Percepção Auditiva

PTC2547 – Princípios de Televisão Digital

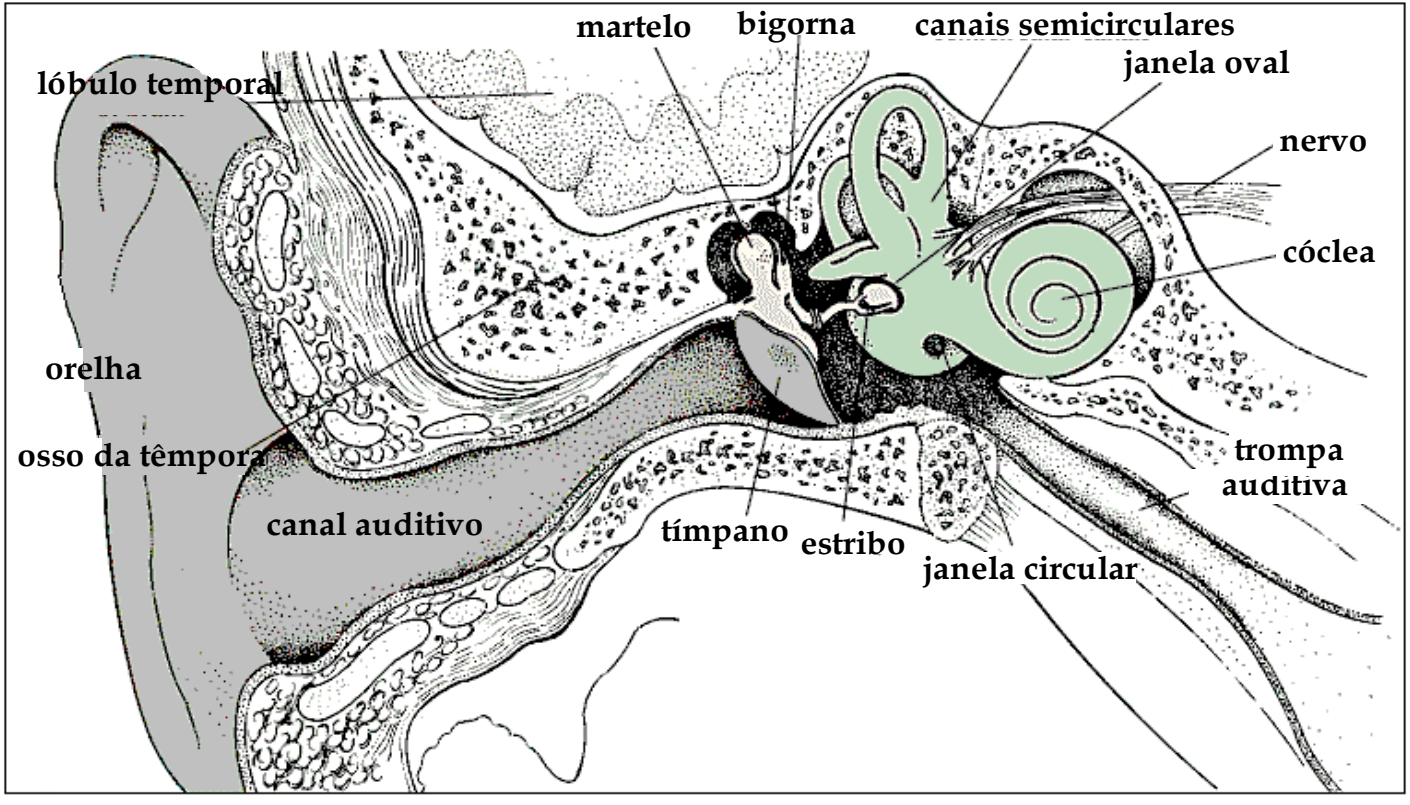
Guido Stolfi – 10 / 2016

Tópicos Abordados

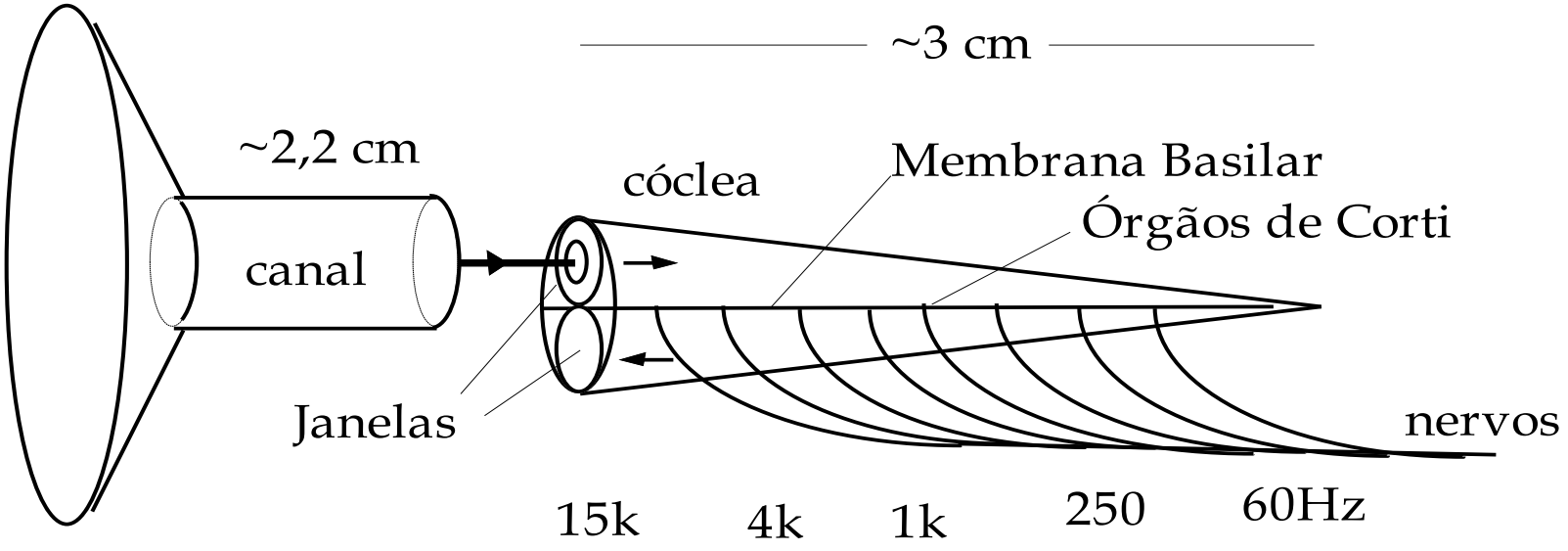
- Fisiologia do Ouvido Humano
- Acústica Física
- Mascaramento
- Voz Humana e Vocoder
- Modelamento de Sistemas Dinâmicos
- Questões Polêmicas

Fisiologia do Ouvido Humano

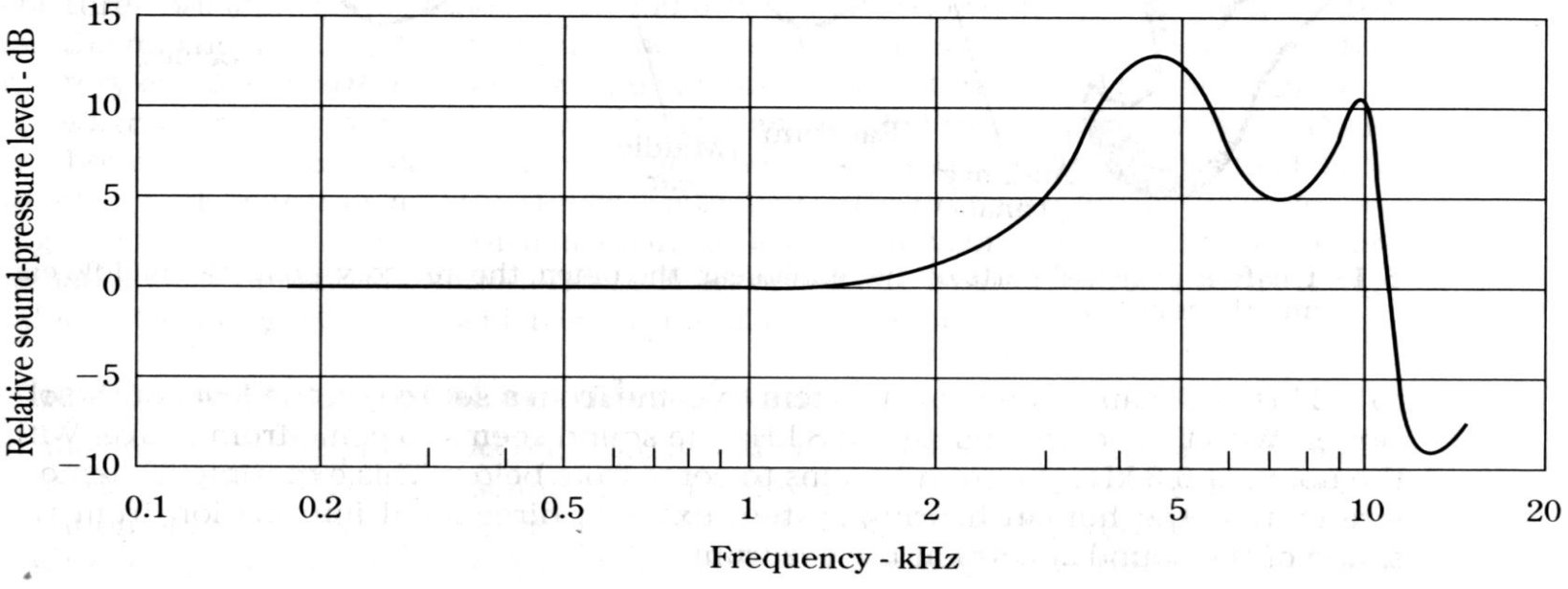
Estrutura do Ouvido Humano



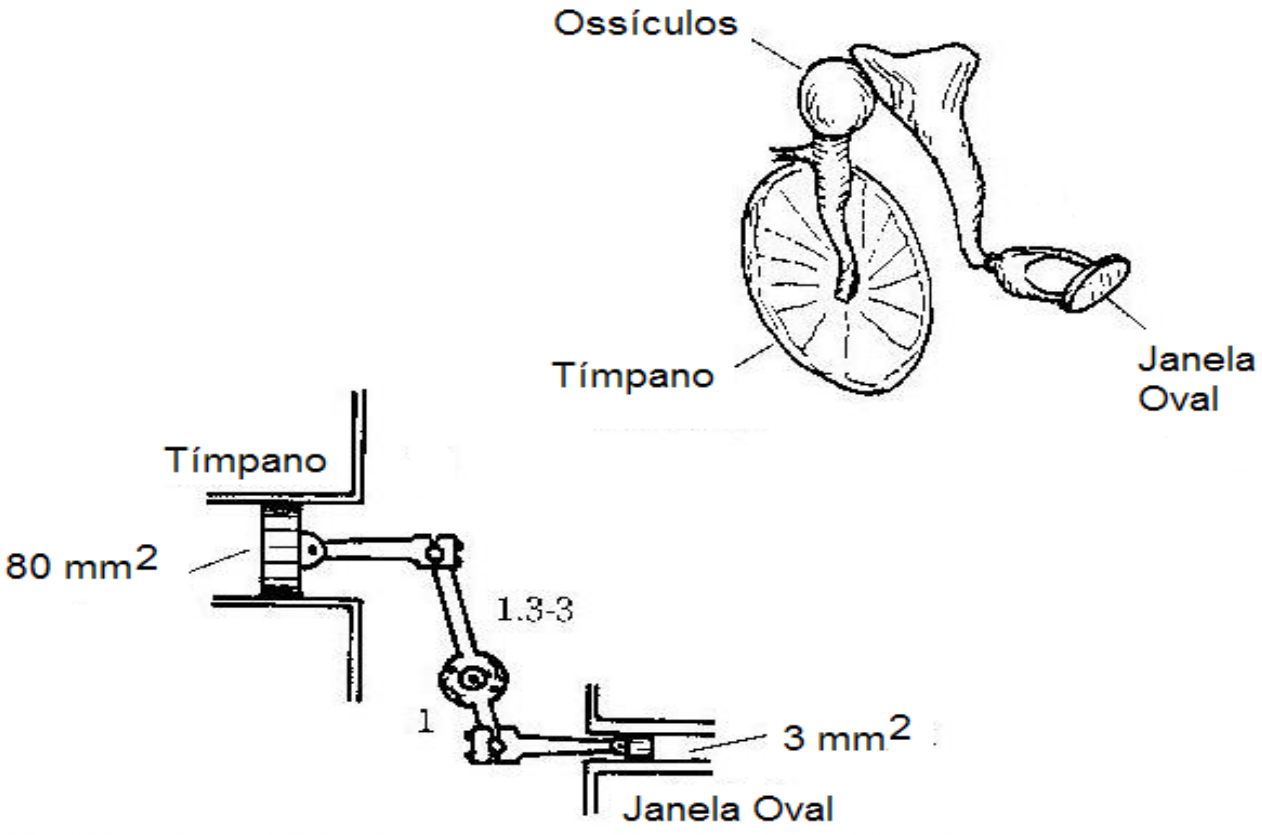
Representação Esquemática do Ouvido



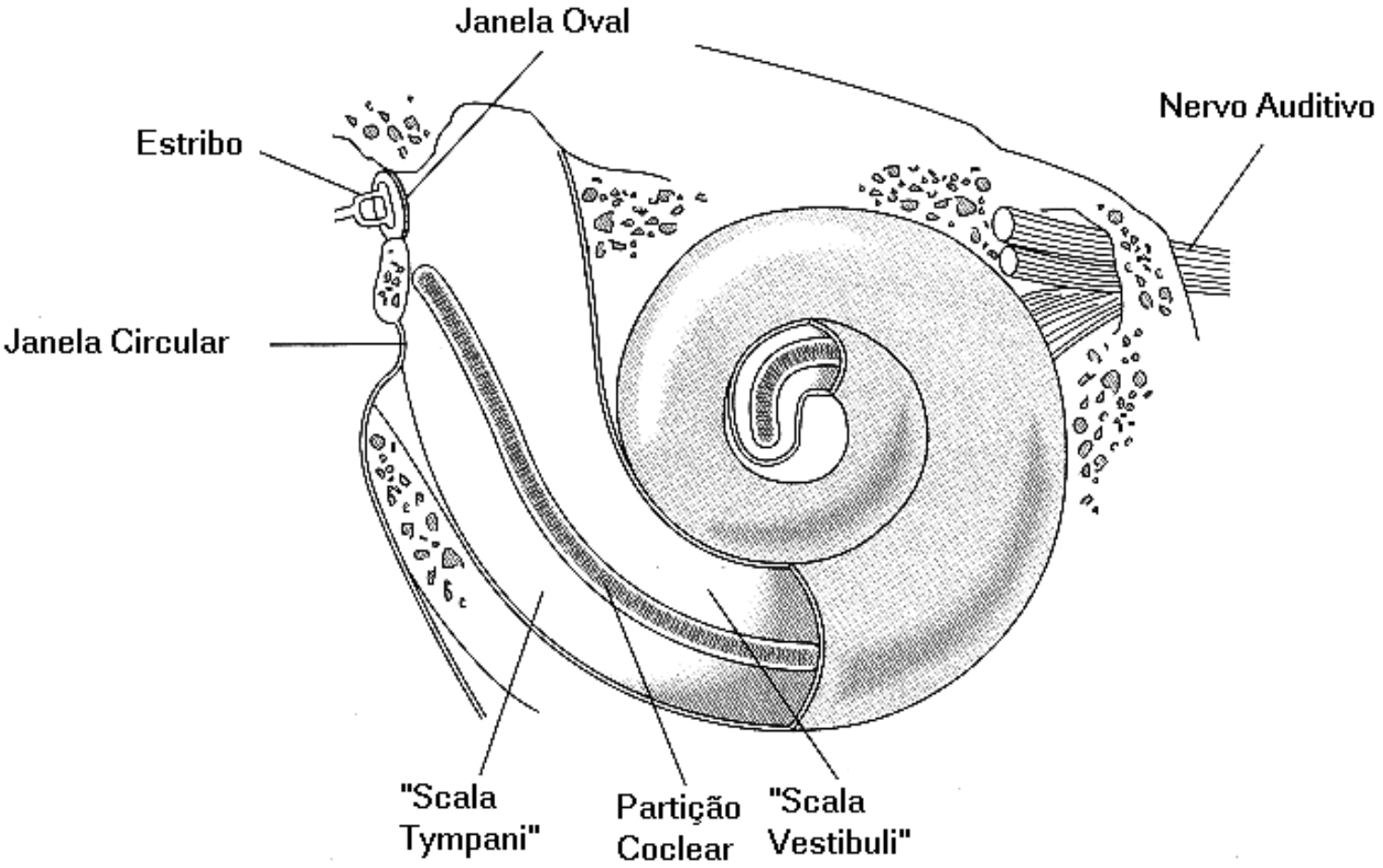
Resposta em Frequência do Canal Auditivo



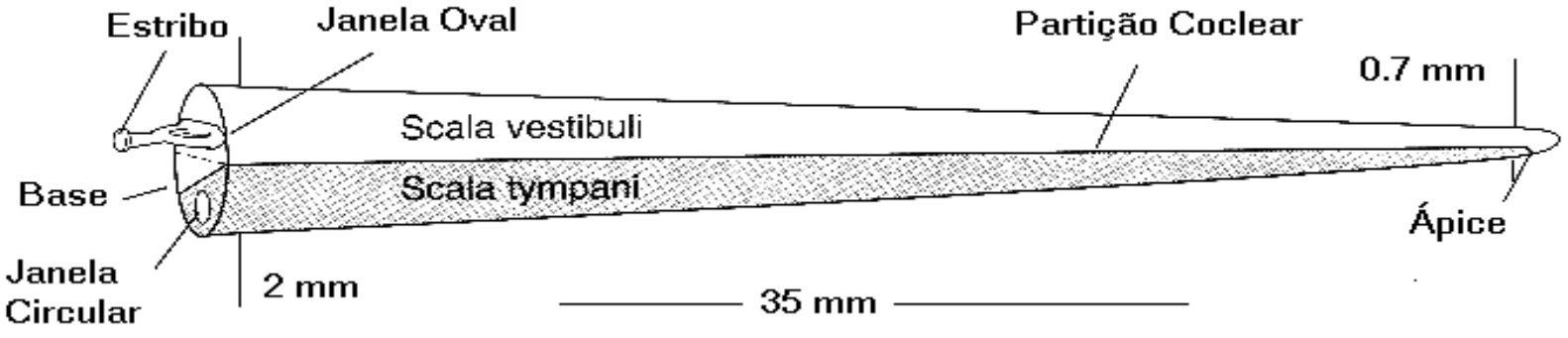
Ouvido Médio



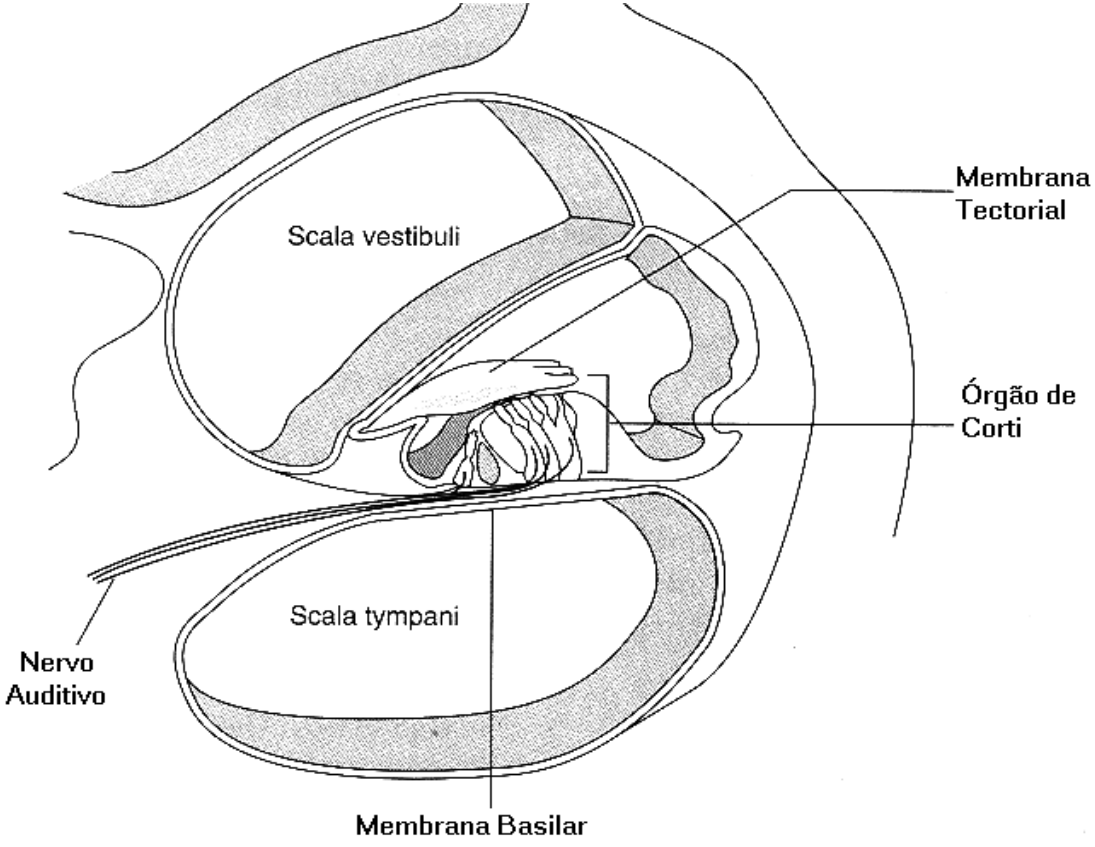
Estrutura da Cóclea



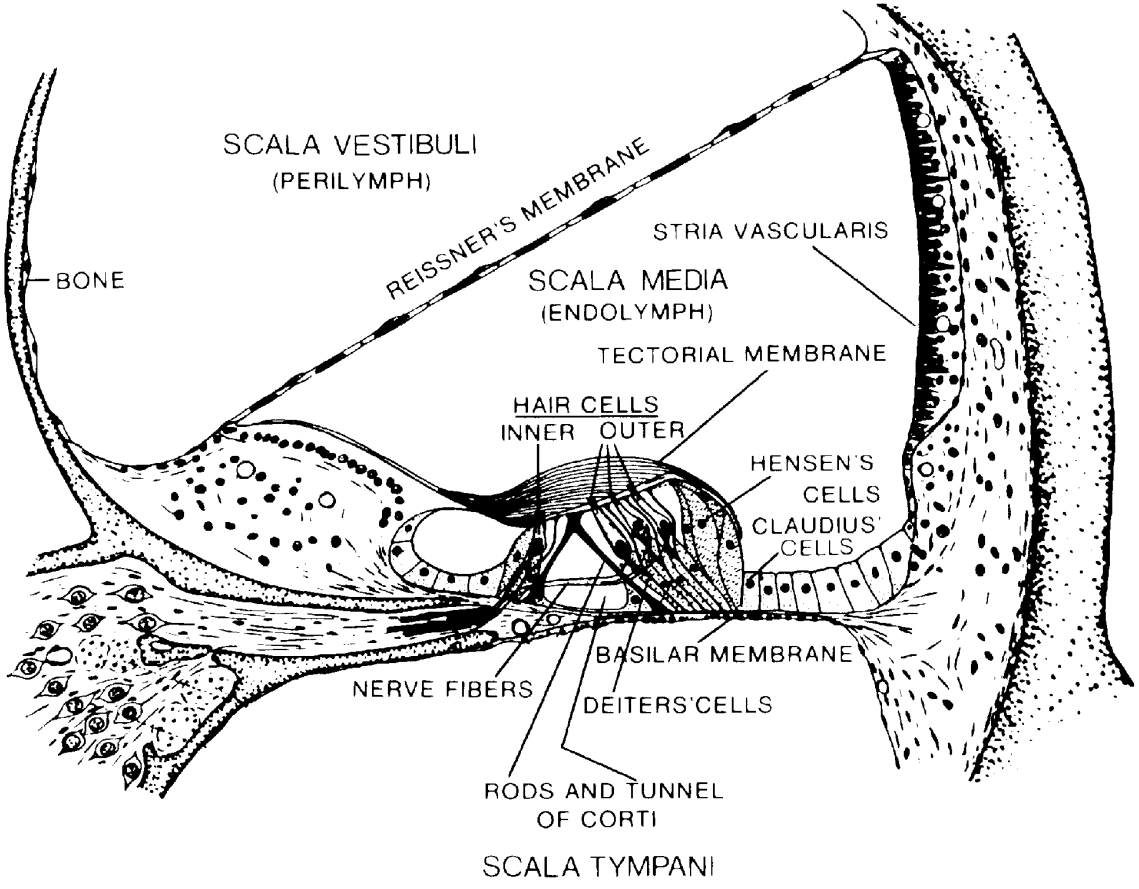
A Cóclea Distendida



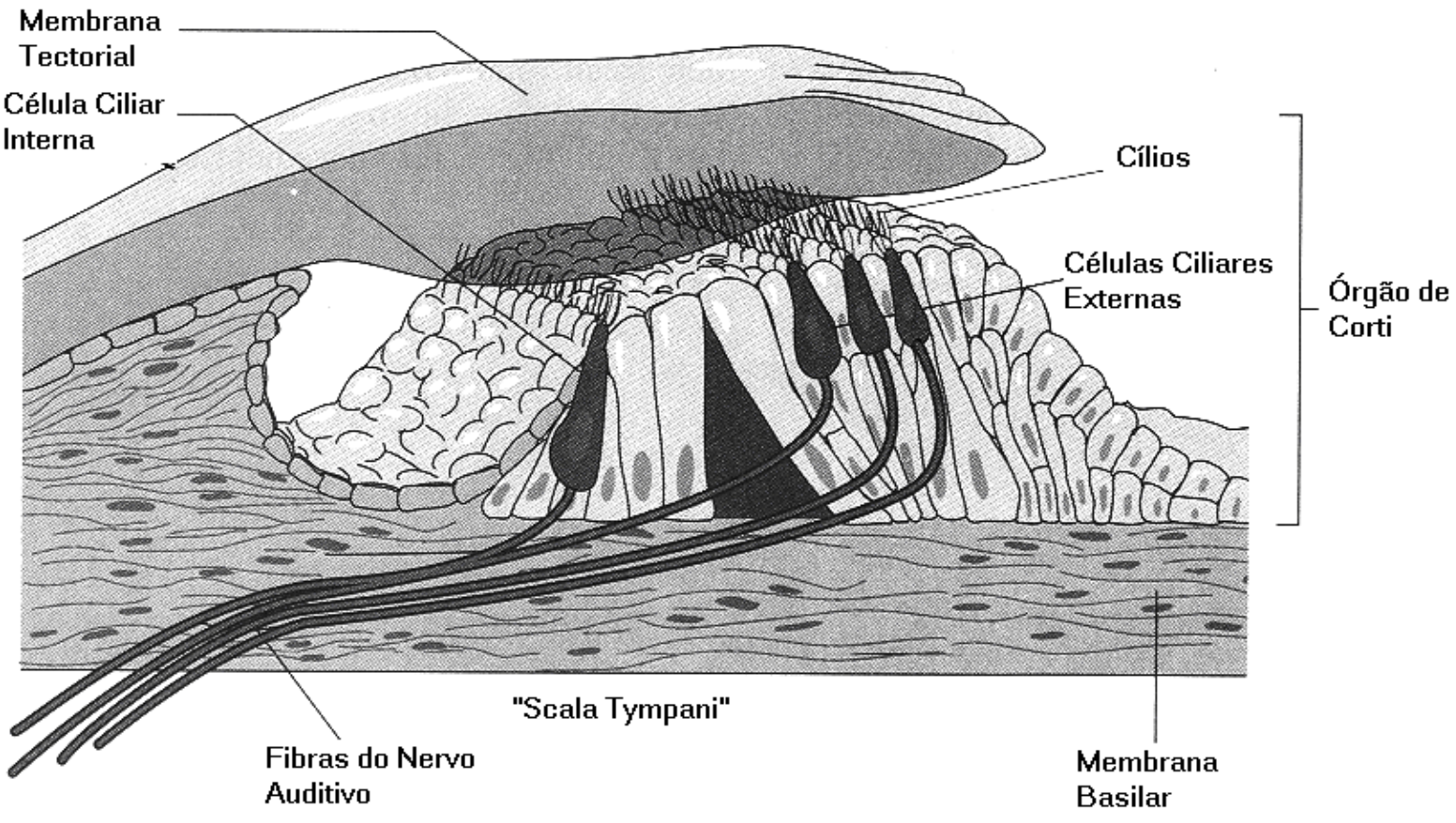
Secção da Cóclea



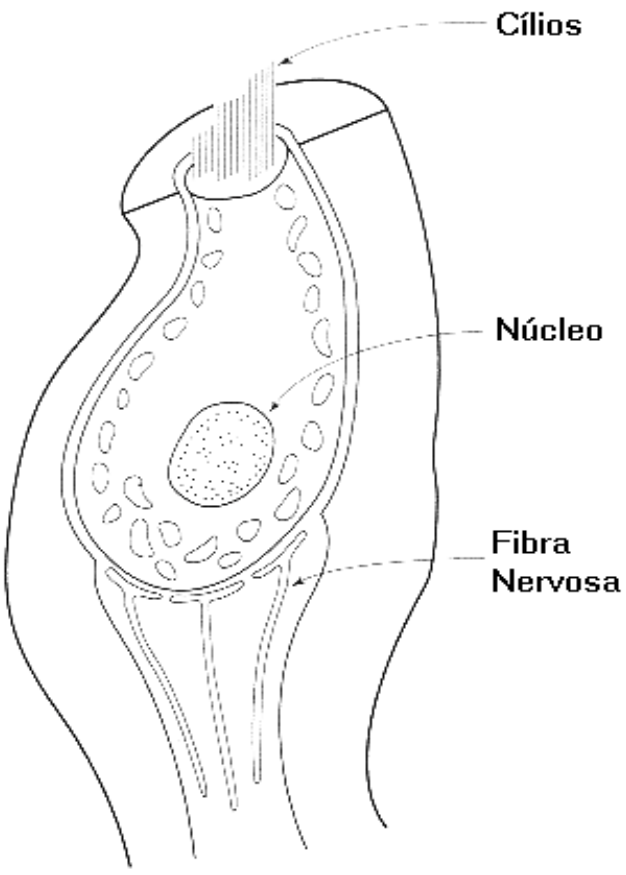
Secção da Cóclea



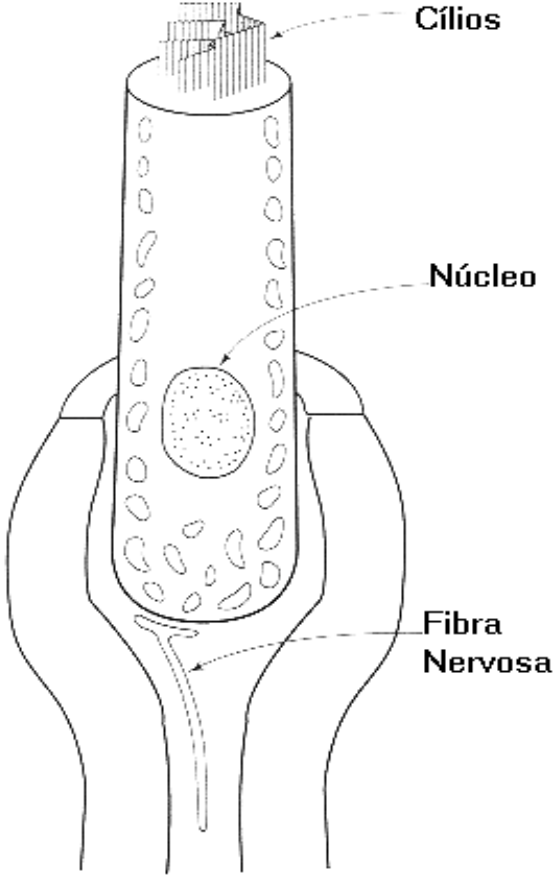
Detalhe dos Órgãos de Corti



Células de Corti

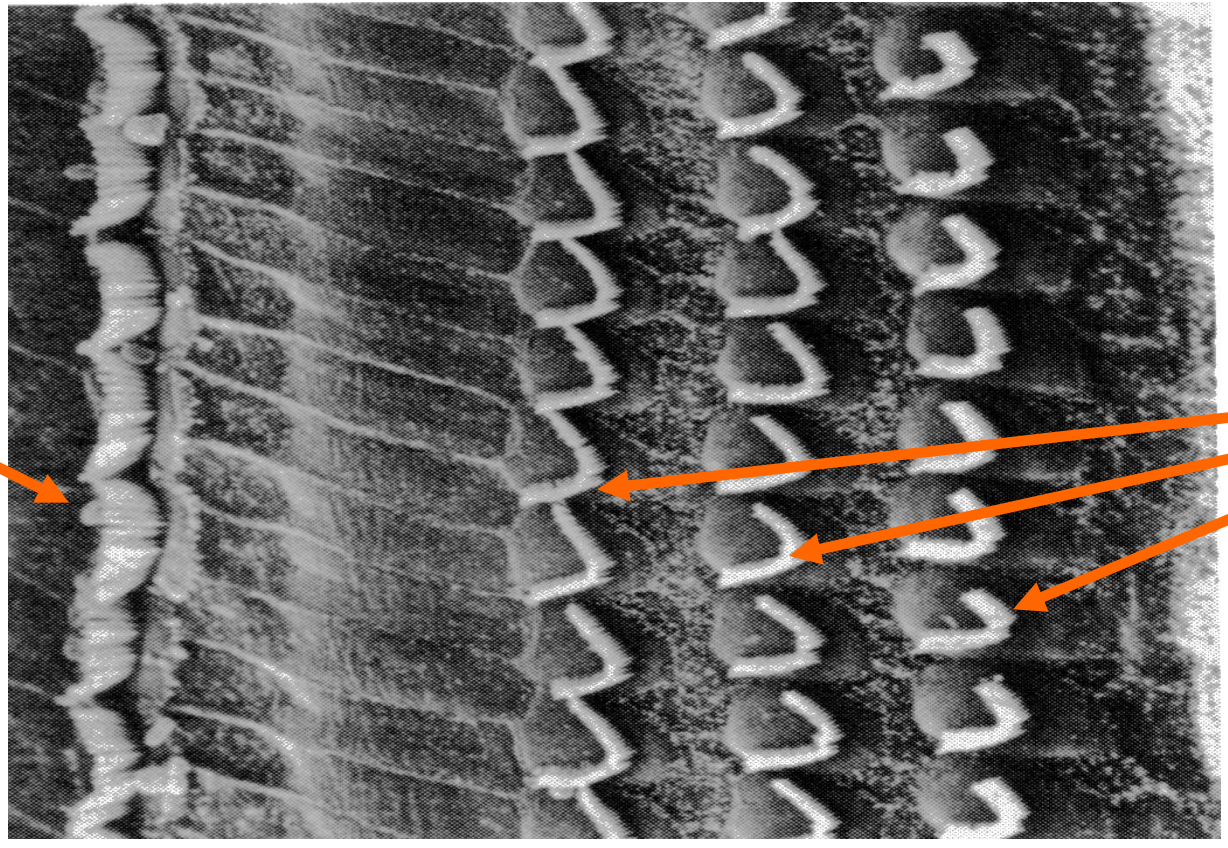


Célula Ciliar Interna

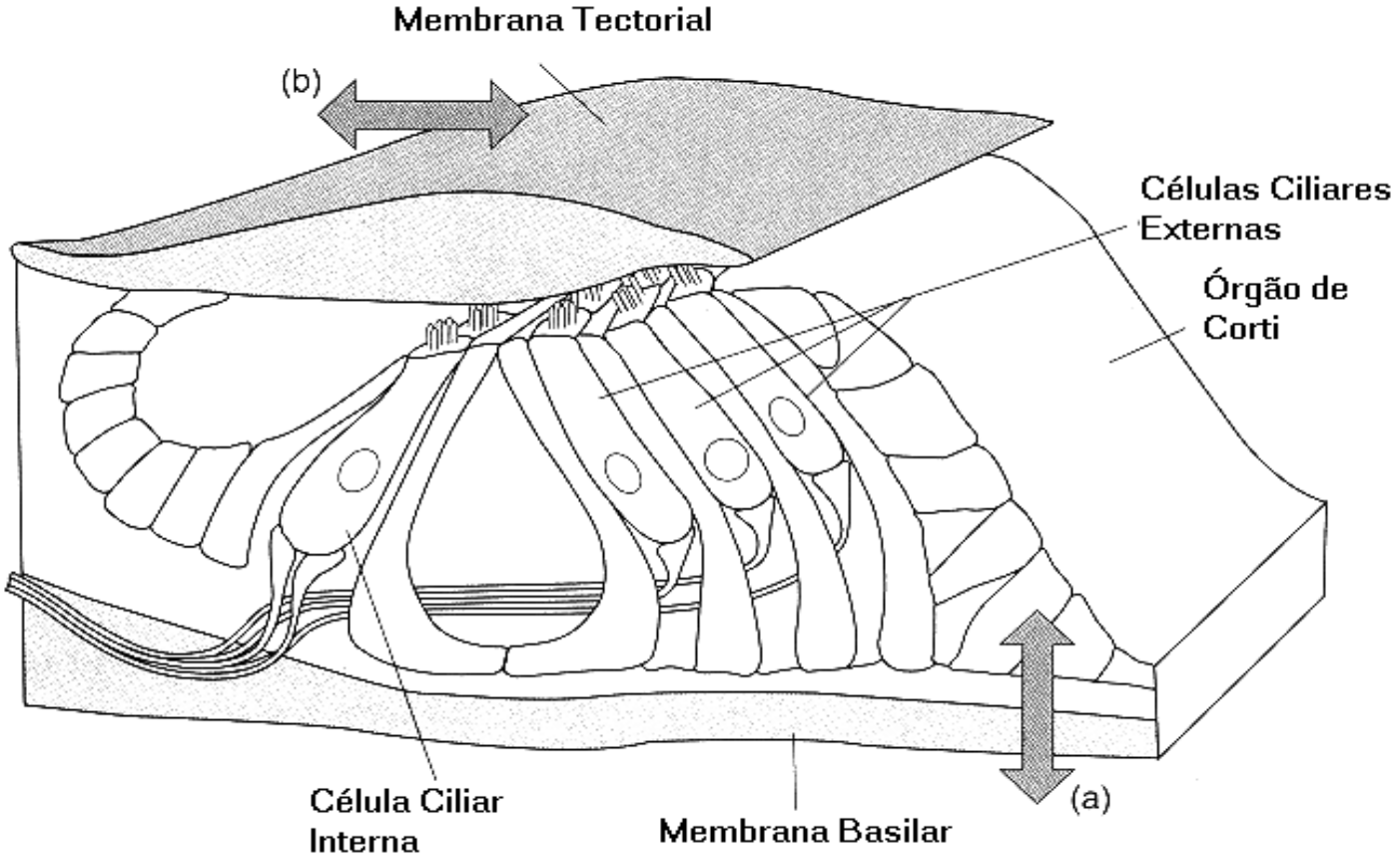


Célula Ciliar Externa

Células de Corti



Atuação das Células Ciliares

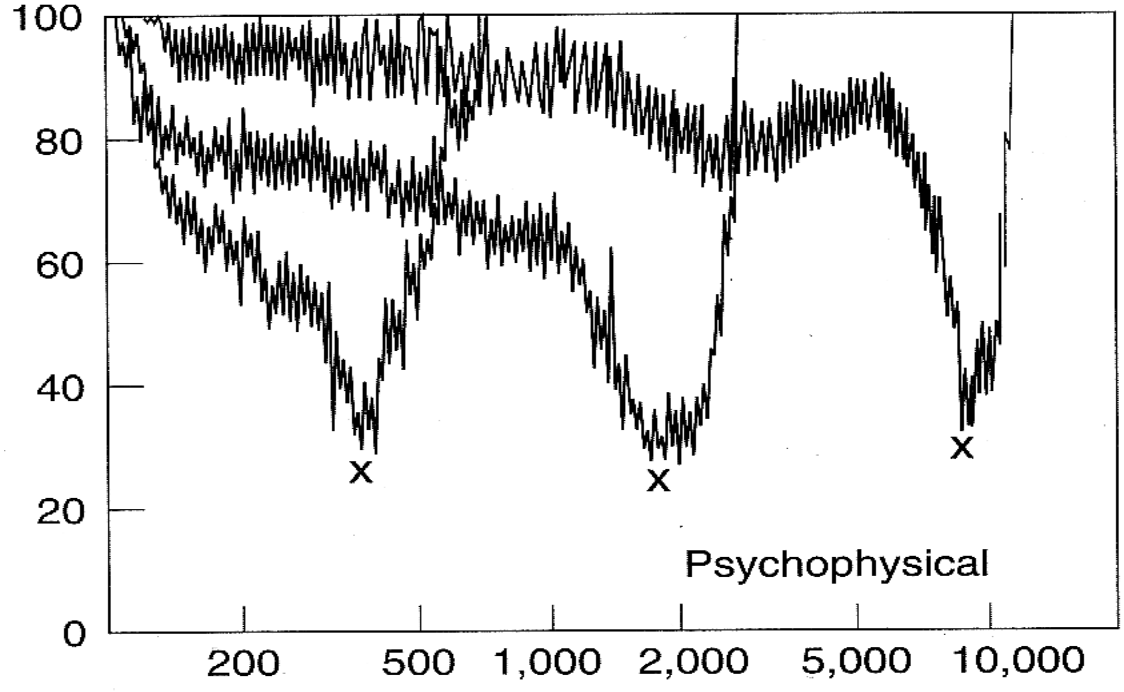


Alguns Números:

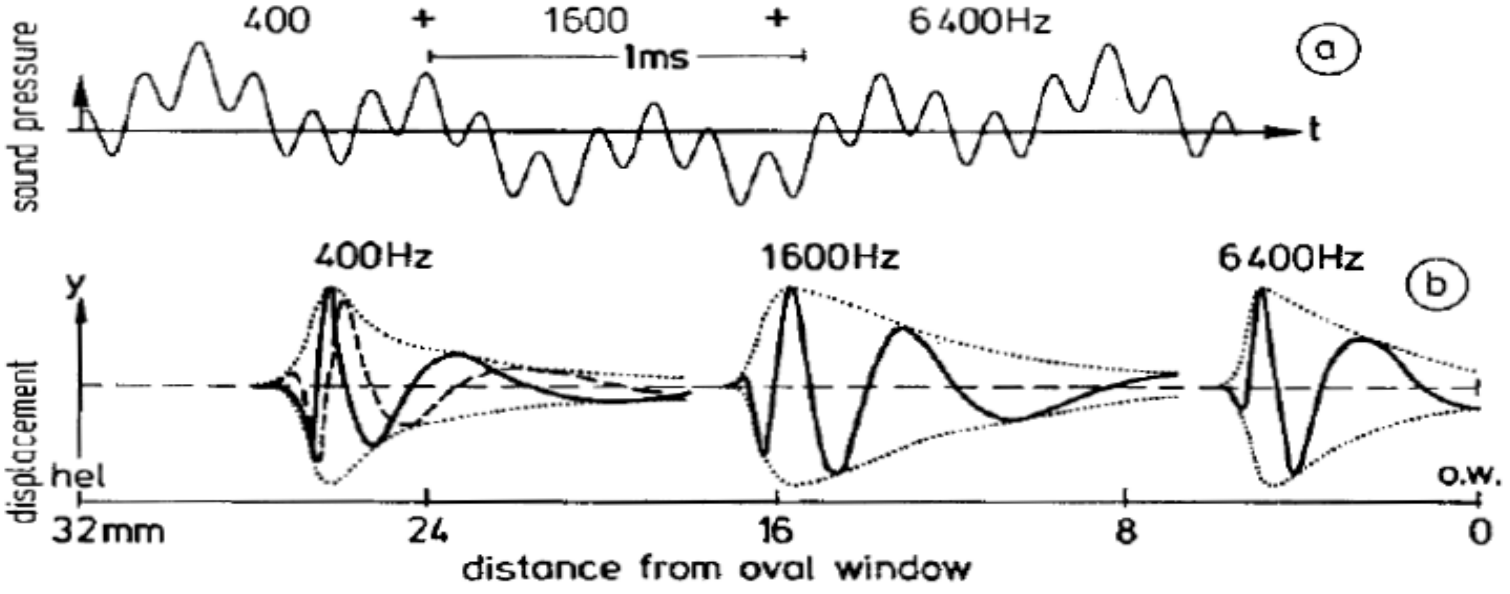
- 3.500 Células Ciliares Internas
- 12.000 Células Ciliares Externas
- 30.000 Fibras Nervosas Aferentes:
  - 90~95% provenientes das Células Internas
  - Células Internas com até 20 sinapses

- ~500 Fibras Nervosas Eferentes:
  - Destinadas às Células Externas

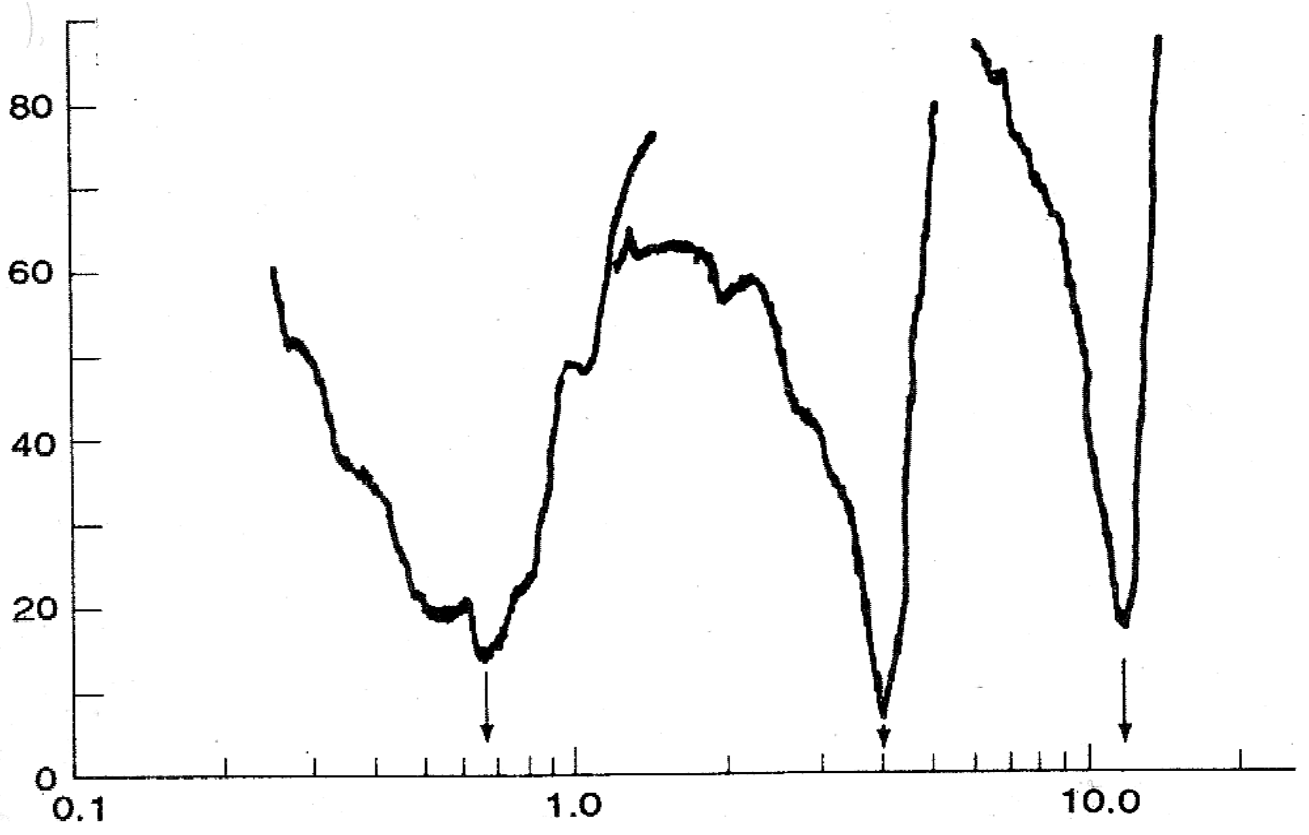
Seletividade da Membrana Basilar



Seletividade da Membrana Basilar



Seletividade das Células de Corti

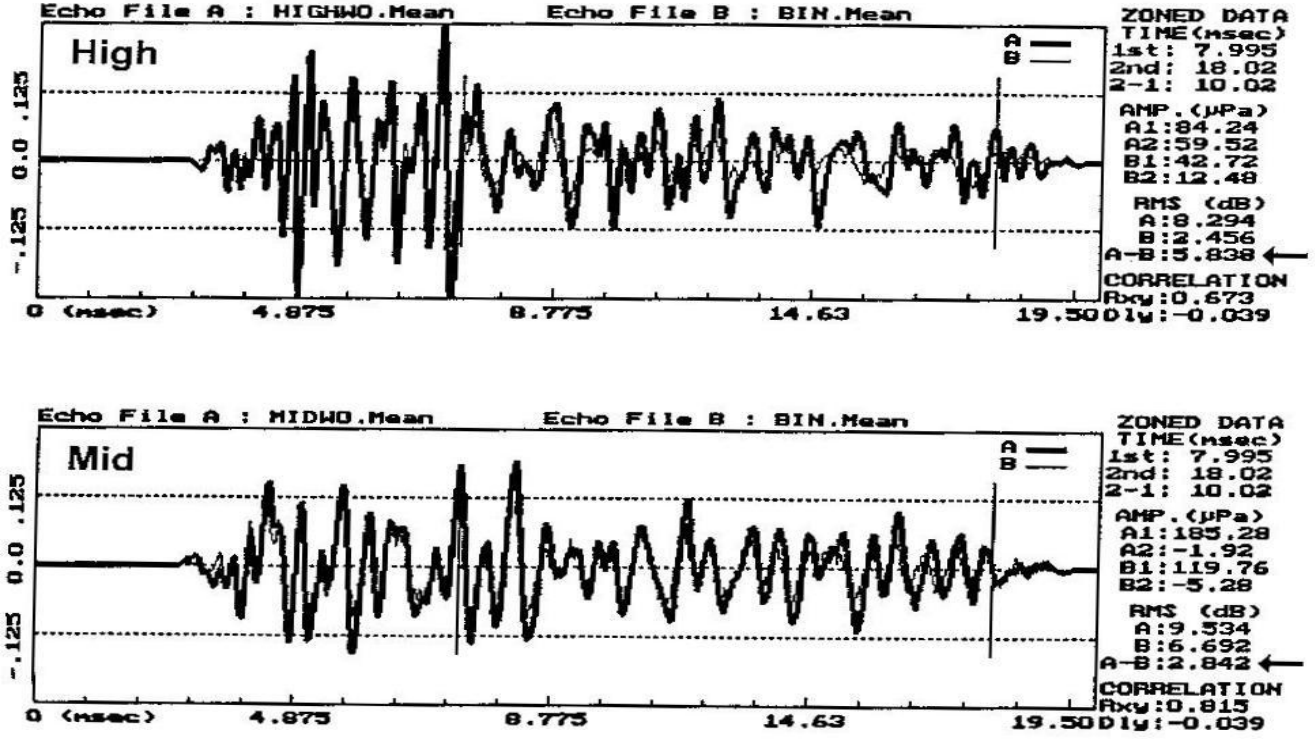


Funções das Células Ciliares

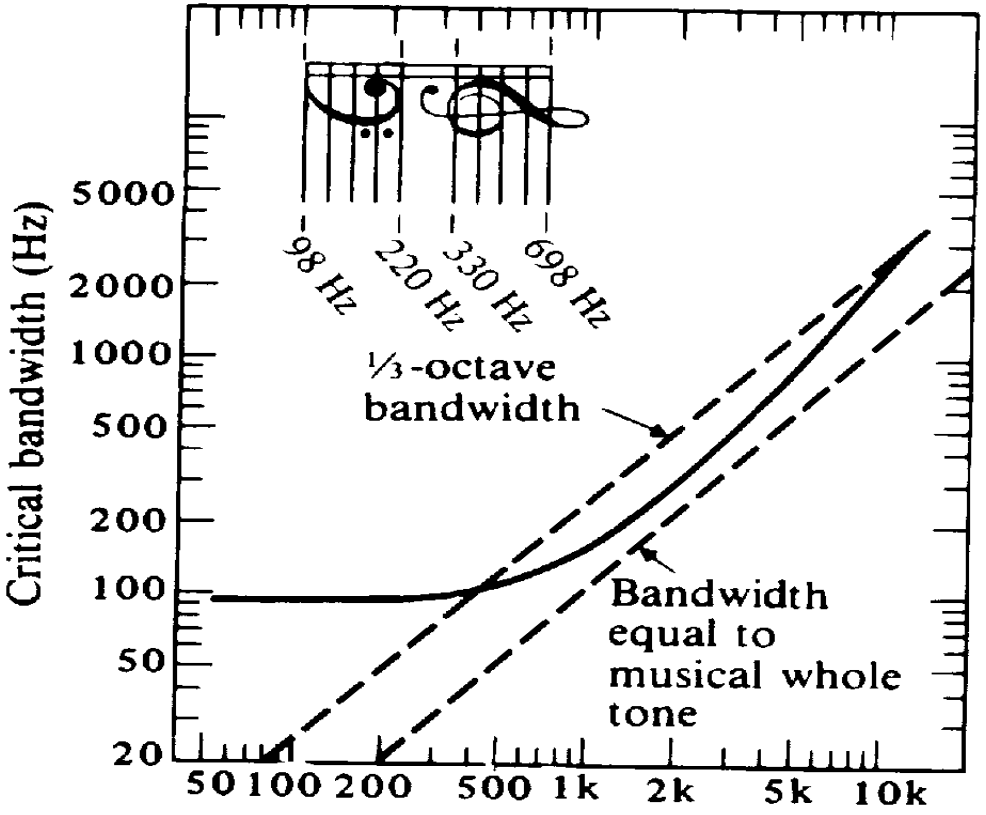
- Células Ciliares Internas:
  - Sensores
  - Faixa dinâmica: ~ 20 a 60 dB
- Células Ciliares Externas:
  - Atuadores
  - Responsáveis pela extensão da faixa dinâmica para até 120 dB
  - Responsáveis pelo aumento de seletividade

Emissões Oto-acústicas Evocadas

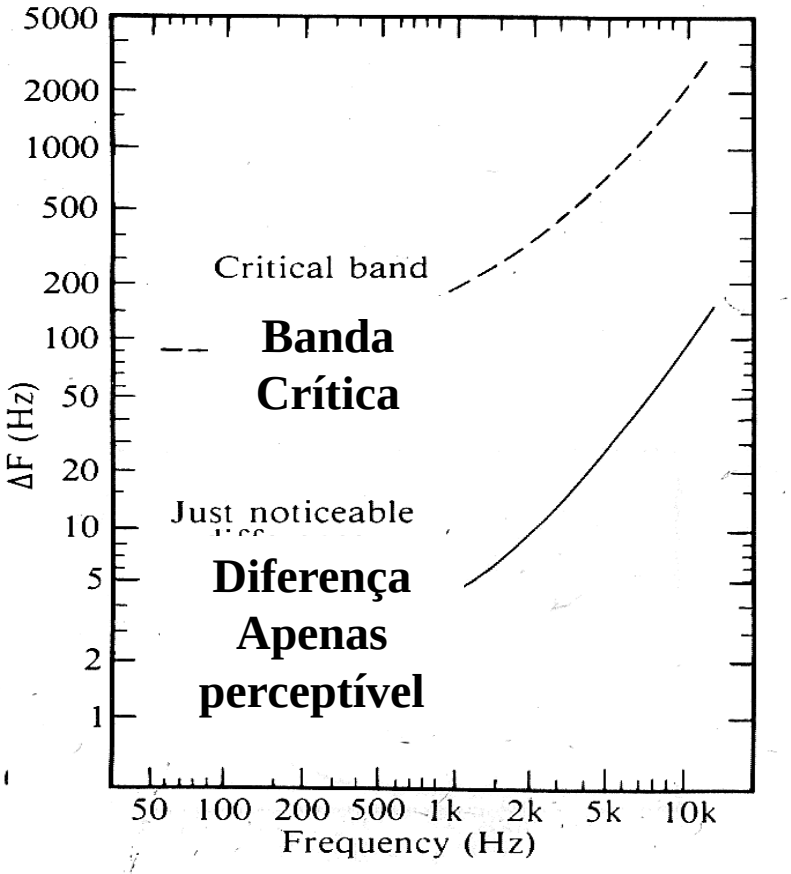
TEOAE (*Transient Evoked Oto-Acoustic Emission*)
- Evidências da atividade das células externas



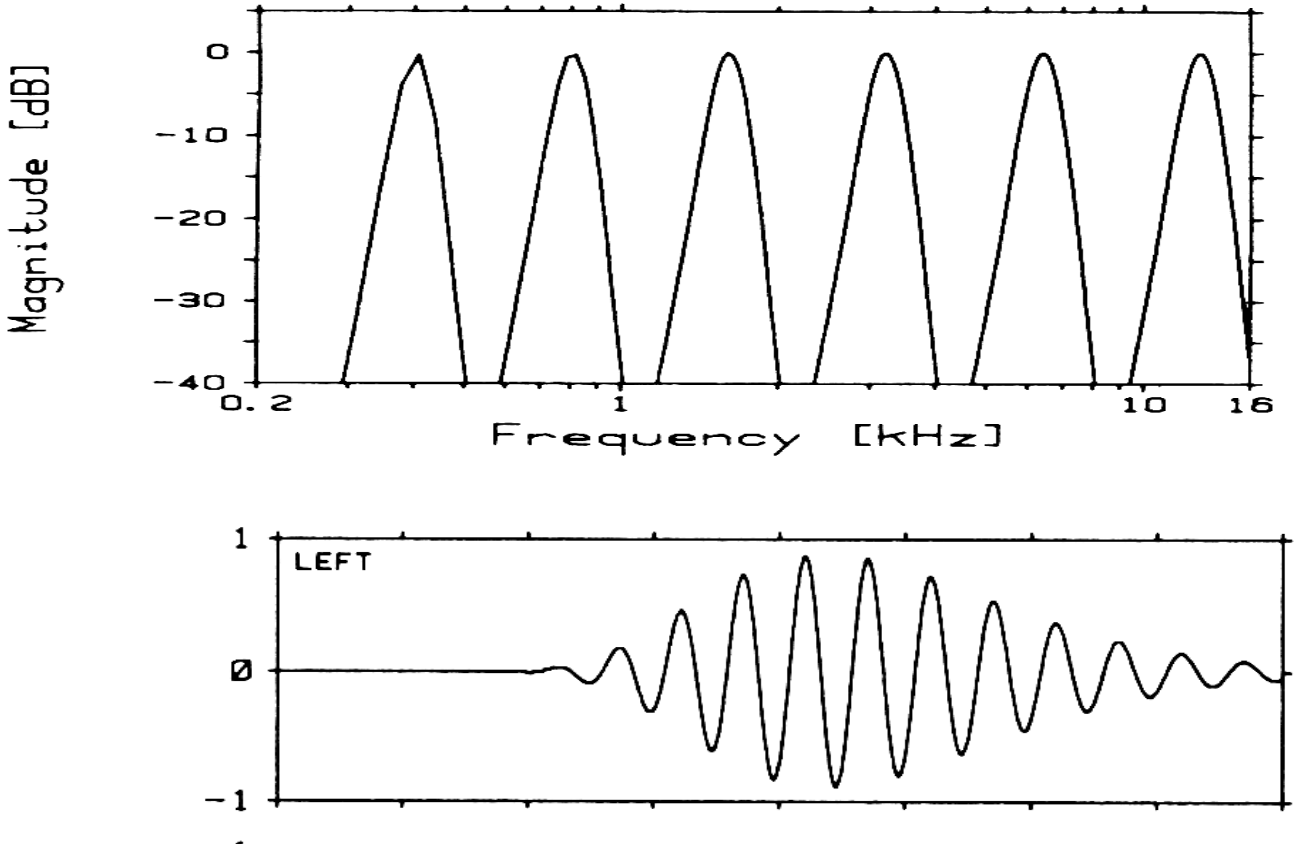
Bandas Críticas da Audição Humana



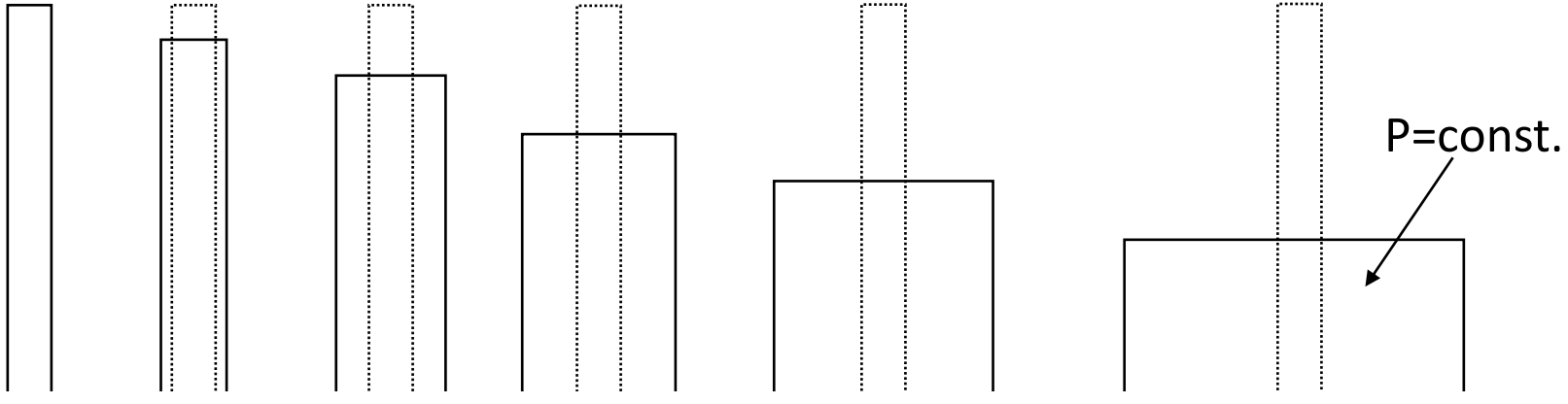
Limiar de Percepção de Frequências



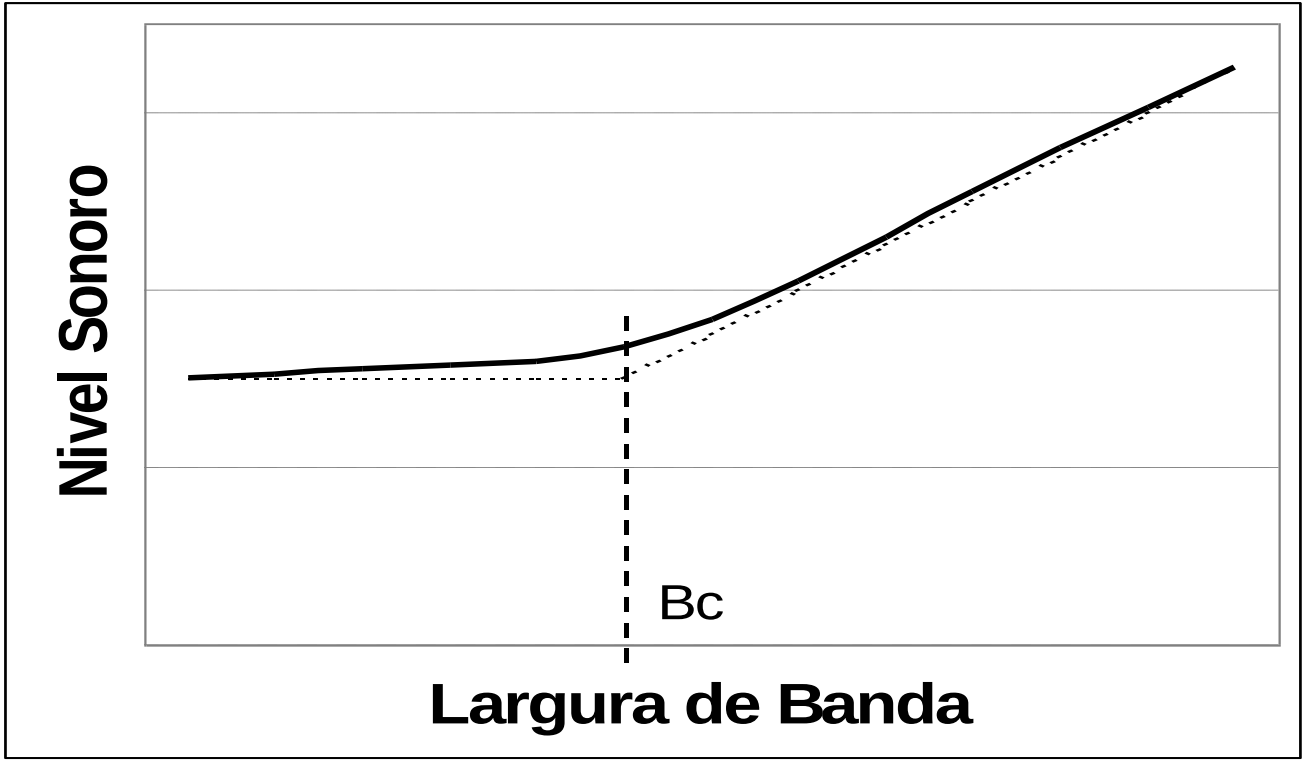
Resposta Temporal das Bandas Críticas



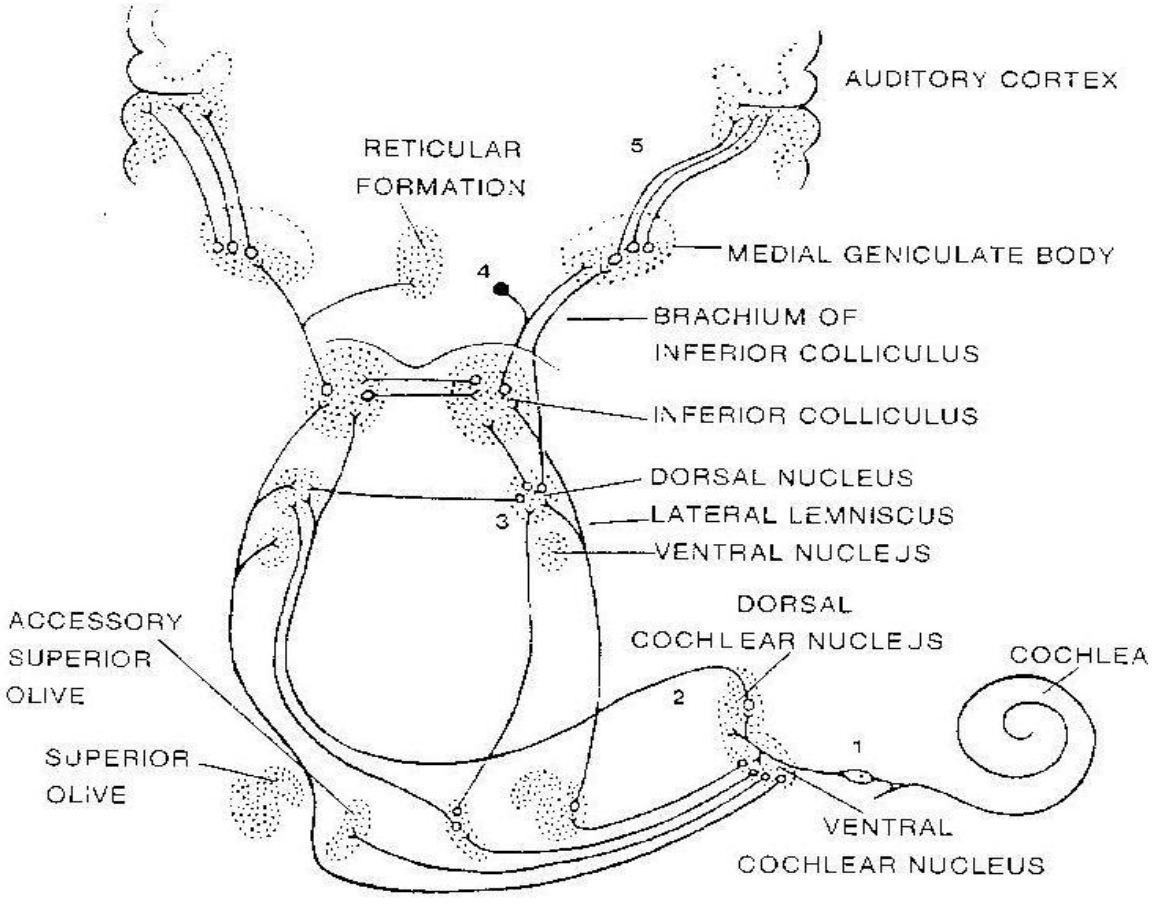
Determinação da Largura das Bandas Críticas



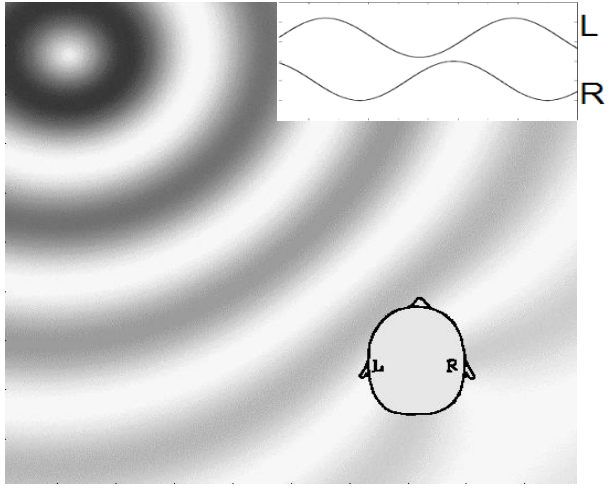
Determinação da Largura das Bandas Críticas



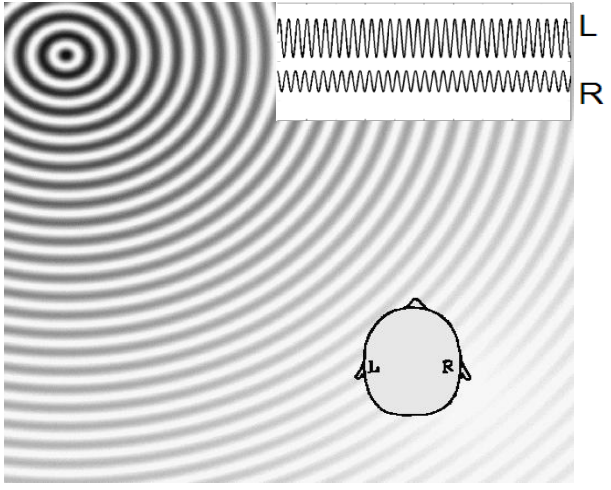
Pós-processamento nos Gânglios Auditivos



Percepção de Direção (Azimute)

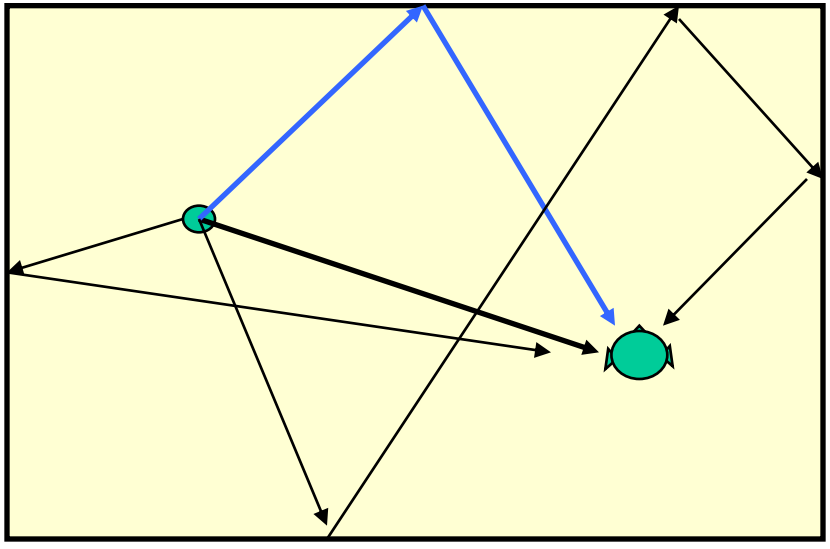


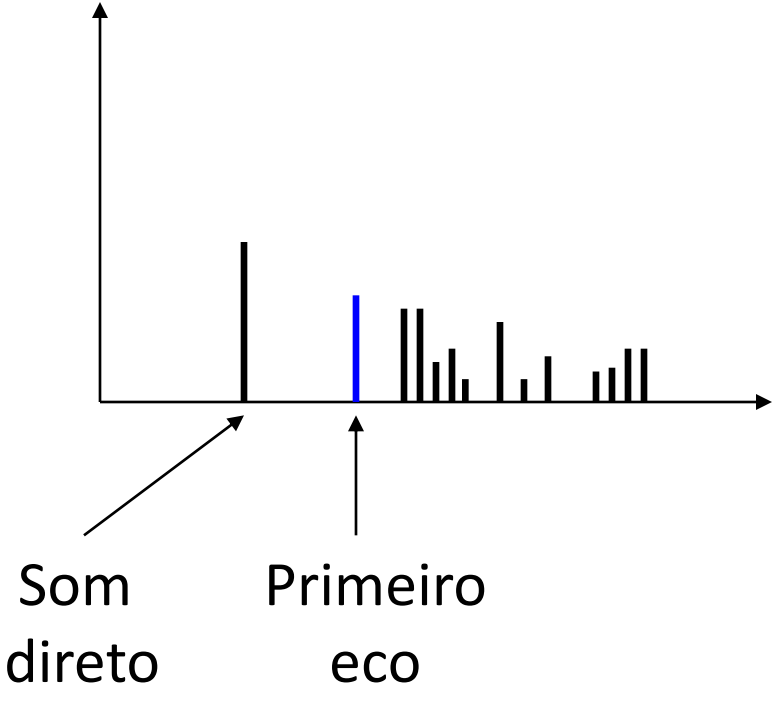
Baixas Frequências:
Diferença de Fase
(tempo de Percurso)



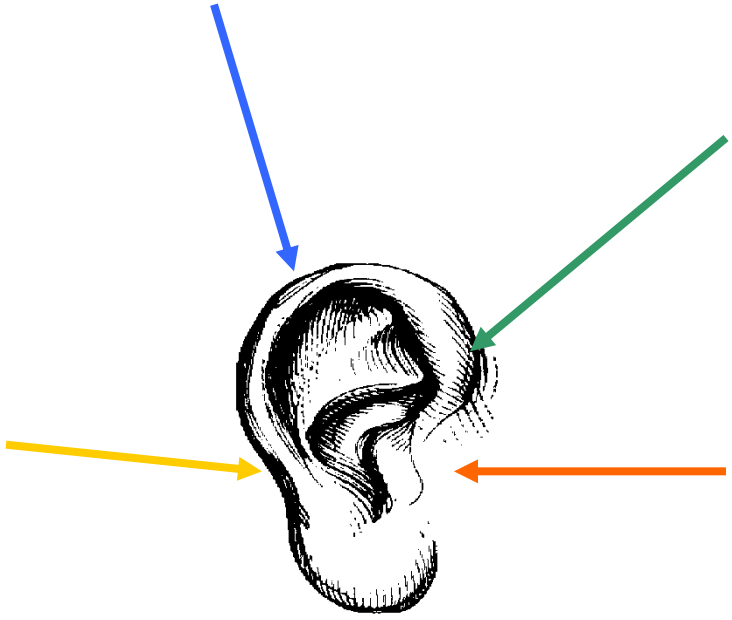
Altas Frequências:
Diferença de Intensidade
(difração)

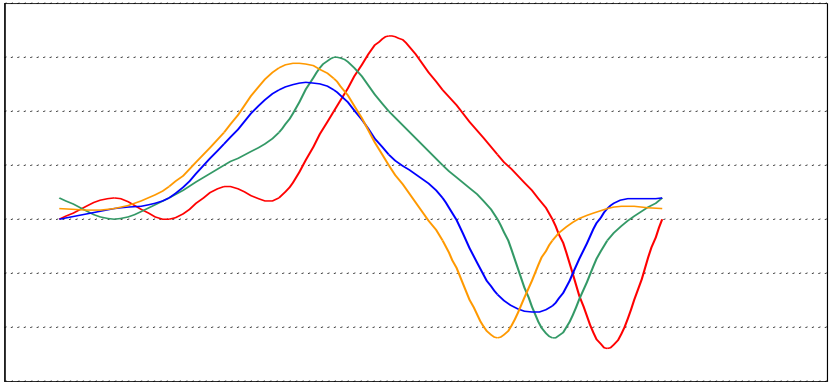
Percepção de Ambiente: Reflexões Múltiplas



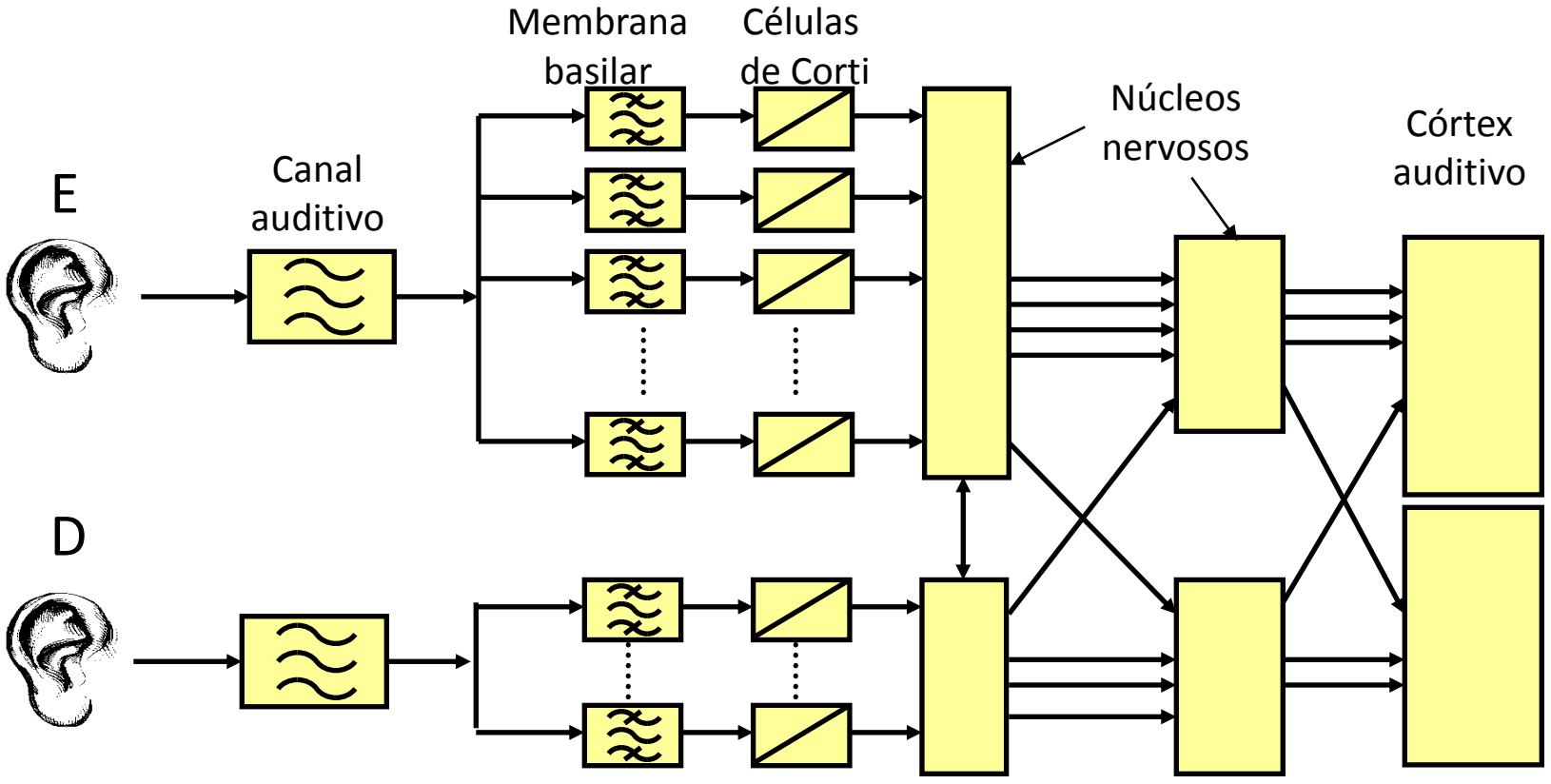


Percepção de Direção (Elevação)





Um Diagrama de Blocos da Audição



Acústica Física

Pressão Sonora: Unidade de Medida

- Limiar de Audibilidade: 20 μPa *rms* = $2x10^{-5}$ N/m^2 @ 1 kHz (1 pw/ m^2)

$$\text{Pressão Sonora} = 20\log\frac{p}{p_o} \qquad (\text{dbSPL}) \text{ onde } p_o = 20\mu\text{Pa}$$

- (Pressão atmosférica: ~ 100 kPa)

Nível Sonoro: Unidade de Medida

- Volume Sonoro = grandeza subjetiva associada à percepção sensorial da intensidade de um som

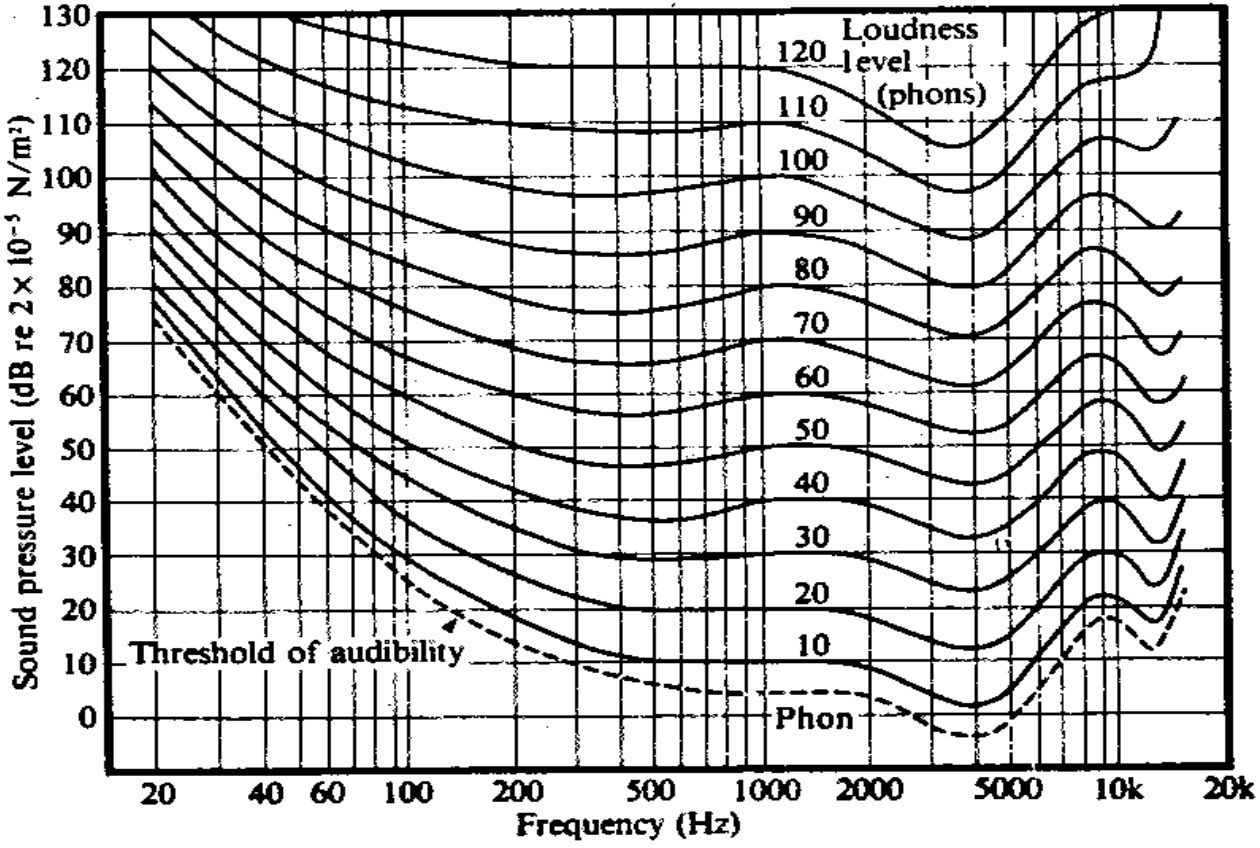
- 1 Phon = 1 dB SPL @ 1 kHz

Situações Sonoras

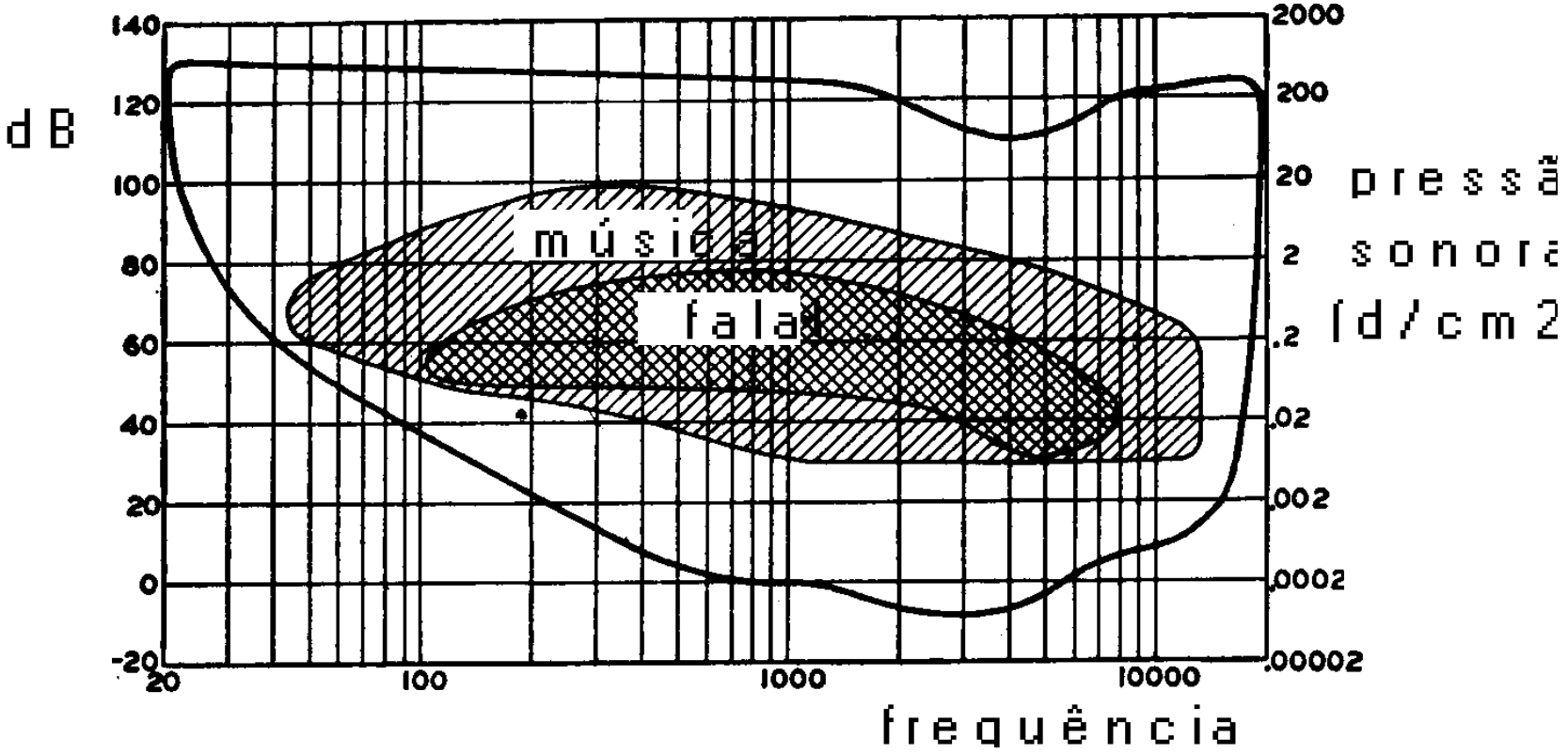
	dB SPL
Foguete	195
Avião a jato	155
Limiar de dor	140
Limiar de desconforto	120=1W/m^2
Orquestra fortíssimo	110
Rebitadeira	100
Fábrica	78
Tráfego pesado	68
Escritório ruidoso	65

	dB SPL
Conversação	65
Restaurante	60
Residência urbana	40
Casa de campo	30
Orquestra pianíssimo	30
Estúdio de gravação	20
Folhagens na brisa	10
Limiar de audição	0=1pw/m^2
Ruído térmico do ar	-10

Curvas de Igual Volume Sonoro (tons senoidais)

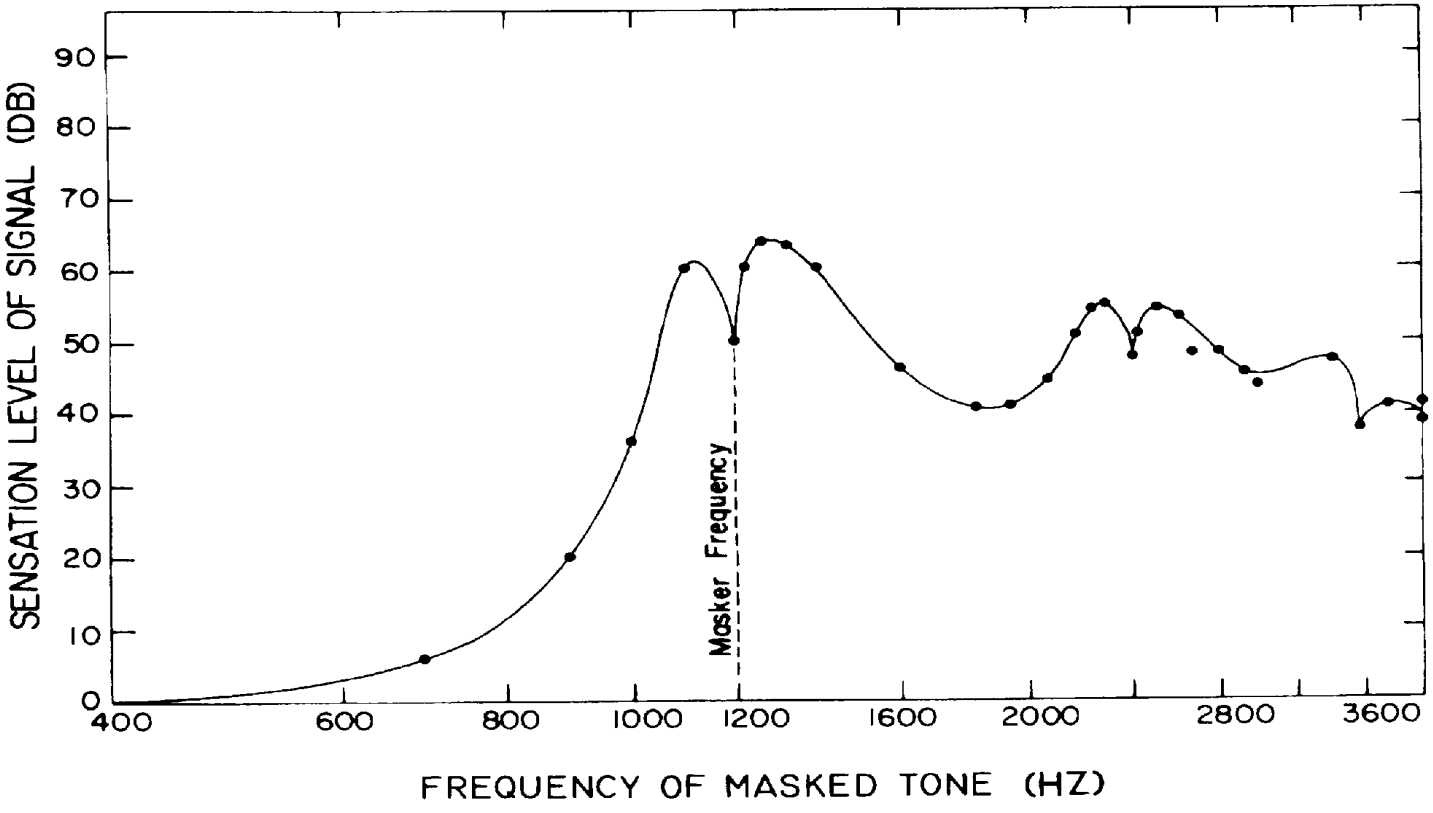


Faixa Dinâmica para Música e Voz

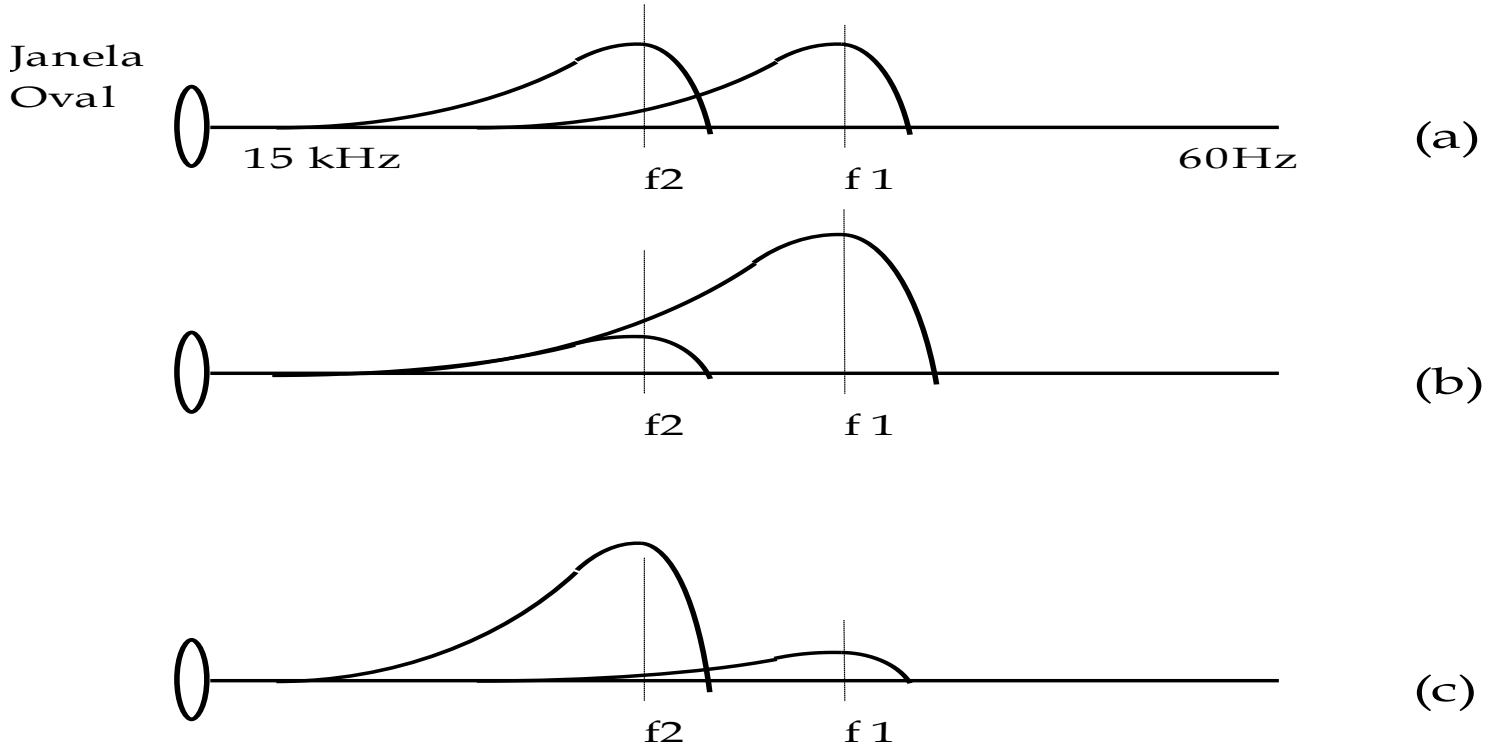


Mascaramento

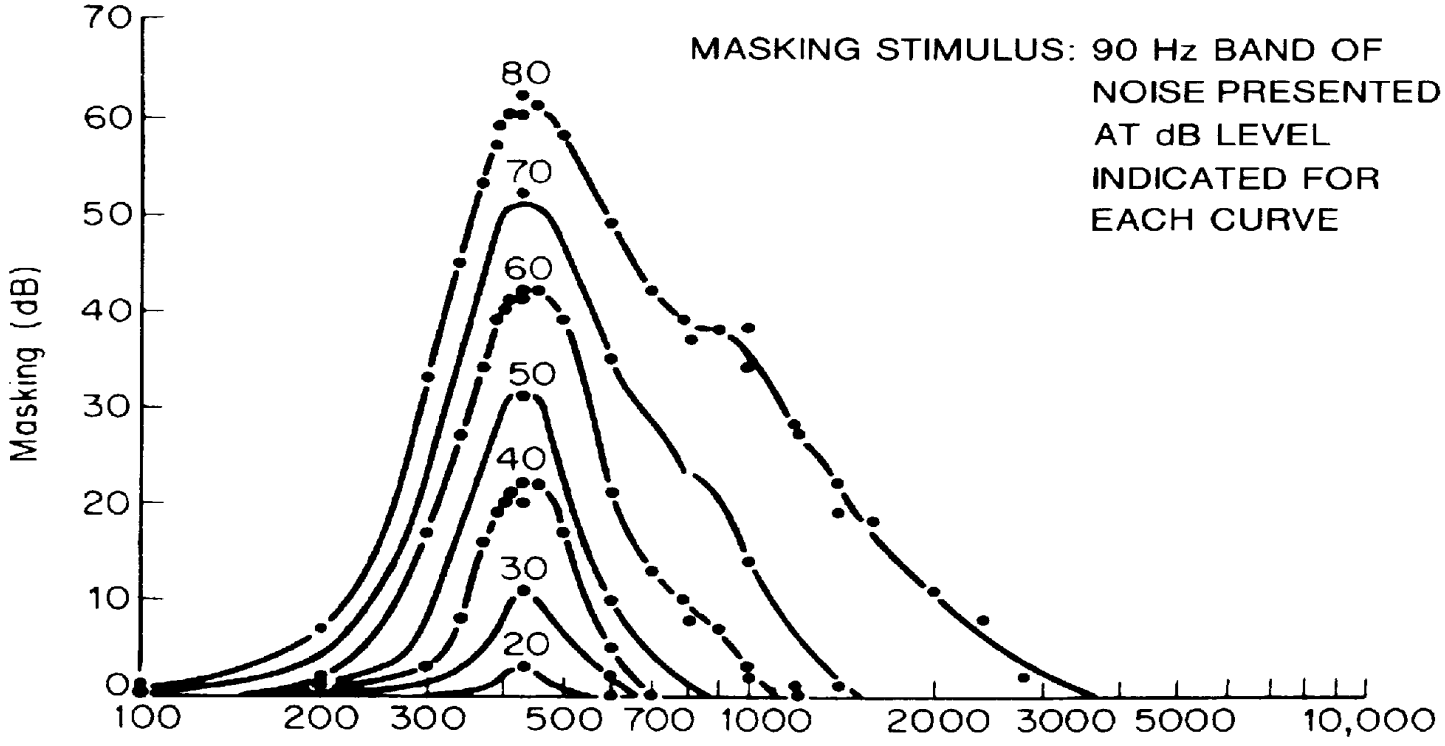
Mascaramento para Tom Senoidal



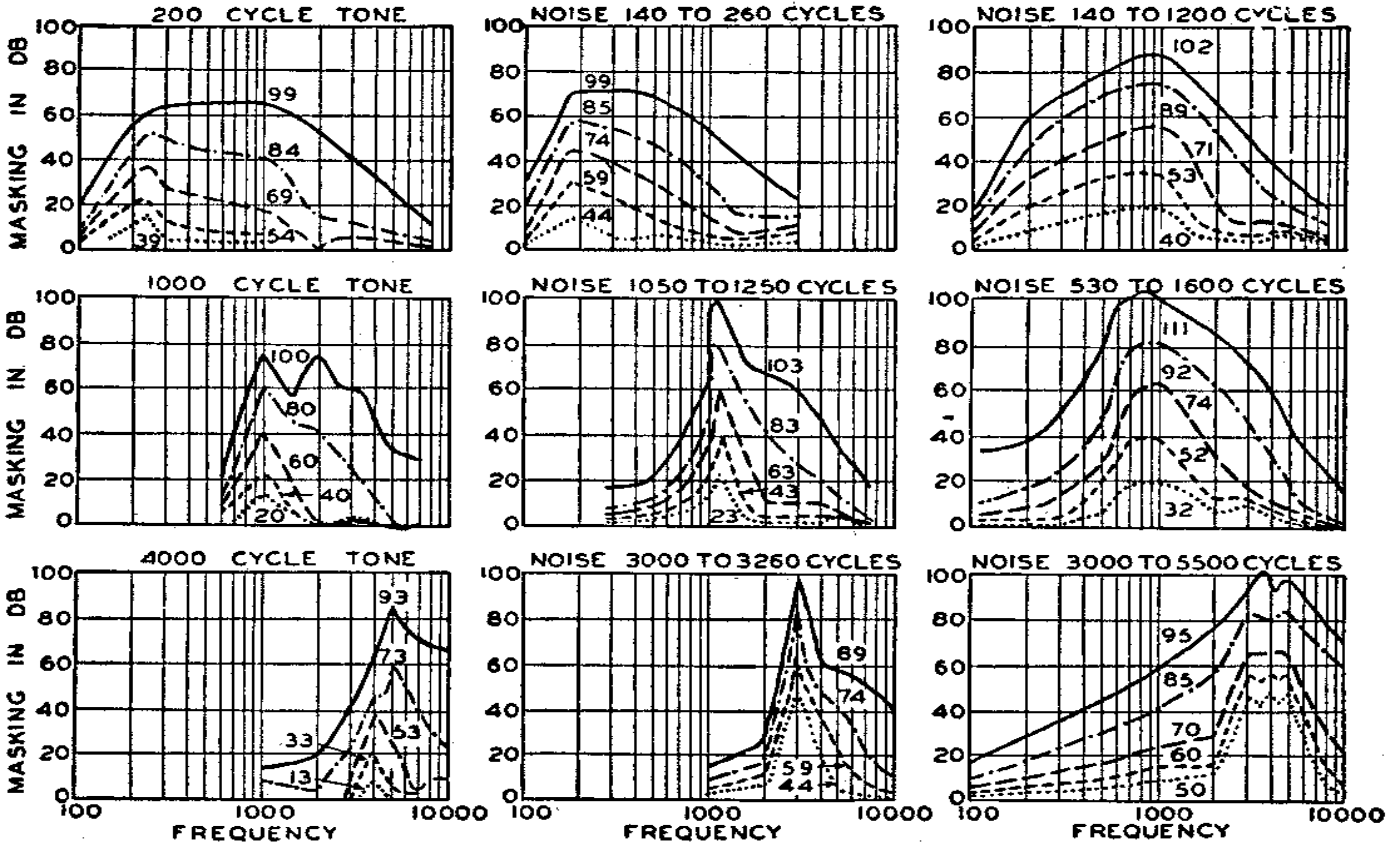
Resposta da Membrana Basilar para 2 Tons



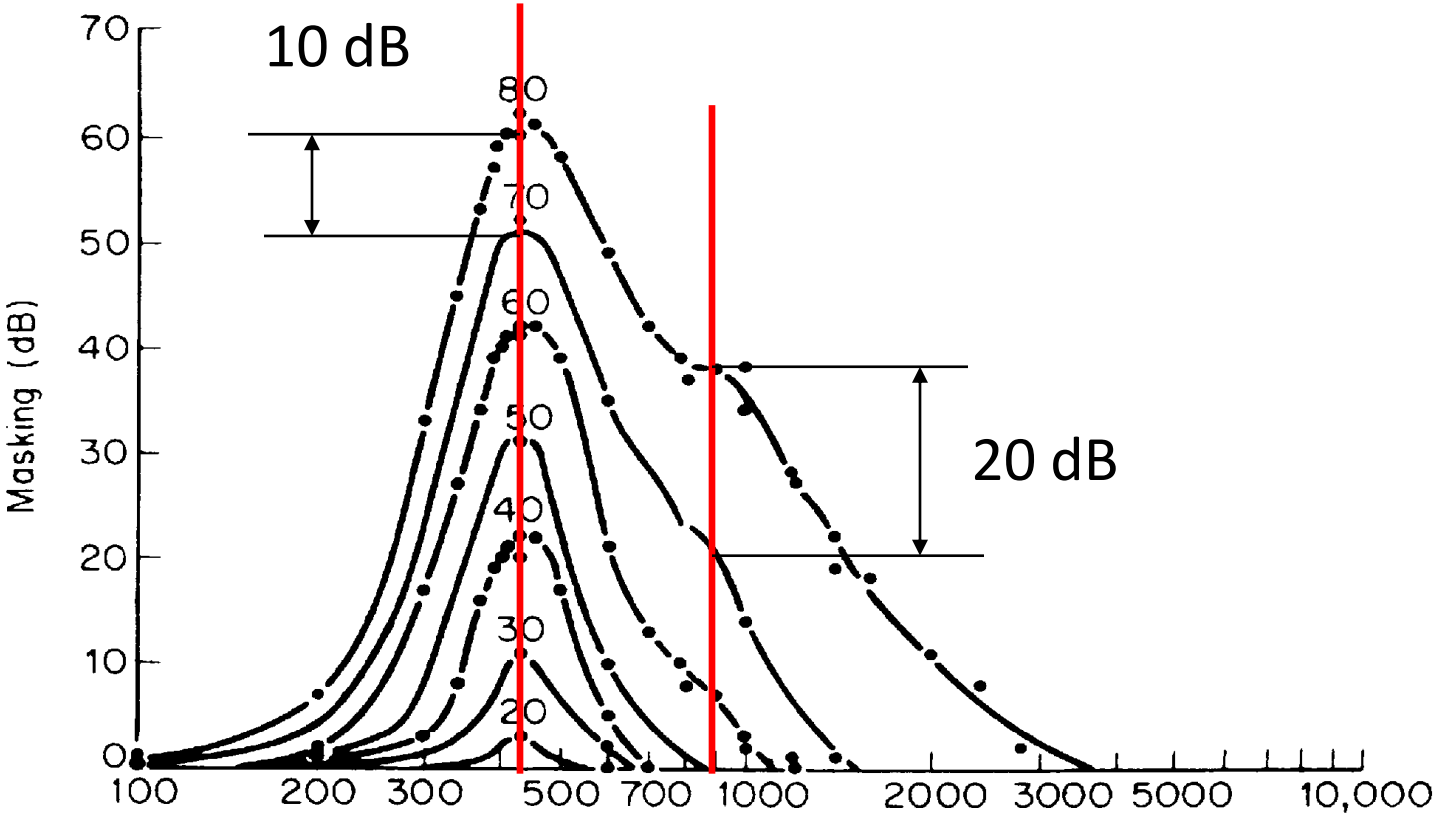
Mascaramento para Ruído de Banda Estreita



Curvas de Mascaramento (ex.)

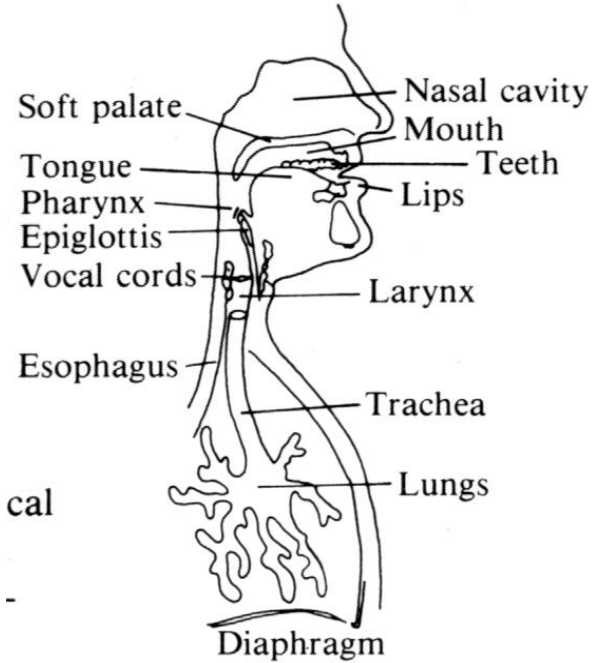


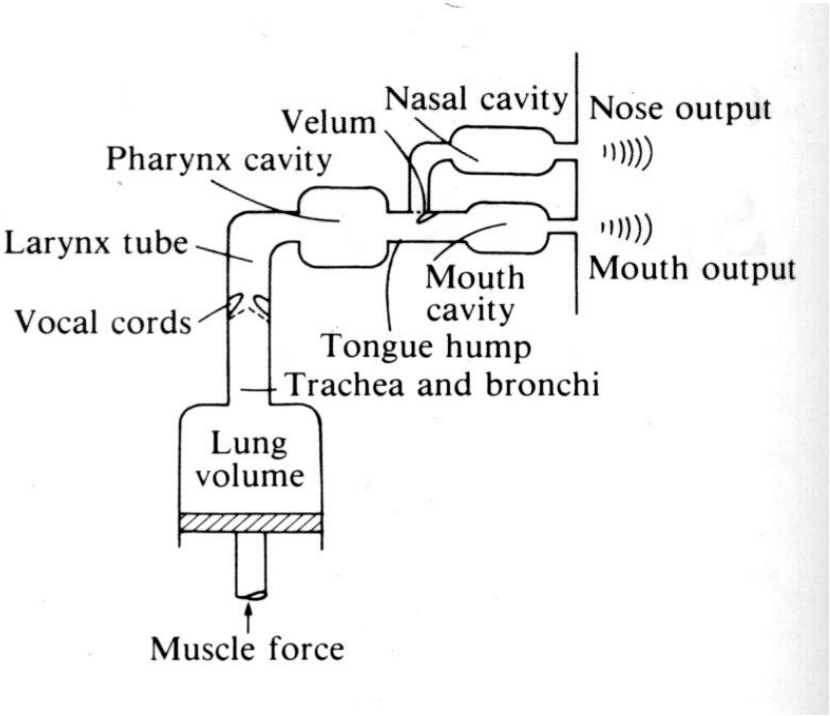
Evidências de Distorção Aural



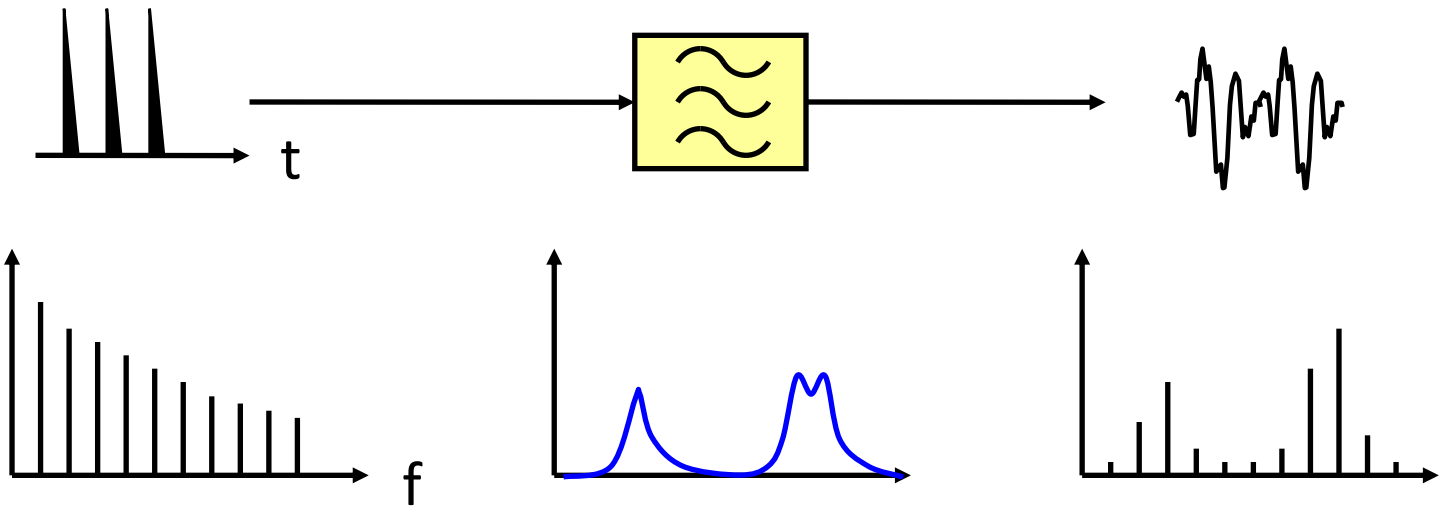
A Voz Humana e o Vocoder

A Voz Humana


Soft palate
Nasal cavity
Mouth
Tongue
Teeth
Pharynx
Lips
Epiglottis
Vocal cords
Larynx
Esophagus
Trachea
Lungs
cal
-
Diaphragm


Velum
Nasal cavity
Nose output
Pharynx cavity
Larynx tube
Mouth output
Mouth cavity
Vocal cords
Tongue hump
Trachea and bronchi
Lung volume
Muscle force

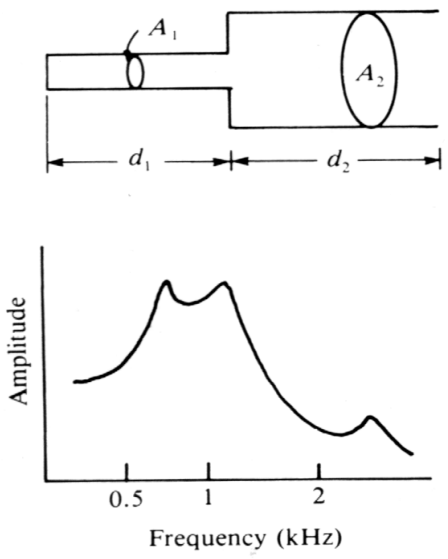
Características Espectrais da Voz



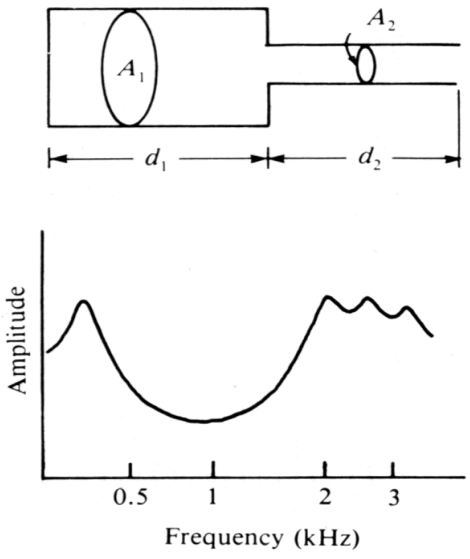
Fundamental:
100-150 Hz (M)
200-300 Hz (F)

Formantes:
250-800 Hz (1)
800-2300 Hz (2)
2200-3000 Hz (3)

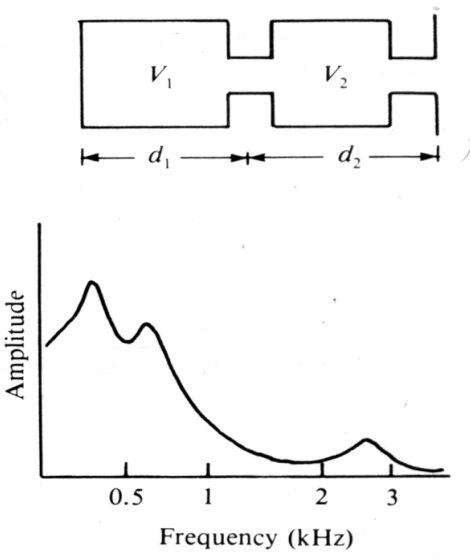
Características das Formantes Espectrais



"a"

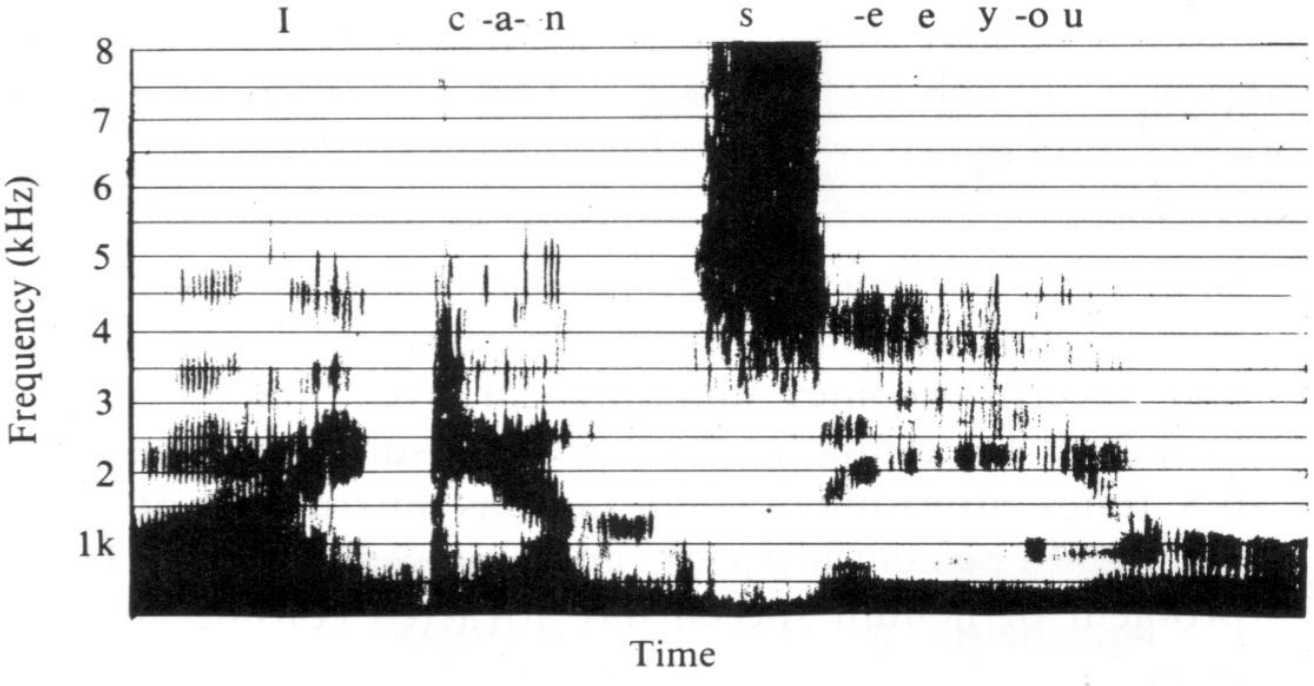


"i"

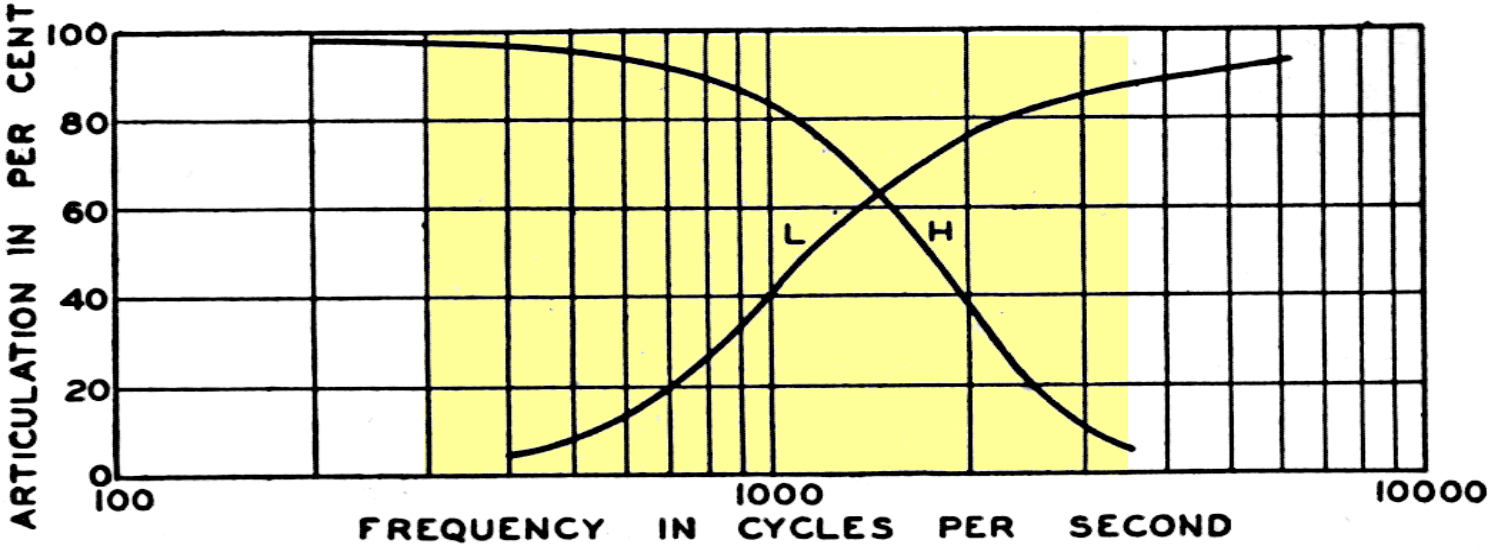


"u"

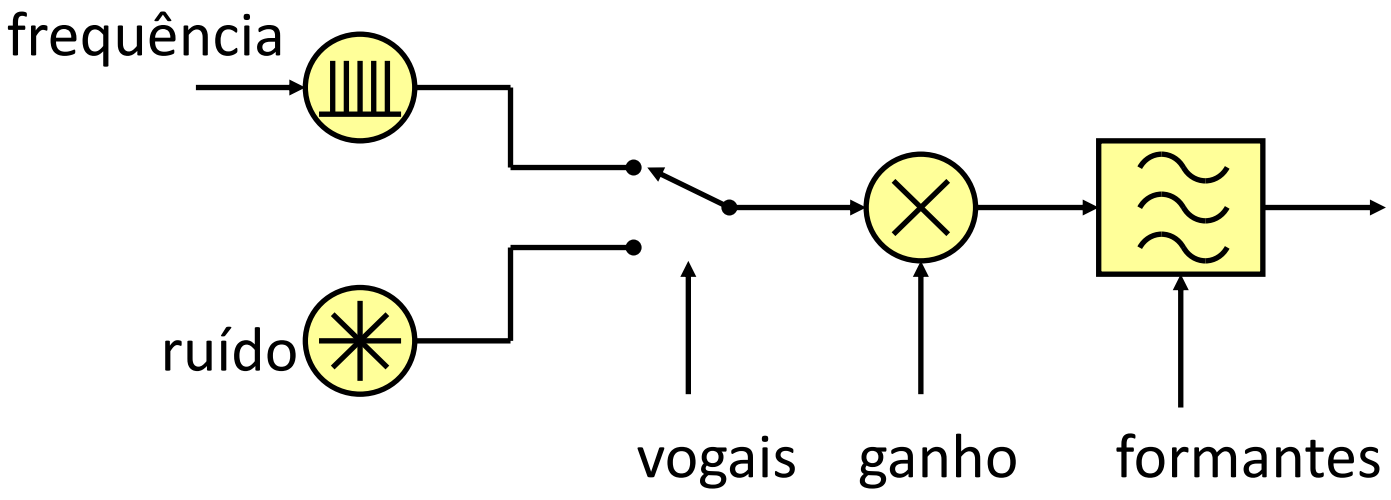
Espectrograma (Tempo / Frequência) da Voz



Inteligibilidade da Fala



Síntese de Voz (Vocoder)



Exemplo de Compressão: 1 Minuto de Áudio

- PCM 16 bits, 44.1 kHz, estéreo:
  - 85 Mbits
- MPEG-1 Nível 3 (MP3) a 128 kb/s:
  - 8 Mbits
- LPC (Linear Predictive Coding - voz):
  - 400 kbits
- MIDI (Musical Instruments Digital Interface):
  - 250 kbits típico

Modelamento de Sistemas Dinâmicos

- Modelo Elétrico de um Sistema Mecânico
  - Corrente $\Leftrightarrow$ velocidade
  - Tensão $\Leftrightarrow$ fôrça
  - Carga $\Leftrightarrow$ Deslocamento
  - Resistência $\Leftrightarrow$ Atrito Viscoso
  - Indutância $\Leftrightarrow$ Massa
  - Capacitância $\Leftrightarrow$ Compliância (Elasticidade^{-1})

Analogias Dinâmicas

- Indutância

$$e = L\frac{di}{dt}$$

- Capacitância

$$v = \frac{q}{C} = \frac{1}{C}\int i\ dt$$

- Resistência

$$e = R \times i$$

- Massa

$$f = M\frac{dv}{dt}$$

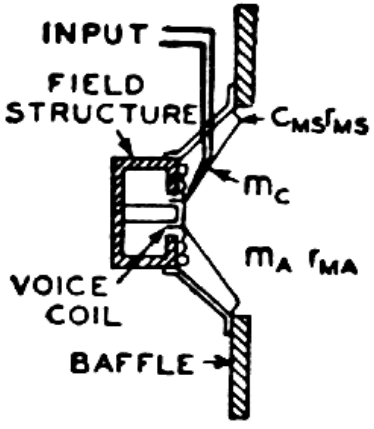
- Compliância

$$f = \frac{x}{C_M} = \frac{1}{C_M}\int v\ dt$$

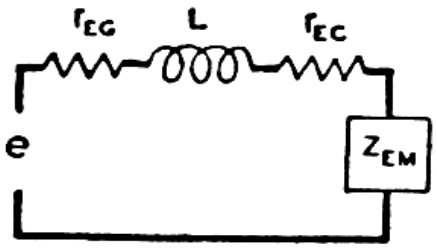
- Viscosidade

$$f = R_M \times v$$

Modelo de um Alto-Falante

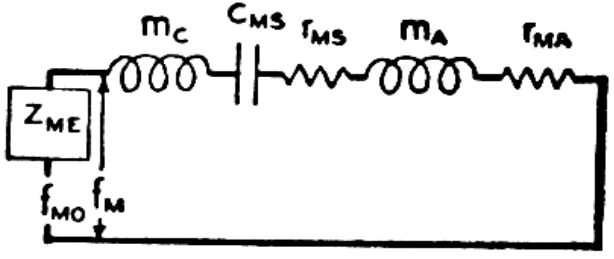


CROSS-SECTIONAL VIEW



VOICE COIL

ELECTRICAL CIRCUIT



MECHANICAL CIRCUIT OF THE MECHANICAL SYSTEM

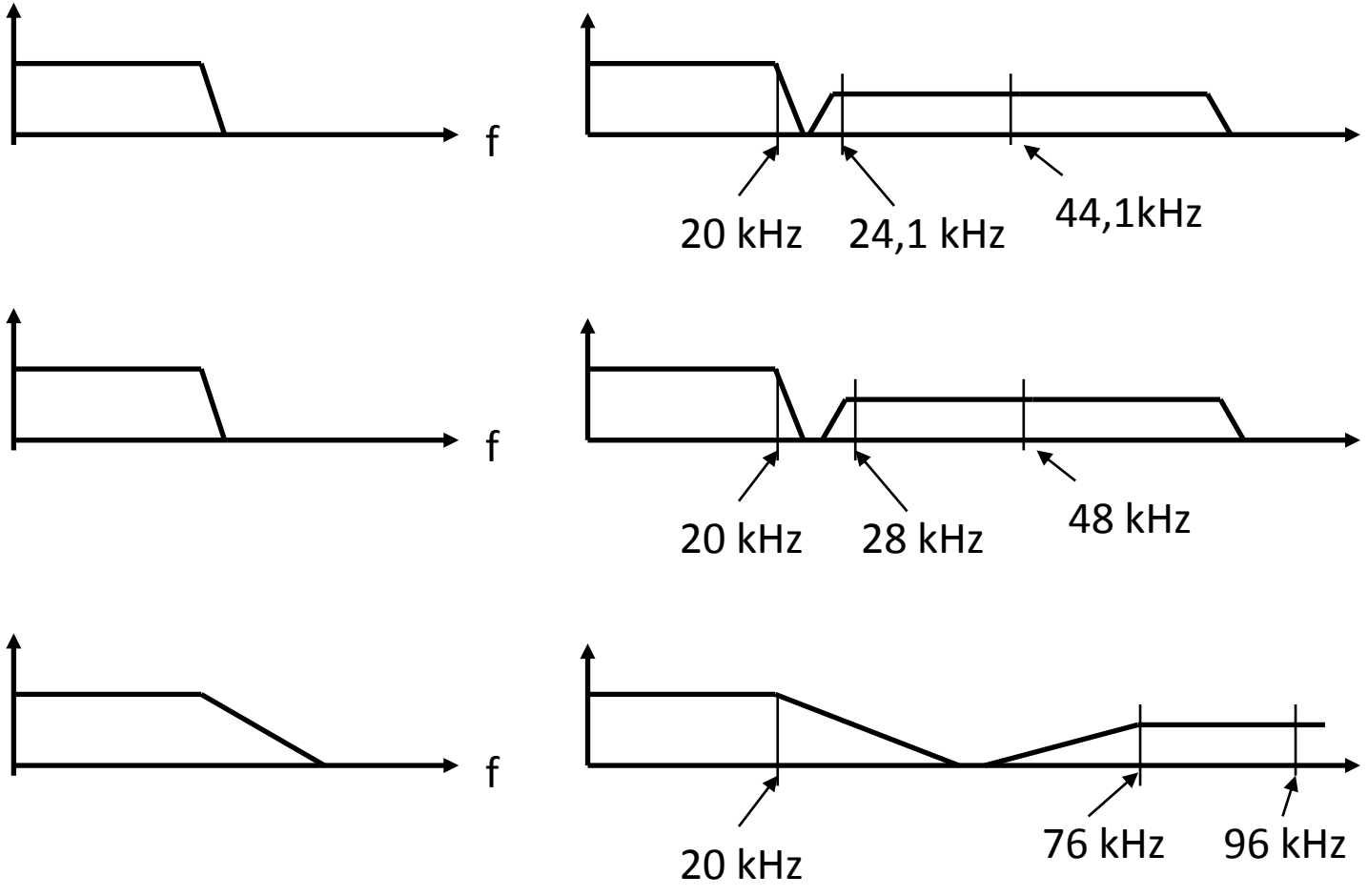
Questões Polêmicas

É Necessário Amostrar em 96 kHz?

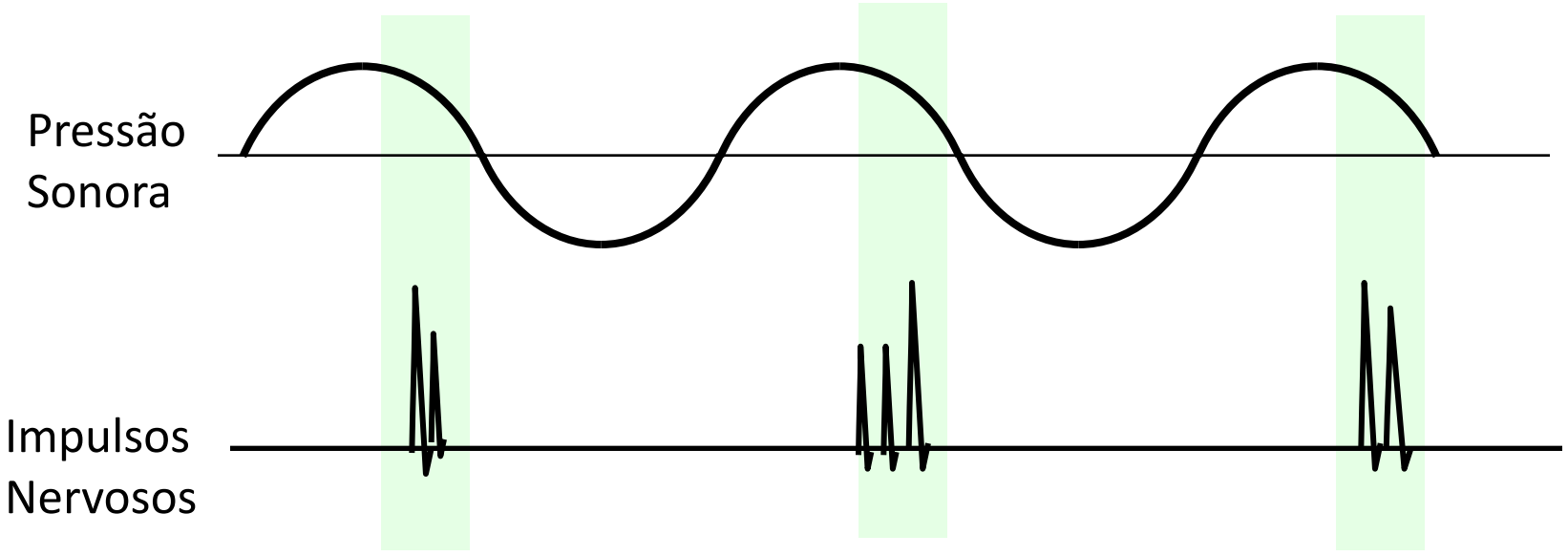
f_a = 48 kHz

f_a = 96 kHz

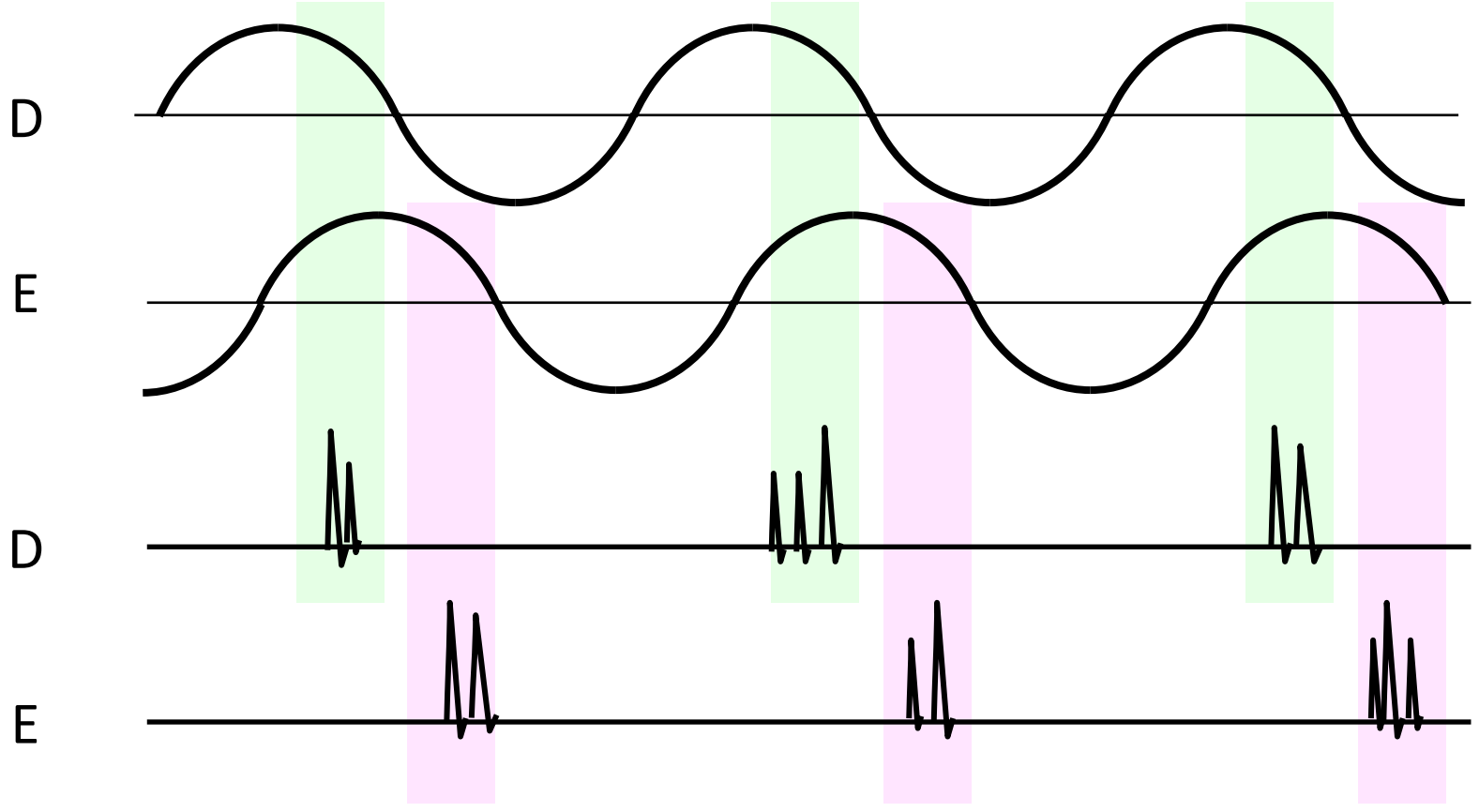
Vantagens da Amostragem em 96 kHz



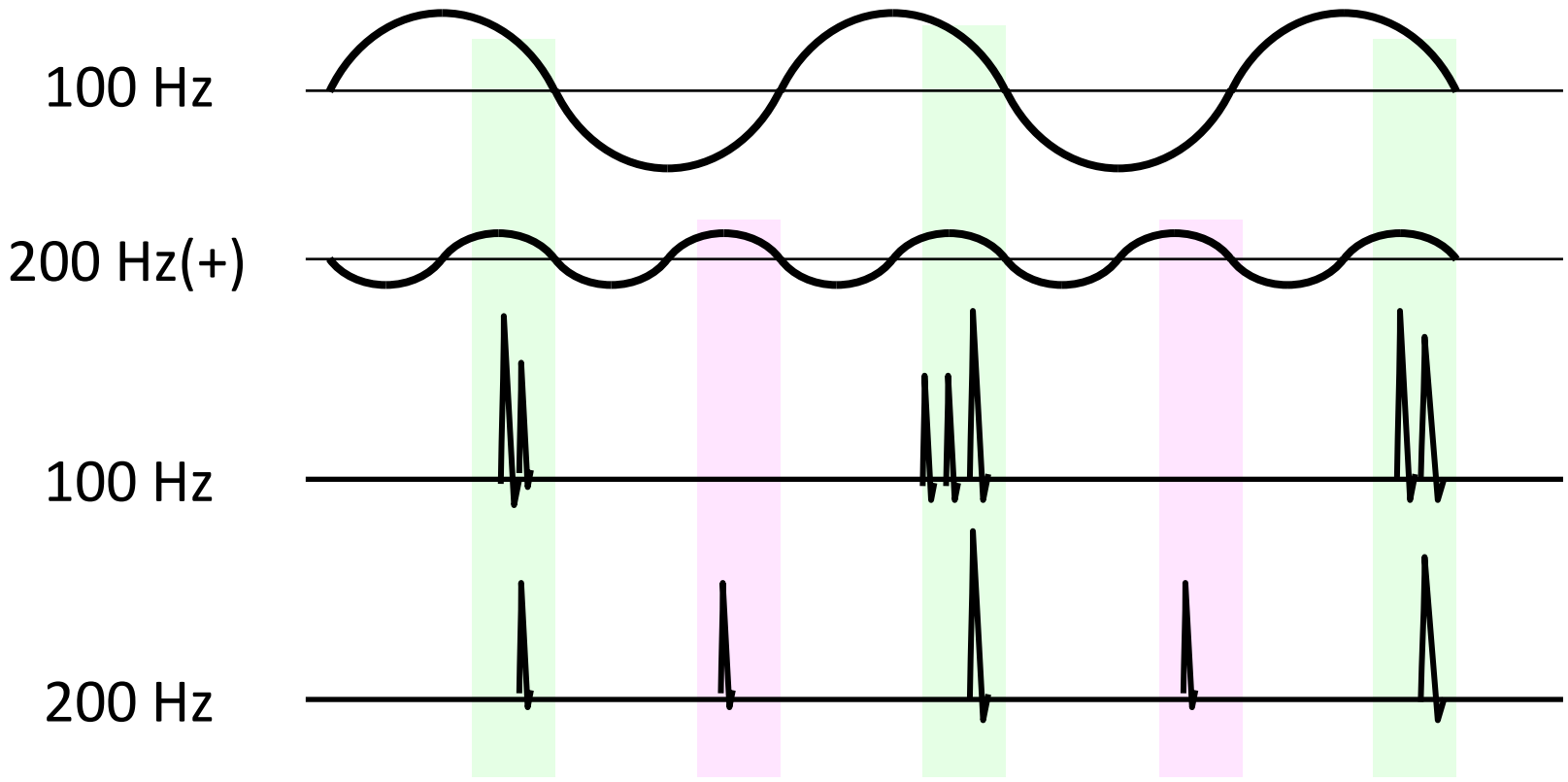
Questões Polêmicas – Audibilidade de Fase



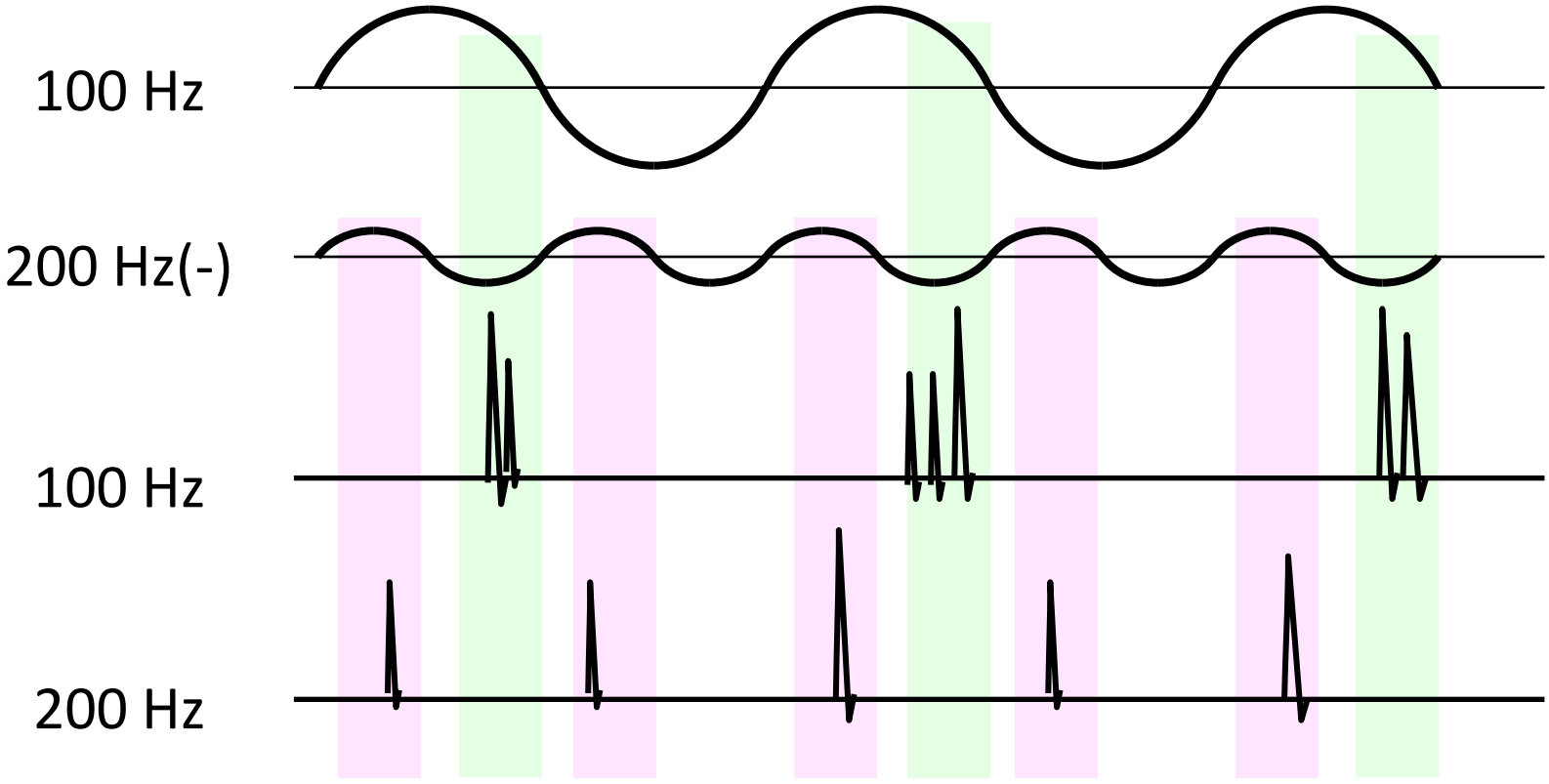
Determinação de Direção pela Defasagem



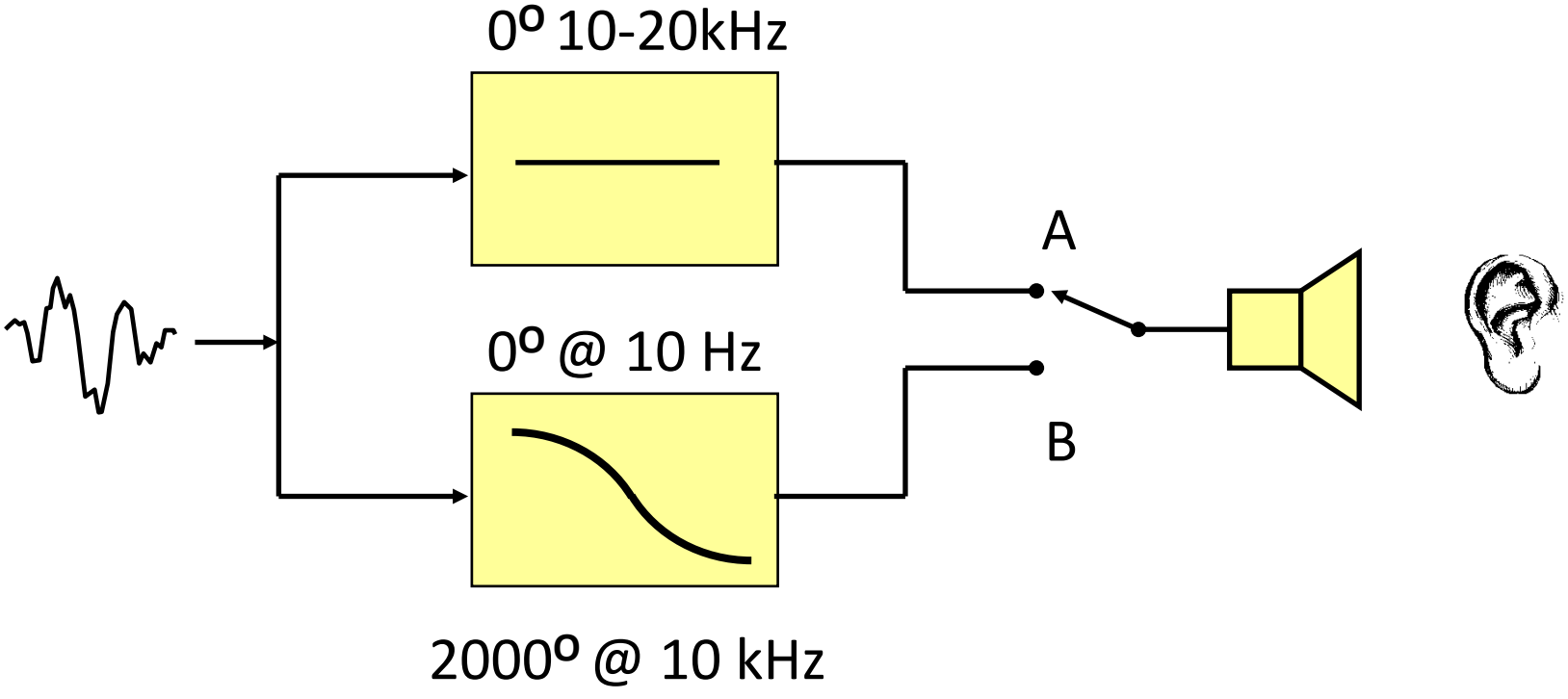
Fase Relativa de Harmônicas



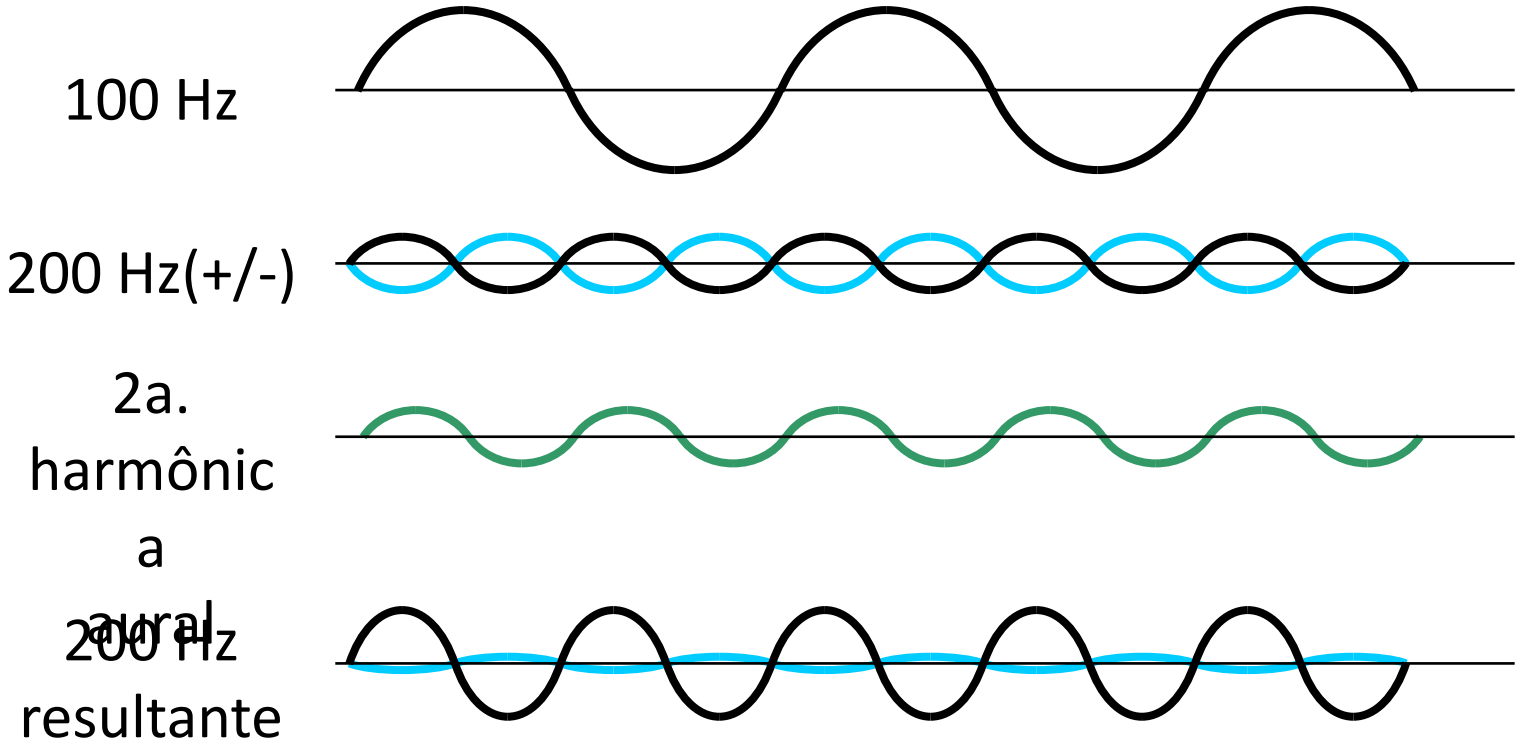
Fase Relativa de Harmônicas



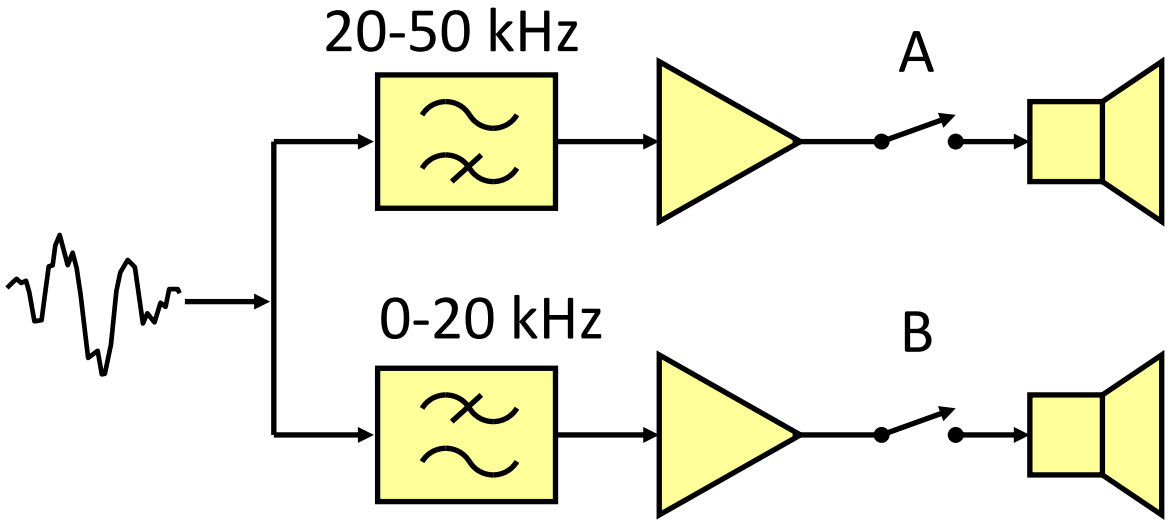
Experimento: Audibilidade de Fase



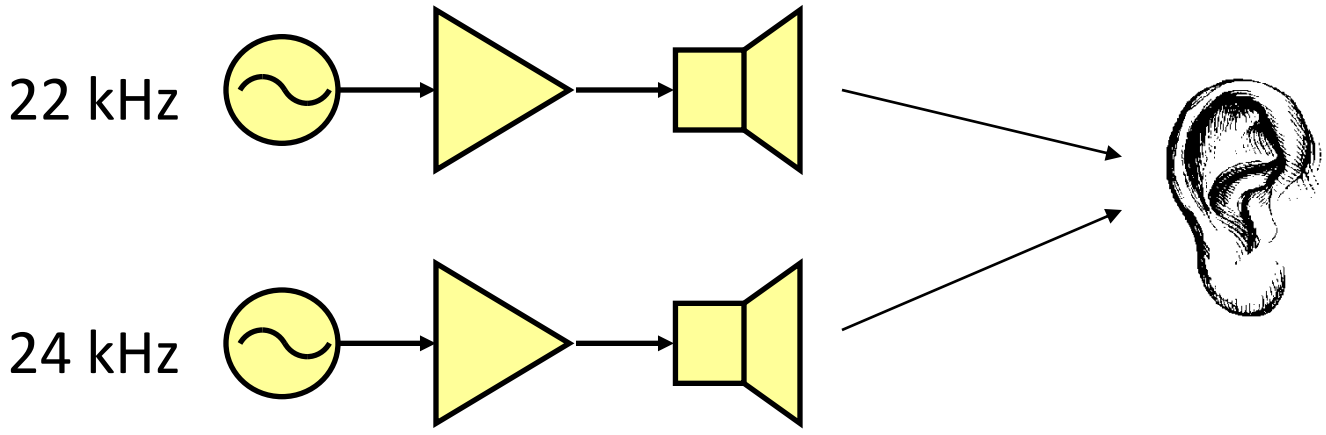
Condições de Audibilidade de Fase



Experimento: Audibilidade Ultrassônica



Condições de Audibilidade Ultrassônica



Audibilidade Ultrassônica: Distorção Aural

22 kHz
+
24 kHz

Após
distorção
aural

2 kHz
audível