

Procesamiento digital de señales de audio

Percepción auditiva

Instituto de Ingeniería Eléctrica, Facultad de Ingeniería
Universidad de la República, Uruguay

Grupo de Procesamiento de Audio

① Fisiología del sistema auditivo

② Psicoacústica

③ Modelado computacional

Fisiología del sistema auditivo

Etapas de procesamiento

- captación y procesamiento mecánico de ondas sonoras
- conversión de la señal mecánica en impulsos nerviosos
- procesamiento neural, producción de la sensación sonora

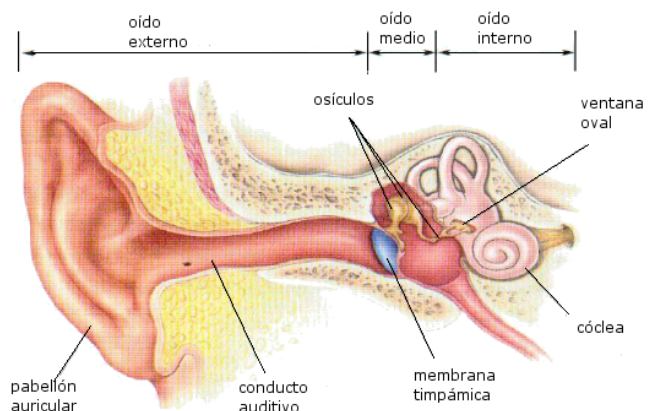
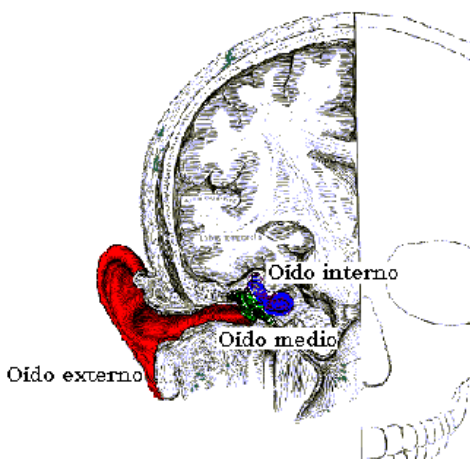
Sistema auditivo

- **periférico** de ondas mecánicas a señales electroquímicas
- **central** se transforma la información en sensaciones auditivas

Fisiología del sistema auditivo

Sistema auditivo periférico

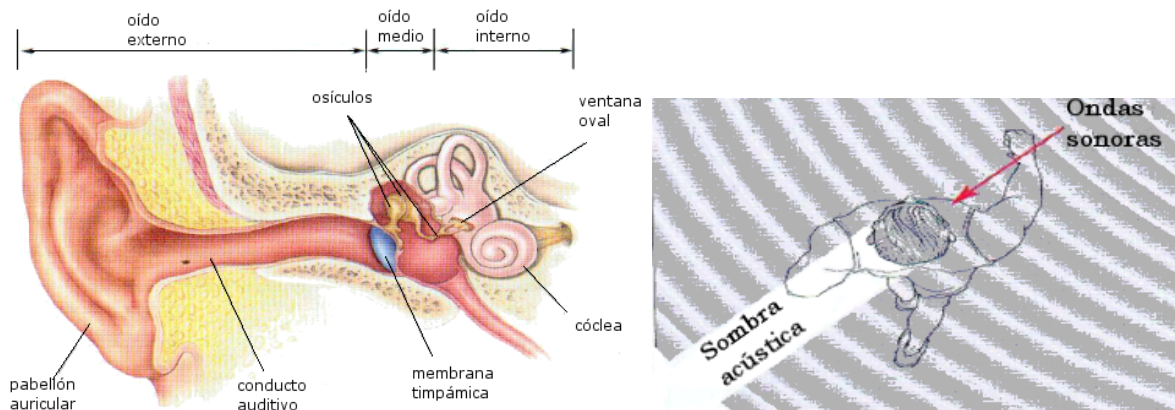
- se divide en: **oído externo**, **oído medio**, **oído interno**
- traduce variaciones de presión a disparos en nervio auditivo



Sistema auditivo periférico

Oído externo

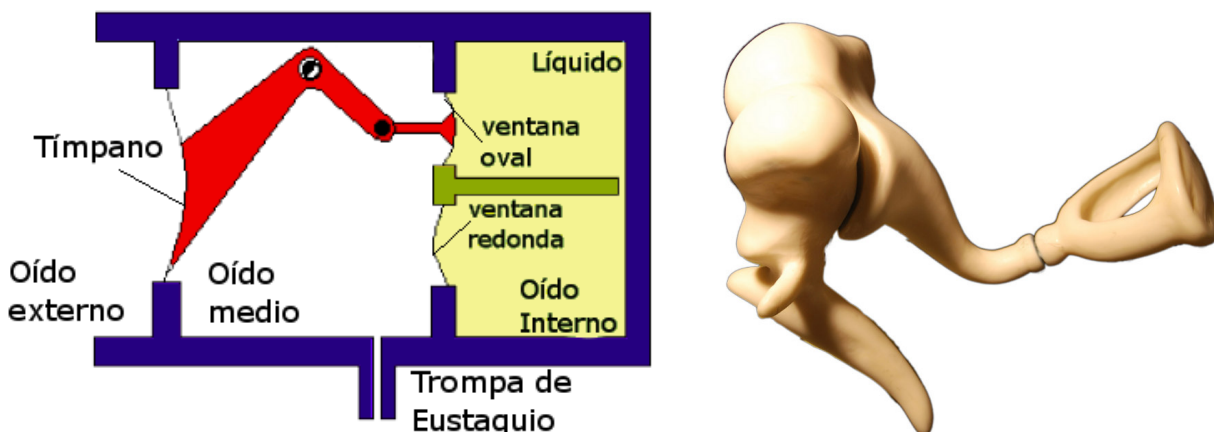
- compuesto por: **pabellón auricular** y **conducto auditivo**
- pabellón auricular:
 - ▶ capta y concentra las ondas acústicas (resonancia en 4-5 kHz)
 - ▶ produce sombra acústica según dirección del sonido
- conducto auditivo:
 - ▶ protege el oído medio e interno (resonancia en 3-4 kHz)



Sistema auditivo periférico

Oído medio

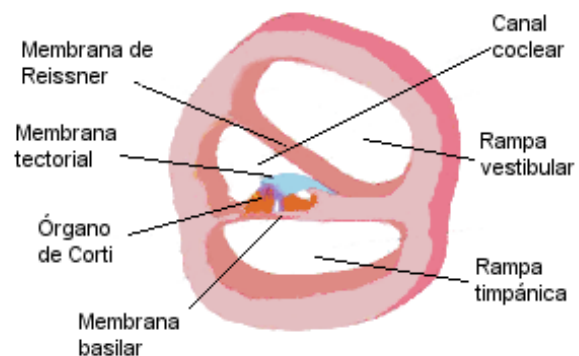
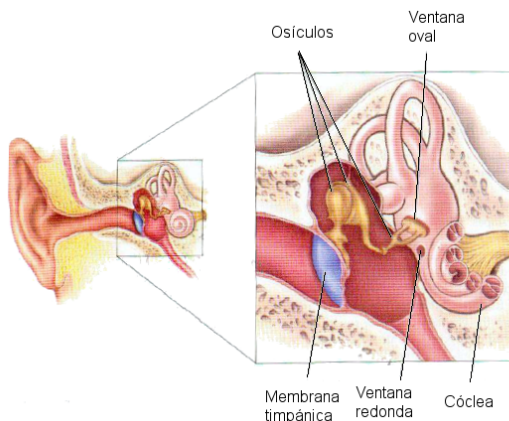
- ubicado en **caja timpánica**: cavidad en el hueso temporal
 - ▶ **tímpano**: membrana que separa el oído externo del medio
 - ▶ **osículos**: cadena de huesecillos (martillo, yunque, estribo)
- reflejo timpánico: contracción muscular frente a sonidos intensos
- adaptación de impedancias: palanca y relación de áreas



Sistema auditivo periférico

Oído interno

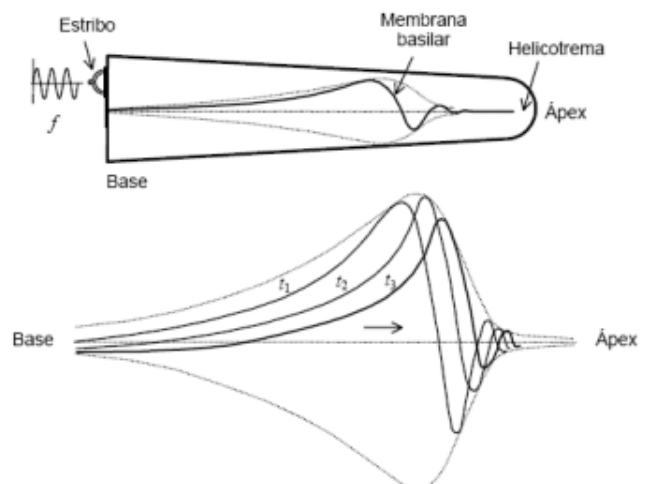
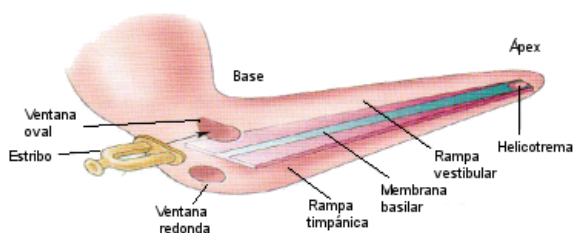
- compuesto por: **canales semicirculares** y **cóclea**
- **cóclea**: conducto rígido en espiral (arrollado 2.5 veces)
 - ▶ conectado al oído medio por ventanas oval y redonda
 - ▶ dividido por membranas basilar y de Reissner
- **órgano de Corti**: sobre la membrana basilar, recibe su movimiento
 - ▶ células ciliadas que generan impulsos nerviosos hacia el nervio auditivo



Sistema auditivo periférico

Oído interno

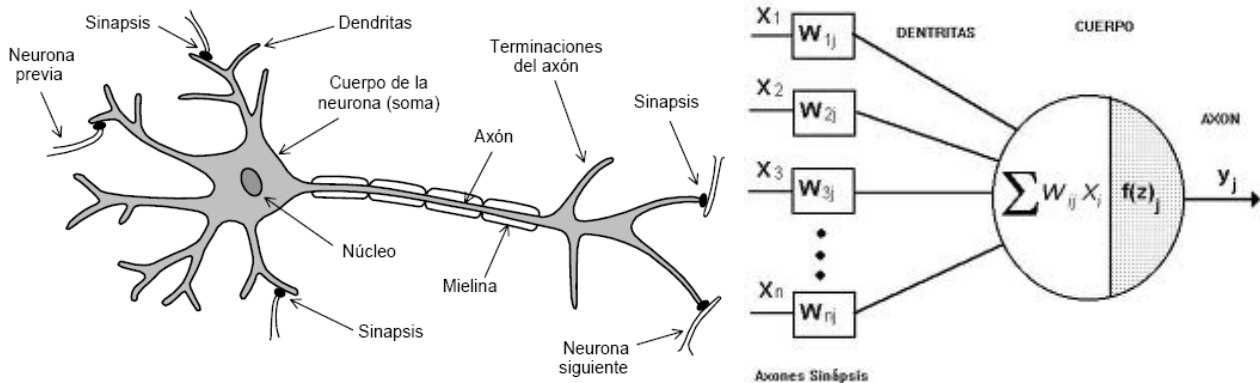
- **onda viajera**: se propaga en la membrana basilar
- la membrana basilar varía en grosor y rigidez
- punto de deformación máxima depende de la frecuencia (log)
- una señal de espectro complejo se manifiesta en distintas posiciones



Sistema auditivo central

Nervio auditivo

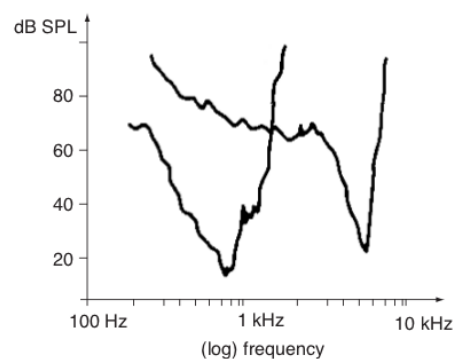
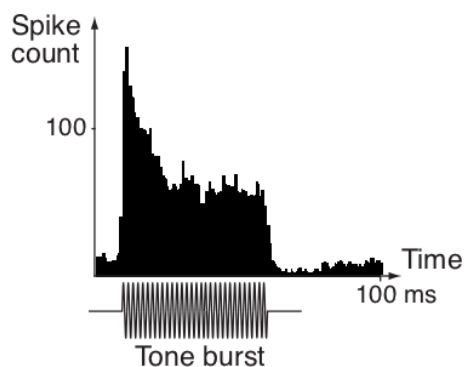
- potencial de acción de células ciliadas viaja hacia nervio auditivo
- se inicia por movimiento de cilias en una única dirección (rectificación)
- ciclo de neuronas: descarga - inactividad (período refractario)
- se genera un tren de impulsos “periódico” en el nervio auditivo
- intervalo entre pulsos aprox. múltiplo del período de la señal



Sistema auditivo central

Nervio auditivo

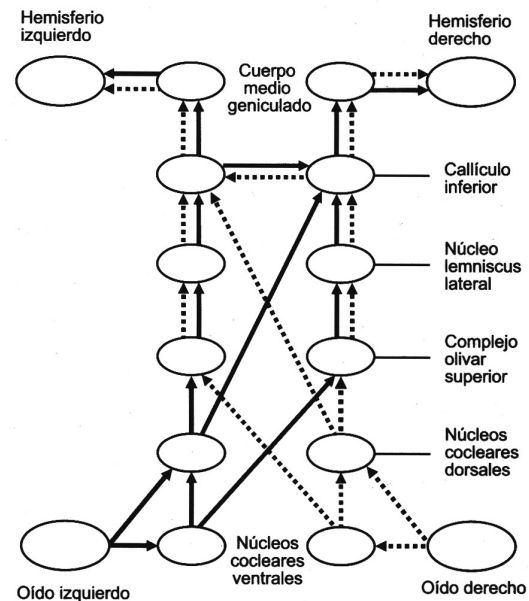
- fibras del nervio auditivo responden espontáneamente sin estímulo
- la tasa de disparo está relacionada con el nivel del estímulo (sigmoide)
- adaptación: inicio marcado y decaimiento en estado estacionario
- selectividad en frecuencia, se mantiene **tonotopía** de membrana basilar



Sistema auditivo central

Nervio auditivo a corteza cerebral

- vía auditiva
 - ▶ **aférente**: hacia la corteza
 - ▶ **eférente**: hacia el oído
explica fenómenos activos,
e.g. emisiones otoacústicas
- neuro-estructuras intermedias
- neuronas especializadas:
periodicidad, intensidad, amplitud,
modulación de amplitud y frecuencia,
diferencias interaurales de tiempo e
intensidad, etc.
- hemisferios
 - ▶ izquierdo: voz, lenguaje
 - ▶ derecho: música



Psicoacústica

Motivación

- comprensión limitada de estructuras cerebrales superiores
- se recurre a descripción **psicoacústica** de fenómenos perceptuales

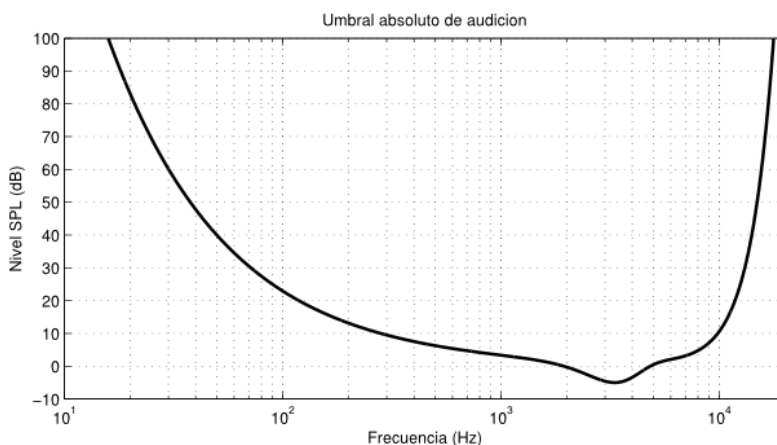
Psicoacústica

- rama de la **psicofísica**, estudia relación entre:
 - ▶ características físicas de un estímulo (sonoro)
 - ▶ respuesta psicológica que provoca en un sujeto
- análisis estadístico de experimentos que buscan medir la respuesta subjetiva de sujetos a estímulos de propiedades físicas cuantificadas

Psicoacústica

Objetivos

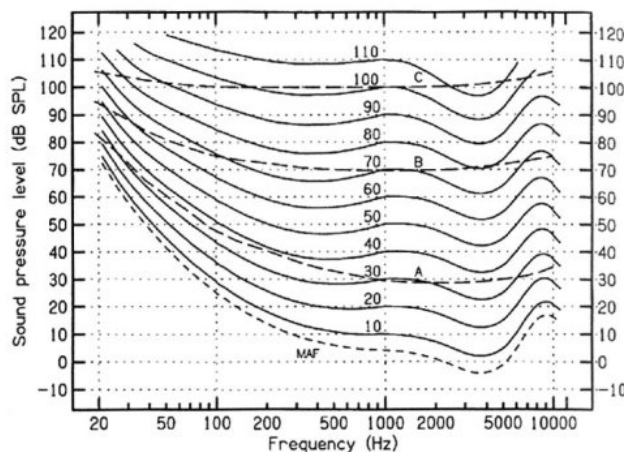
- cuantificar relación entre estímulos físicos y respuesta perceptiva
- establecer umbrales absolutos de sensación de cada parámetro
 - ▶ valor de parámetro físico en que la sensación comienza o desaparece
- establecer umbrales diferenciales de percepción
 - ▶ mínima diferencia perceptible, mínima variación perceptible
- la respuesta de sensación a varios estímulos simultáneos



Psicoacústica

Sonoridad (loudness)

- relacionado en principio con la intensidad del sonido (SPL)
$$SPL(dB) = 20 \log(p/p_{ref}) \quad p(\text{pa}) \quad p_{ref} = 20\mu\text{pa}$$
- depende además de: frecuencia, ancho banda, espectro, duración
- **curvas isófonas** igual sonoridad para distinta frecuencia



sonoridad: ejemplo 1

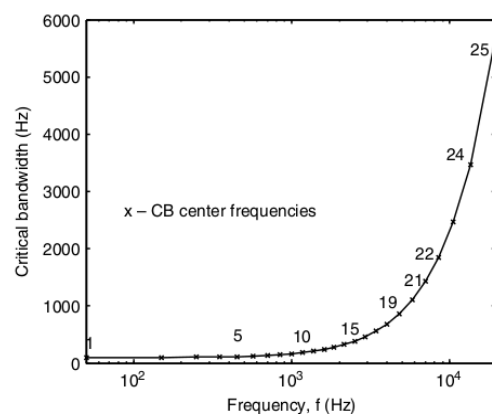
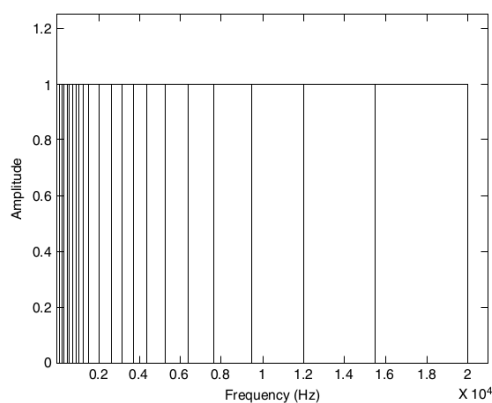


ejemplo 2

Psicoacústica

Sonoridad y banda crítica

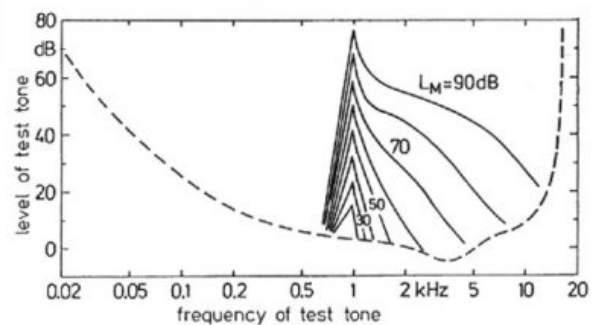
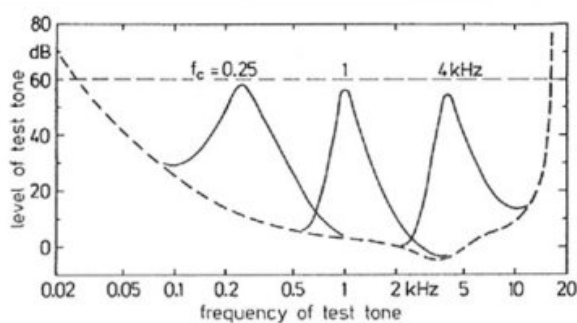
- experimento:
 - ▶ ruido de banda angosta, se incrementa su ancho de banda
 - ▶ sonoridad permanece incambiada hasta **ancho de banda crítico**
 - ▶ luego cualquier aumento de ancho de banda incrementa sonoridad
- **ancho de banda crítico**
 - ▶ constante (100 Hz) hasta 500 Hz y luego un 20% de frec. central
 - ▶ aproximación: escala **Bark**, $z_b(f) = 600 * \operatorname{asinh}(f/600)$



Psicoacústica

Enmascaramiento

- un sonido hace inaudible a otro (sonido enmascarante y prueba)
- depende de diferencia de: intensidad, frecuencia y tiempo
- **umbral de enmascaramiento**: corrimiento del umbral de audibilidad
- patrón no simétrico, se extiende más hacia los agudos
- además de **simultáneo**, **pre** y **pos** enmascaramiento



Psicoacústica

Altura (pitch)

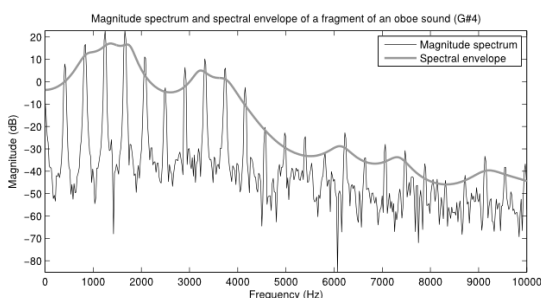
- respuesta subjetiva a los patrones periódicos del estímulo sonoro
 - sonidos armónicos:** componentes se funden en una sensación de altura definida (correspondiente a f_0 , aún si no está presente)
 - sonidos inarmónicos:** se pierde sensación de altura, aunque depende del grado de inarmonicidad (e.g. piano vs campana)
 - bandas de ruido:** no hay sensación definida de altura
- teorías de percepción de altura:
 - ▶ **teoría del lugar:** reconocimiento de patrón específico de componentes distribuidos en diferentes partes de la cóclea
 - ▶ **teoría temporal:** sincronismo temporal que permite derivar periodicidad del estímulo por correlación
 - ▶ probablemente se utilice una combinación de información

🔊 ejemplos 🔊 paradoja perceptiva: glissando de Shepard-Risset

Psicoacústica

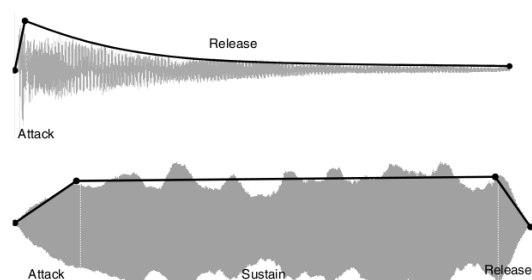
Timbre

- percepción que integra diversos grados de generalidad (e.g. material, familia de instrumentos, instrumento, registro)
- información de la fuente: material, forma y tamaño, modo excitación
- principales aspectos que lo determinan:
 - ▶ distribución de energía en el espectro (resonador y sus formantes) y su evolución temporal
 - ▶ comportamiento temporal de la envolvente de intensidad
 - ▶ transitorios del sonido que aparecen en el ataque



magnitud del espectro y envolvente espectral

🔊 sin resonador 🔊 con resonador



envolvente de amplitud para nota de guitarra y violín

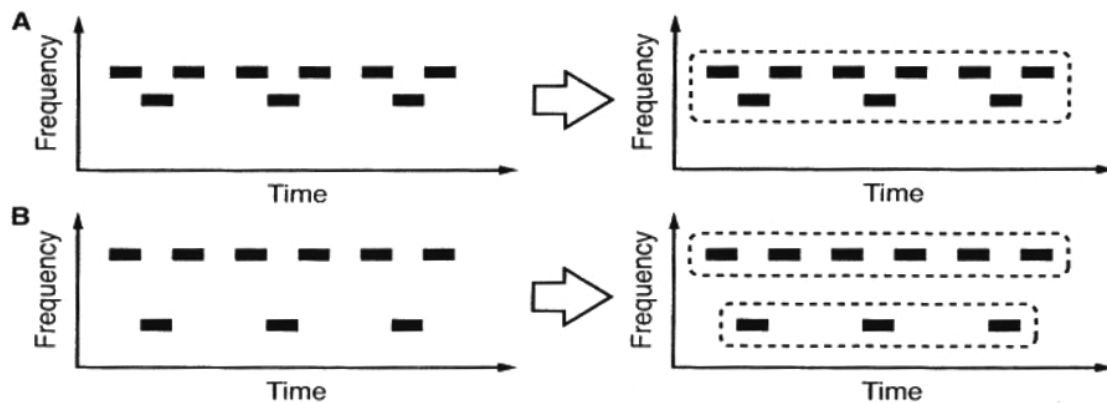
🔊 bell 🔊 tam-tam 🔊 sin y 🔊 con ataque

Psicoacústica

Análisis del panorama auditivo

Auditory Scene Analysis (ASA)

- cocktail party problem: agrupamiento y separación de sonidos
- principios de agrupamiento: primitivos y basado en esquemas
- principales principios primitivos:
 - ▶ proximidad en frecuencia y tiempo, periodicidad (armonicidad), transición continua o suave, inicio y fin común, modulación de amplitud y frecuencia, ritmo, localización espacial común



Modelado computacional

Aplicación de modelos computacionales de percepción auditiva

- sistema auditivo periférico motiva representaciones intermedias
- psicoacústica y análisis de panorama auditivo brindan principios sobre funcionamiento de la audición que pueden ser explotados
- algoritmos inspirados en modelos computacionales de audición
 - ▶ algunas aplicaciones se han visto muy beneficiadas (e.g. compresión)
 - ▶ algunos principios brindan ventajas concretas (e.g. log-freq-filterbank)
 - ▶ en general costo computacional alto y menor interpretabilidad



Esquema de un sistema típico de CASA (Computational Auditory Scene Analysis)

Modelado computacional

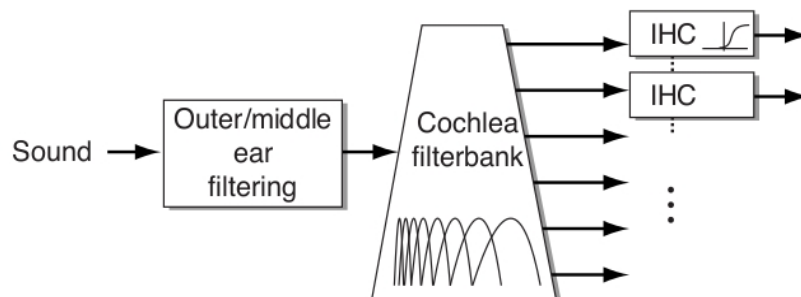
Modelado del sistema auditivo periférico

1 banco de filtros auditivos

- ▶ modela selectividad en frecuencia de la cóclea
- ▶ frecuencia central distribuída uniformemente en escala logarítmica
- ▶ respuesta simula el movimiento de la membrana basilar

2 simulación de células ciliadas Inner Hair Cell (IHC)

- ▶ compresión de nivel, rectificación de media onda, filtrado pasabajos



Esquema de un modelo típico de sistema auditivo periférico

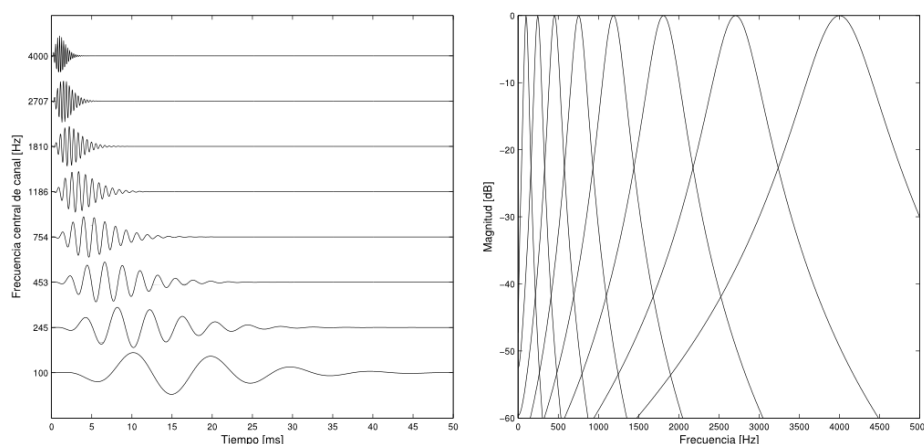
Modelado computacional

Banco de filtros auditivos

gammatone aproximación de respuesta de nervio auditivo

$$g_c(t) = (2\pi b)^n / \Gamma(n) t^{n-1} e^{-2\pi b t} \cos(2\pi f_c t + \phi)$$

- ▶ ancho de banda y distribución de f_c dado por bandas críticas (Equivalent Rectangular Bandwidth, ERB)

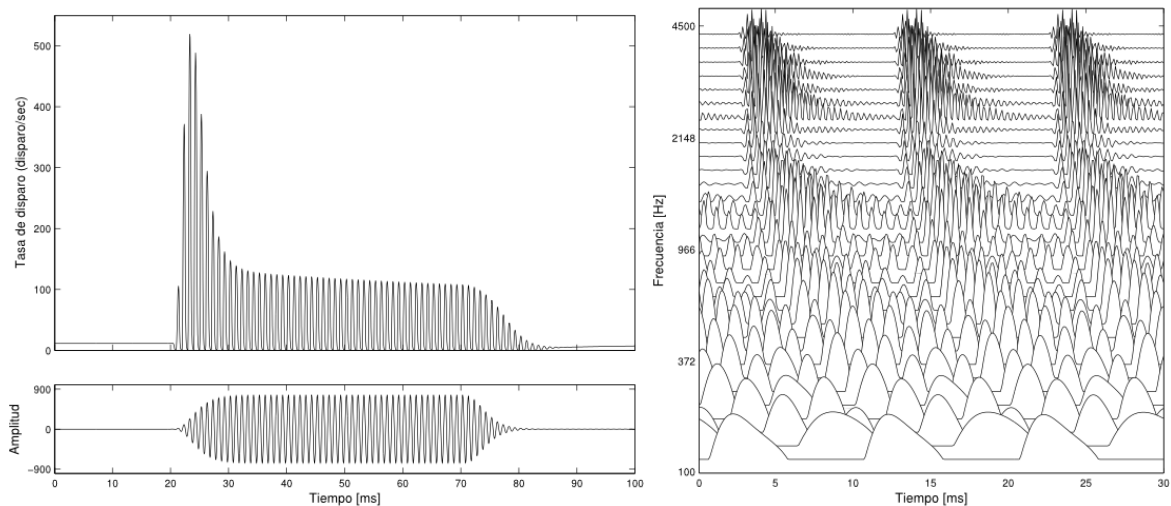


Modelado computacional

Modelo de célula ciliada

simula la tasa de disparos en el nervio auditivo

- ▶ adaptación a estado estacionario resalta inicio de eventos



Modelado computacional

Modelado del sistema auditivo central

- 1 análisis de periodicidad en cada canal auditivo (correlograma)
- 2 integración de información entre canales (e.g. summary, cross-correlogram)

