRTF necesarias para el efecto que se quiere lograr, y el sonido que se quiere tratar. Mediante la operación de la convolución se logra obtener la señal definitiva para cada oído. Eso sí, este método, y en general el audio binaural, no está exento de problemas, los cuales incluso lo han llevado a fracasar comercialmente (hasta ahora).

1. En primer lugar las curvas son distintas para cada persona, puesto que dependen de su sistema auditivo en particular y de cómo su cabeza distorsiona las ondas. En general, se usan curvas promedio, lo cual afecta la calidad del sonido.
2. Esto sólo funciona cuando escuchamos el audio con audífonos. Al usar parlantes, se obtiene un fenómeno llamado cross-talk. Es decir, las ondas de sonido de cada canal se superponen, generando efectos no deseados. Sin embargo, este problema ya ha sido estudiado y corregido, aunque solo si el espectador se encuentra en cierta zona con respecto a los parlantes.
3. El último, y más grave problema, es que estas funciones dependen de la posición de la cabeza de la persona. Como estas curvas cambian con la posición relativa del sonido, si giramos nuestra cabeza, el sonido debería cambiar también.

<sup>19</sup>Fuente: <http://alumnus.caltech.edu/~franko/thesis/hrtffreq.gif>

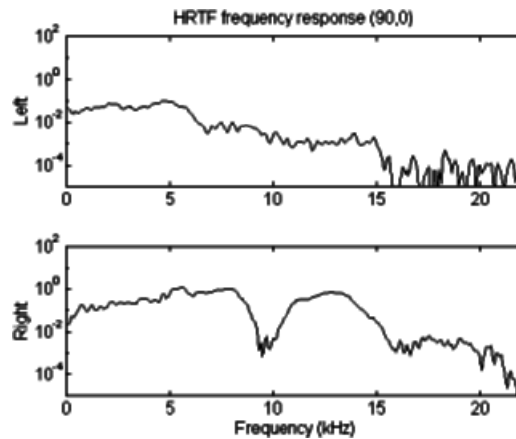


Figura 7.6: Ejemplo de curvas HRTF.

Entonces, necesitamos usar algún sistema de head-tracking, el cual nos permita tener pistas dinámicas calculadas en tiempo real. Esta es una de las razones por la cual los sistemas de audio “surround” que se venden, no son sistemas de audio 3D.

Es interesante notar que cuando se escucha música normalmente con audífonos, la música se escucha dentro de la cabeza de uno mismo. El hecho de estar escuchando con audífonos y que no se localice el audio en ningún lado, hace que se escuche dentro de la cabeza. El único sonido que se escucha así de forma natural es nuestra propia voz. Cuando movemos la cabeza, al estar localizada dentro de nosotros, la voz no cambia. Es por esto que es tan extraño escuchar nuestra propia voz desde una fuente externa.

Finalmente, es importante estudiar el efecto de los sonidos cuando hay eco. En la realidad, las señales de audio rebotan lo que causa que a nuestro sistema auditivo nos llegue una gran cantidad de señales de audio repetidas, variando su intensidad y frecuencia, debido a que nos llegan de varios lugares distintos. Sin embargo, no percibimos esa infinidad de señales repetidas. El sistema auditivo es capaz de detectar dichos patrones, y al tener menor intensidad y cierto desfase, asociarlos a un sonido anterior original. De esta forma, suprime las diferentes repeticiones de dicho sonido dándole más sentido a lo que escuchamos. Esto es llamado **efecto de precedencia**. Probamos esto en clases mediante un experimento, en el cual se escucha un golpe de martillo en tres condiciones de eco distintas y luego se reprodujeron al revés. Según el tiempo de retardo del sonido, lo escuchamos como eco o simplemente no lo escuchamos. Pero si se invierten los sonidos, es decir más débil primero y más fuerte al final, escuchamos todos los sonidos, ya que no tiene sentido que un eco venga antes que el sonido original.

## 8. Percepción de timbre

El timbre puede ser descrito como la característica personal, el sello de un sonido. Está determinado por varios aspectos, tales como la forma de la onda o el contenido y composición de armónicos, entre otros. Es decir, es un aspecto multidimensional del sonido, mucho más que las otras características previamente analizadas, lo que hace que no podamos hacer una “escala” de timbre. Tampoco es posible escribirlo de forma específica en la música, no tiene una notación. Lo que se hace habitualmente es describir el sonido con palabras. Algunas personas lo intentan definir como todo lo que no es pitch ni loudness, es decir, lo que nos permite distinguir dos sonidos como distintos a pesar de tener mismo pitch y loudness. El timbre también tiene que ver con la estructura del sonido, pudiendo analizar su trayectoria y espectro. Las partes transiente y de reposo definen gran parte del timbre, los cuales representan el ataque y “sustain” de los sonidos. Por otra parte, el brillo del sonido se ha relacionado con el centroide del espectro.

### Estudios multidimensionales

A pesar de no poder hacer una escala de timbre, es posible intentar hacer un mapa multidimensional de timbres. Generalmente, se hace un mapa tridimensional en donde se van juntando sonidos parecidos, es decir con un gran “similarity rating”. Además de este rating, se usan etiquetas VAME (Verbal attribute magnitude estimator). Estos estudios se han podido realizar gracias a los computadores y la síntesis de sonido, los cuales nos permiten mantener loudness y pitch en niveles iguales y modificar a gusto los diferentes parámetros. Así, es posible realizar estos estudios y crear escalas multidimensionales (MDS). El gran trabajo luego es el poder interpretar los ejes de los espacios formados. Uno de los trabajos más famosos del timbre es el realizado por Grey, cuyo resultado se puede apreciar en la figura 8.1. A partir de este estudio, se han realizado varios otros que lo toman como base. Este es el caso del estudio de McAdams, el cual incorpora más instrumentos en el mismo análisis como se muestra en la figura 8.2.

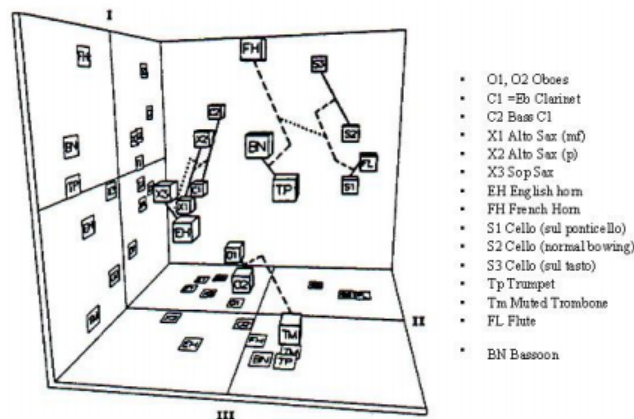


Figura 8.1: Estudio multidimensional de Grey (1977)

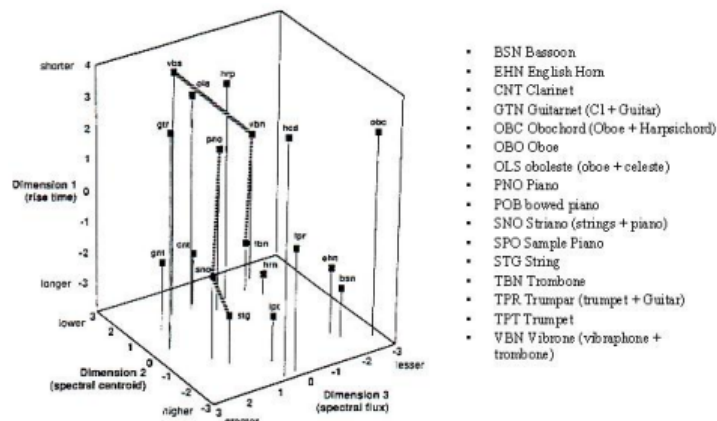


Figura 8.2: Estudio de McAdams, el cual incorpora más instrumentos al estudio de Grey (1995)

Se han realizado más investigaciones para estudiar la composición del timbre de otras formas. El profesor Rodrigo Cádiz junto a Lipscomb hizo un estudio con un enfoque más científico. Trabajaron con doce combinaciones, producto de cuatro envolventes temporales junto a tres formas de onda, presentados en la tabla 8.1. Estas combinaciones fueron presentadas a varios grupos de personas, pidiéndoles que describieran los sonidos de acuerdo a la escala de palabras de VAME, las cuales se pueden ver en la figura 8.3<sup>20</sup>. Lamentablemente, no se pudieron obtener resultados concluyentes de este estudio, aunque los resultados del grupo de gente con entrenamiento musical eran bastante prometedores. Y esto mostró la diferencia que hace el entrenamiento musical en la escucha de sonidos, como se verá más adelante.

Envolventes temporales	Formas de onda
Ataque impulsivo (IA)	Nota
Ataque cerrado (CA)	Nodo
Ataque abierto (OA)	Ruido
Graduado (G)	

Cuadro 8.1: Características de los sonidos

Total absence	Total presence
Not complex	Complex
Not fast	Fast
Not graduated	Graduated
Not linear	Linear
Not nasal	Nasal
Not open	Open
Not rich	Rich
Not sharp	Sharp
Not smooth	Smooth
Not swelled	Swelled

Figura 8.3: Palabras usadas para VAME

Otro punto de importancia es el comportamiento del timbre en el tiempo, puesto que los timbres no son estáticos, sino que dependen de la evolución de armónicos en el tiempo. Parte de los parámetros que afectan es la variación de la energía en cada armónico, cuales están presentes en que instante o que frecuencias se enfatizan en el tiempo. Esto se debe a que cada banda de frecuencia no tiene por qué comportarse igual que el resto en el tiempo, lo que provoca que el timbre vaya evolucionando. Un ejemplo de esto se puede ver en la figura 8.4<sup>21</sup>.

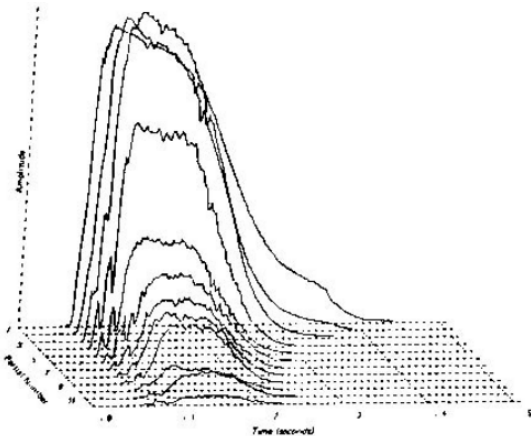


Figura 8.4: Variación de la intensidad en el tiempo de distintos armónicos del sonido de una trompeta

<sup>20</sup>Fuente: Clases  
<sup>21</sup>Fuente: Clases

Por otra parte, también interesa el comportamiento del timbre en distintas notas en el mismo instrumento. El timbre no se traspone simplemente, sino que varía su estructura según la nota tocada. Esto le pone una dificultad extra al sintetizar sonidos, puesto que debemos conocer la variación de la estructura del timbre para cada frecuencia fundamental.

*Esto lo encontré muy interesante desde el punto de vista de que como somos capaces de percibir dos notas como iguales en diferentes escalas, si el timbre varía también y que además pertenezcan al mismo instrumento. Sería interesante estudiar la autocorrelación del timbre en diferentes instrumentos, viendo su parecido al variar la frecuencia.*



## 9. Análisis de la escena auditiva

Este trabajo ha sido realizado principalmente por el psicólogo canadiense Albert Bregman y presenta una teoría completa y unificada de todo el fenómeno y cómo entendemos la escena auditiva. La percepción de escena está muy estudiada a nivel visual, llegando incluso al punto en el cual se han podido programar robots que son capaces de ver e interpretar lo visto. Sin embargo, todavía no somos capaces de fabricar robots que escuchen y entiendan lo escuchado. Este estudio está basado en gran parte en su contraparte visual, tomando varios conceptos de la psicología de Gestalt. Tiene tres componentes fundamentales: segmentación, integración y segregación.

Para lograr la comparación entre la parte visual y la auditiva, hay que reconocer la unidad básica de reconocimiento. Así como en la visión reconocemos distintos objetos visuales, los cuales agrupamos y tratamos como unidad, en el sistema auditivo tenemos “streams” auditivos. Por ejemplo, mientras que visualmente tenemos el fenómeno de size consistency, en el cual un objeto no cambia de tamaño a medida que se acerca, en la audición tenemos timbre consistency. Este fenómeno dice que la voz de una persona no cambia por el ruido que se encuentre en el ambiente. En fin, este estudio da cuenta de como interpretamos la realidad de acuerdo a ciertos criterios. Algunos de los principios gestálticos en la audición son los siguientes:

- **Continuidad:** Si tenemos un sonido con un vacío, lo escuchamos como dos streams distintos. En cambio, si el vacío lo rellenos de un ruido, lo percibimos como un solo stream. Es equivalente a ver un objeto a través de una reja que tapa partes de este. El objeto no lo percibimos como si estuviese cortado, sino que sabemos que tiene continuidad.
- **Proximidad:** Consiste en que si tenemos tonos de pitch y loudness parecidos, bajo ciertos criterios los percibimos como si fuesen del mismo stream.

Por otra parte, nuestro sistema auditivo es capaz de aislar gran parte del ruido externo para concentrarnos en una sola cosa, intentar percibir bien una fuente de sonido. Hacer esto en un computador es muy difícil, puesto que no sabemos de donde viene cada sonido. Además, los streams de cada fuente se unen en el espacio, por lo que a nuestro canal auditivo llega una gran cantidad de streams unidos, distorsionados y con diversos ecos, lo que hace que teóricamente sea de gran dificultad poder distinguir cada parte por separado. Esto se puede ver con una muy buena analogía. En el borde de un lago hacemos dos surcos largos y angostos en los cuales ponemos una hoja que vibra con las ondas en el agua. Si analizamos el movimiento de las hojas es prácticamente imposible responder a preguntar como cuántos botes o barcos se encuentran presentes, cuál está más cerca o si hay viento o si ha caído un objeto o muchas preguntas más. Pero, haciendo la analogía entre los surcos con la hoja con nuestro canal auditivo y tímpanos, vemos que somos capaces de hacer todo esto y de forma rapidísima, casi inmediata!

*Es importante tener en mente la esencia de la percepción. Esto es como cada uno ve el mundo, puede no ser así necesariamente. Es por esto que se dan ciertas “mentiras” perceptuales. Cosas que no están completas las vemos o escuchamos como si lo estuviesen, o encontramos patrones que asociamos a streams. Todo esto es debido a que el cerebro intenta hacer sentido de la información visual y auditiva que nos llega, aunque esto désignifique distorsionar la realidad de cierta forma.*

El análisis de la escena auditiva nos dice que la percepción auditiva tiene como objetivo el poder indentificar o estimar las fuentes sonoras de la mezcla de streams que nos llega. Si percibimos un patrón estructuralmente intacto, este debe provenir de una fuente. En cambio, si percibimos N patrones intactos, entonces tenemos N fuentes diferentes. El estudio de Bregman se centró en gran parte en determinar qué determina que un patrón se perciba estructuralmente intacto. La teoría de respuesta de Bregman asume un proceso en dos etapas: procesos primitivos y procesos esquemáticos. Los procesos primitivos son procesos automáticos, suceden sin esfuerzo alguno y permiten detectar grupos perceptuales. Los procesos esquemáticos con procesos aprendidos, los cuales requieren de esfuerzo al prestar atención a ciertos grupos formados anteriormente.

En clases se realizaron varios experimentos para ver los resultados de Bregman. Estos dependen del método de integración auditiva que usemos. Estos pueden ser tanto por integración temporal como espectral. Estos experimentos fueron:

- **Integración temporal**
  - **Tonos intercalados**, segregación de streams
  - **Patrones con interrupciones**, camuflaje auditivo
  - **Pérdida de información rítmica**, segmentación y ritmo
  - **Streaming de distintos flujos**

- **Efecto de rango de altura** en la separación de streams
- **Efecto de timbre** en la separación de streams
- **Efecto de la posición espectral** en la separación de streams
- **Efecto auditivo de la conexión**, conectividad auditiva
- 
- **Integración espectral**
  - **Argumento de frecuencia fundamental común**
  - **Modulación de frecuencia común**
  - **Modulación micro en percepción de voz**

## 10. Voz humana y habla

La voz humana es un proceso bastante complejo. Desde un punto de vista anatómico, tenemos nuestra fuente de poder son los pulmones, los cuales empujan el aire hacia las cuerdas vocales. Éstas producen los sonidos “hablados” al vibrar con el paso del aire, produciendo ondas sonoras complejas. Luego, estos sonidos pasan al tracto vocal, compuesto por la laringe y cavidades bucal y nasal. El tracto vocal posee ciertas características únicas en cada persona y es lo que da el timbre personal, puesto que posee ciertas resonancias y amplificaciones de ciertas frecuencias. Estas resonancias se llaman **formantes** y se clasifican de acuerdo a ciertas alturas. Los sonidos no hablados son como ruidos producidos como turbulencias u obstrucciones de este tracto. Por otra parte, las diferentes vocales y consonantes se producen de formas distintas. Los sonidos vocales se producen en distintas partes de la cavidad bucal, como se ve en la figura 10.1 <sup>22</sup>, mientras que las consonantes se agrupan según sus características, como se ve en la tabla 10.2 <sup>23</sup>

Los **fonemas** son los grupos de sonidos que se clasifican como idénticos. Estos son las letras de cada idioma. La letra ‘p’ no suena igual en todas las palabras, pero dichos sonidos los asociamos siempre con esa letra. En este caso se dice que esos sonidos son **alófonos** de dicho fonema. Por otra parte, tenemos los elementos **fonéticos**, los cuales representan el sonido producido al hablar. En este caso se hace la distinción en cada inflexión que se pueda producir. Acá se pueden introducir diferencias en cada idioma y como afectan a nuestra percepción. Si en un idioma no se hace la distinción entre dos sonidos y se tratan como alófonos, entonces esas personas no serán capaces de distinguir dicha diferencia en un idioma distinto. Esto muestra que el proceso del habla es aprendido y los sonidos que distinguimos dependen de nuestra cultura idiomática.

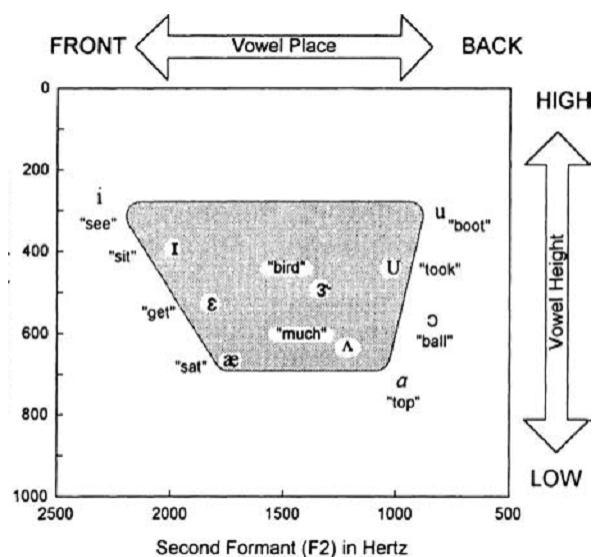


Figura 10.1: Formación de vocales

Los formantes son muy relevantes para el análisis de la voz humana, puesto que son éstos los que nos dan la información necesaria para identificar y procesar correctamente lo que se dice. Los dos formantes más bajos son los encargados de determinar la identidad de la mayoría de las vocales, y pueden cambiar hasta en dos octavas. Se puede ver el contenido de frecuencias de varias palabras, junto a la forma de la cavidad bucal asociada en la figura 10.3 <sup>24</sup>. En cambio, los formantes más altos no contribuyen a la calidad de la vocal formada, pero sí al timbre personal. Además, la trayectoria de los formantes nos entregan la combinación consonante-vocal, lo cual se puede ver en la figura 10.4 <sup>25</sup>.

Se realizó una tarea, en la cual se pidió analizar nuestro registro de voz. En primer lugar se analizaron las formantes de las vocales en cinco alturas de pitch distintas. Esto nos permitió observar que la estructura de formantes de cada vocal permanece prácticamente intacta. Luego, se nos pidió analizar la trayectoria de las formantes al realizar combinaciones consonante-vocal viendo una consonante de cada grupo anteriormente descritos. Esto nos mostró que la consonante inicial tiene ciertas características definidas las cuales cambian suavemente en el tiempo para dar paso a la vocal requerida.

<sup>22</sup>Fuente: Clases

<sup>23</sup>Fuente: Clases

<sup>24</sup>Fuente: Clases

<sup>25</sup>Fuente: Clases

Table 14.1 Consonants of English.

	Bilabial	Labiodental	Linguadental	Alveolar	Palatal	Velar	Glottal
Stops							
Voiceless	p			t		k	
Voiced	b			d		g	
Fricatives							
Voiceless	ɱ(wh <del>i</del> ch)	f	θ (th <del>i</del> ng)	s	ʃ(shoe)		
Voiced		v	ð (th <del>i</del> s)	z	ʒ(beige)		
Affricates							
Voiceless					tʃ(catch)		h
Voiced					dʒ(dodge)		
Nasals <sup>a</sup>	m			n		ŋ (sing)	
Liquids <sup>a</sup>					r, l		
Glides <sup>a</sup>	w					j (yes)	

<sup>a</sup>The nasals, liquids, and glides are voiced.

Figura 10.2: Tipos de consonantes

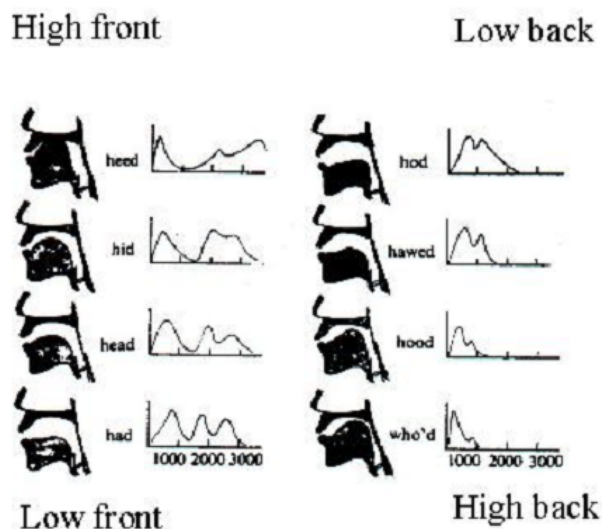


Figura 10.3: Formación de vocales según la posición de la lengua y tracto vocal

Finalmente, es interesante notar que al cantar se modifican las frecuencias de las formantes, las cuales pueden caer incluso en la categoría de otra vocal. Además, modificamos nuestra altura de voz (pitch). En general las frecuencias fundamentales del habla están alrededor de los 110 y 200 Hz (hombres - mujeres), pero al cantar se sube por lo menos a los 350 Hz. Esto se debe a que al cantar, se intenta mover la frecuencia de las formantes hacia la fundamental que se quiere alcanzar. Usualmente esto se hace modificando la apertura de la mandíbula, incrementando la amplitud de las formantes. Esto se puede apreciar en la figura 10.5 <sup>26</sup>.

<sup>26</sup>Fuente: Clases

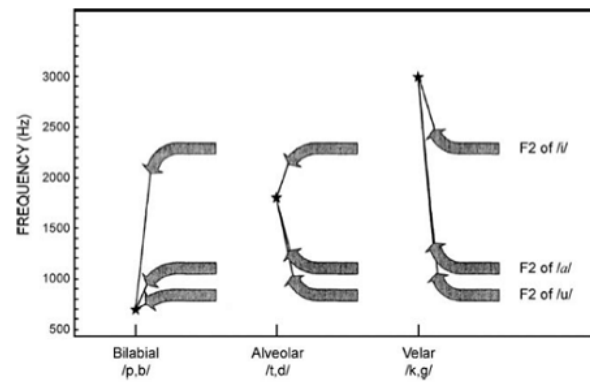


Figura 10.4: Transición de las formantes cuando se juntan vocales y consonantes

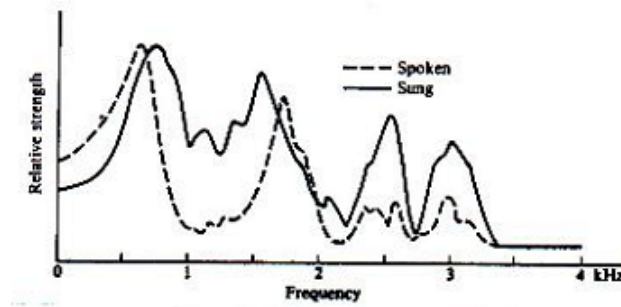


Figura 10.5: Cambio de frecuencias de las formantes al cantar

## 11. Música y cognición

La cognición se puede definir como la representación mental de un proceso. En otras palabras, es cómo procesamos la información a partir de la percepción, conocimiento adquirido y otros aspectos subjetivos<sup>27</sup>. En el aspecto del curso, podemos decir que la cognición a nivel musical es el cómo representamos la audición o la música en nuestra cabeza. De acá salen las preguntas: ¿Cómo queda la música en nuestra cabeza? ¿Qué cortezas se activan? ¿Cómo improvisamos o recordamos música?. En fin, es un tema bastante complejo que podría tener un curso completo destinado a entender los procesos cognitivos musicales, pero de todas formas intentaremos ver un par de conceptos:

1. Interacción con la música en presencia o ausencia de información “extra musical”
2. Relación visión - audición

### 11.1. Interacción e información extra musical

En primer lugar, es importante ver las razones por las cuales escuchamos música. En general, las pudimos encapsular en las siguientes:

- **Placer:** Va ligado con las sensaciones, la necesidad emocional de las personas. Es un concepto bastante amplio, pero que enfrasca la emocionalidad de las personas. Incluso cuando queremos escuchar canciones que nos dan tristeza, lo hacemos con ciertas ganas de llegar a dicho estado emocional, lo que lleva un tipo de placer. Esto lo hacemos a un nivel tanto de melodía como de la letra de las canciones, a las cuales les otorgamos cierto valor emocional.
- **Bloqueo:** Otra opción es para bloquear otros sentidos. Hay estudios que avalan este método. Al estar sometido a cierto estímulo auditivo que decidimos “desechar”, bloqueamos dicho sentido, pudiendo otorgar mayor atención a otros o evitando ser distraídos por otras señales auditivas.
- **Inquietud intelectual:** Una opción que se asocia más a la música docta o a la exploración de nuevos materiales. Acá queremos ver distintas obras que han sido creadas por una curiosidad que va más allá del placer. Este tipo de escucha sirve para alimentar el conocimiento musical o, desde un punto de vista menos abstracto, para ver distintas interpretaciones del mundo.

Analizando la escucha musical de la mayoría de la gente, escuchamos lo que se nos presenta, lo que otras personas nos muestran. Es más, escuchamos de la forma en la que nos enseñan a escuchar, lo que nos limita de gran forma a apreciar distintas formas de arte musical. Por ejemplo, personas sin acercamiento a música docta, no será capaz de apreciar lo realizado por el compositor, pero podrá disfrutar de la música comercial sin ningún problema. Pero al empezar a exponer a estas personas a otros tipos de arte musical, con el tiempo son capaces de detectar y aprender a escuchar de forma diferente. Sin embargo, también hay componentes naturales en nuestra escucha. Se ha demostrado que ciertos acordes son preferidos por recién nacidos, incluso mientras están en gestación, lo que indica cierta predilección instintiva por cierto tipo de música.

*Esto lo encuentro sumamente interesante, ya que reafirma que existen conceptos más que estudiados para la creación de “música envasada” o popular. Al estudiar las respuestas inconscientes del humano frente a distintos tipos de acordes, progresiones u otros, es fácil predecir la respuesta de la gente frente a la creación musical. También es posible condicionar a las personas de esta forma, lo que se ve aplicado en locales comerciales diariamente.*

Para analizar este concepto se realizó un ejemplo en clases. En un comienzo se nos hizo escuchar una obra sin ninguna información. Luego, se nos dio el nombre de la obra y la escuchamos nuevamente. Finalmente, se nos entregó toda la información referente a la obra, el autor y la nota de programa, para escucharla por una última vez. En cada una de estas veces teníamos que fijarnos en el paisaje sonoro, las posibles fuentes y las sensaciones que nos daba dicha obra.

#### 1. Sin información adicional

Escuché bombos y vientos creando tensión, suspenso. También sonaba algo parecido a puertas crujiendo o parlantes malos y ramas quebrándose. Había una sensación extraña en el ambiente, como llegando al límite de algo. Luego, empezó la invasión de algo. Finalmente, se pierden los bajos y la obra tiene menos sentido, puesto que antes se apreciaba una estructura rítmica.

#### 2. Sólo título - Nombre de la obra: Estrellas compactas

Esta vez se me vinieron a la mente fractales tridimensionales. Como si estuviesen ocurriendo varios big bangs

---

<sup>27</sup>Fuente: Wikipedia

simultáneos y sucesivos. Las formas se comprimían y explotaban para aumentar su tamaño. Además, el crecimiento era ramificado. Estaban hechas de algún tipo de material que se podía quebrar, pero también endurecer en cada posición nueva. Al final, cuando se pierden los bajos, suena como campanas. Es como si se estuviesen descascarando estas moles y caen pequeños diamantes de estrellas.

3. **Información completa** - Autor: Federuco Schumacher, fecha: 2003, en la nota de programa cuenta que son los diferentes sonidos que produce el tubo de su aspiradora más un homenaje a los variados ruidos de su apartamento que intentó grabar con una máquina, pero que no pudo grabarlos completamente.

Esta vez se fueron la mayoría de los aspectos abstractos anteriores. Me imaginaba el tubo de PVC de la aspiradora con cosas atascadas que les costaba pasar. Sonidos rugosos y rebotes a lo largo del tubo, como si estuviese absorbiendo una pelota. Ruidos de latas de bebida abriéndose, puertas y suelos crujiendo. Finalmente, campanillas de ventanal.

Esto nos muestra como influye la información extra-musical en la cognición. A mayor cantidad de información que poseemos, menor espacio tenemos para imaginar. *Nuestro cerebro intenta dar sentido a las cosas que escuchamos, por lo que no podemos ir en contra de la información ya adquirida. Provocaríamos una especie de colapso al intentar ir en contra de dicha información o incluso al intentar ignorarla, ya que incluso inconscientemente sabemos que va a hacer menos sentido.*

## 12. Percepción rítmica

(Clase basada en el trabajo de Juan Pablo Cáceres, “Performance Issues and Strategies”)

Al analizar las diferentes problemáticas a la hora de querer tocar música con otras personas a través de internet, llegamos al problema fundamental del retraso de las señales. Este problema es natural y también se puede dar en vivo, en el caso de grandes orquestas. Se podría llegar a pensar que lo óptimo sería tener una red completamente libre de retraso, pero este caso no es el óptimo. Resultados experimentales mostraron que un buen sistema de música por internet incorpora un leve retraso de aproximadamente 15 ms. Esto se debe a la diferencia de las velocidades de la luz y el sonido, puesto que estamos acostumbrados a compensar dicho retraso en la vida real. Hay tres posibles efectos de descompensación temporal: lead (cuando la señal auditiva tiene muy poco retraso), synch (cuando está bien calibrada) y lag (cuando el retraso es mayor al ideal).

Otra forma de analizar el retraso es en orquestas. Se da un retraso natural, debido a las distintas posiciones de los músicos y al tiempo de respuesta de cada uno, lo que hace que la obra no esté 100 % coordinada. Sin embargo, al ser presentados frente a una orquesta coordinada perfectamente, con perfección rítmica, (logrado de forma artificial por medio de computadores) se tiene la sensación de música no natural. Logramos notar que es producida por máquinas al quitarle dichas imperfecciones sutiles a nivel rítmico.

*Encuentro muy interesante que seamos capaces de discernir si algo suena “perfecto” o no, incluso a nivel rítmico. Además, como se ha visto durante todo el curso, el cerebro intenta, por todos los medios posibles, hacer sentido de las señales que nos llegan constantemente. Así, al ser presentados una obra con perfección rítmica, somos capaces de discernir la no-naturalidad de la fuente, asumiendo la imposibilidad de tal nivel de perfección en la naturaleza.*