

# Córtex Auditivo e Modelagem do Processamento Auditivo

**ESTI019 – Codificação de Sinais Multimídia**

**Profs. Celso Kurashima, Kenji Nose**

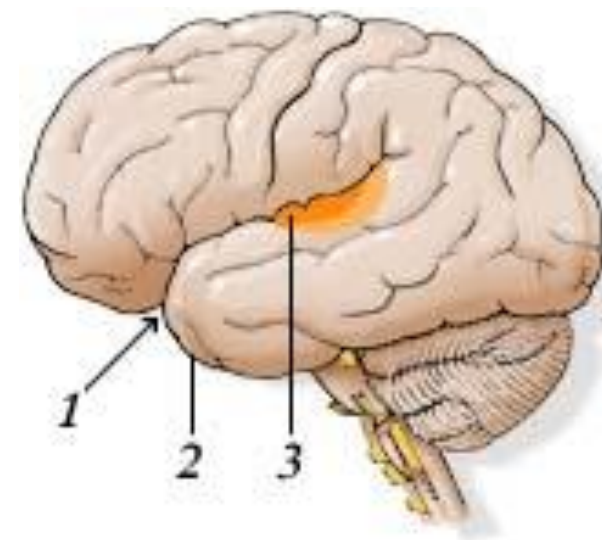
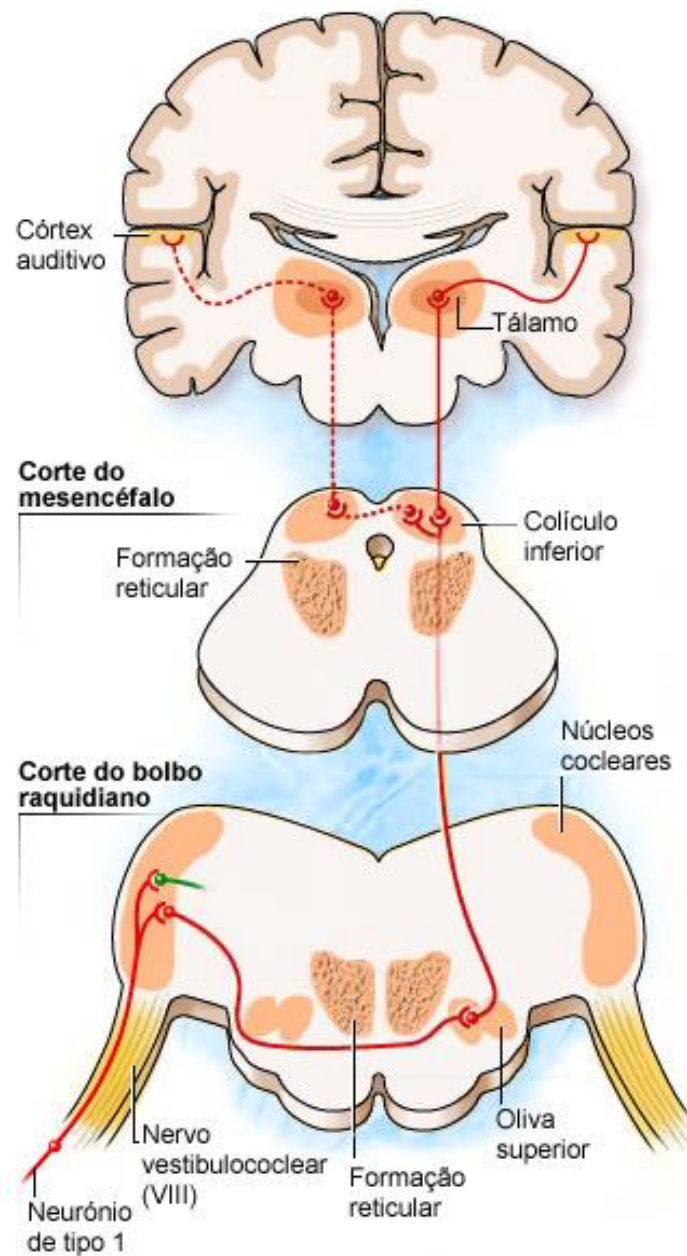
**Mário Minami**



# Organização Tonotópica Paralela

- O sistema auditivo responde setorialmente a tons (frequências) diferentes desde a Cóclea (membrana basilar).
- As fibras neuronais carregam a informação em paralelo, em ambos os hemisférios.
- Altamente redundante para garantir várias vias de transporte da informação até o Córtex Central.

# Caminho Auditivo Ascendente da Cóclea ao Córtex Primário



- 1 – Sulco Lateral
- 2 – Lóbulo Temporal
- 3 – Córtex Auditivo

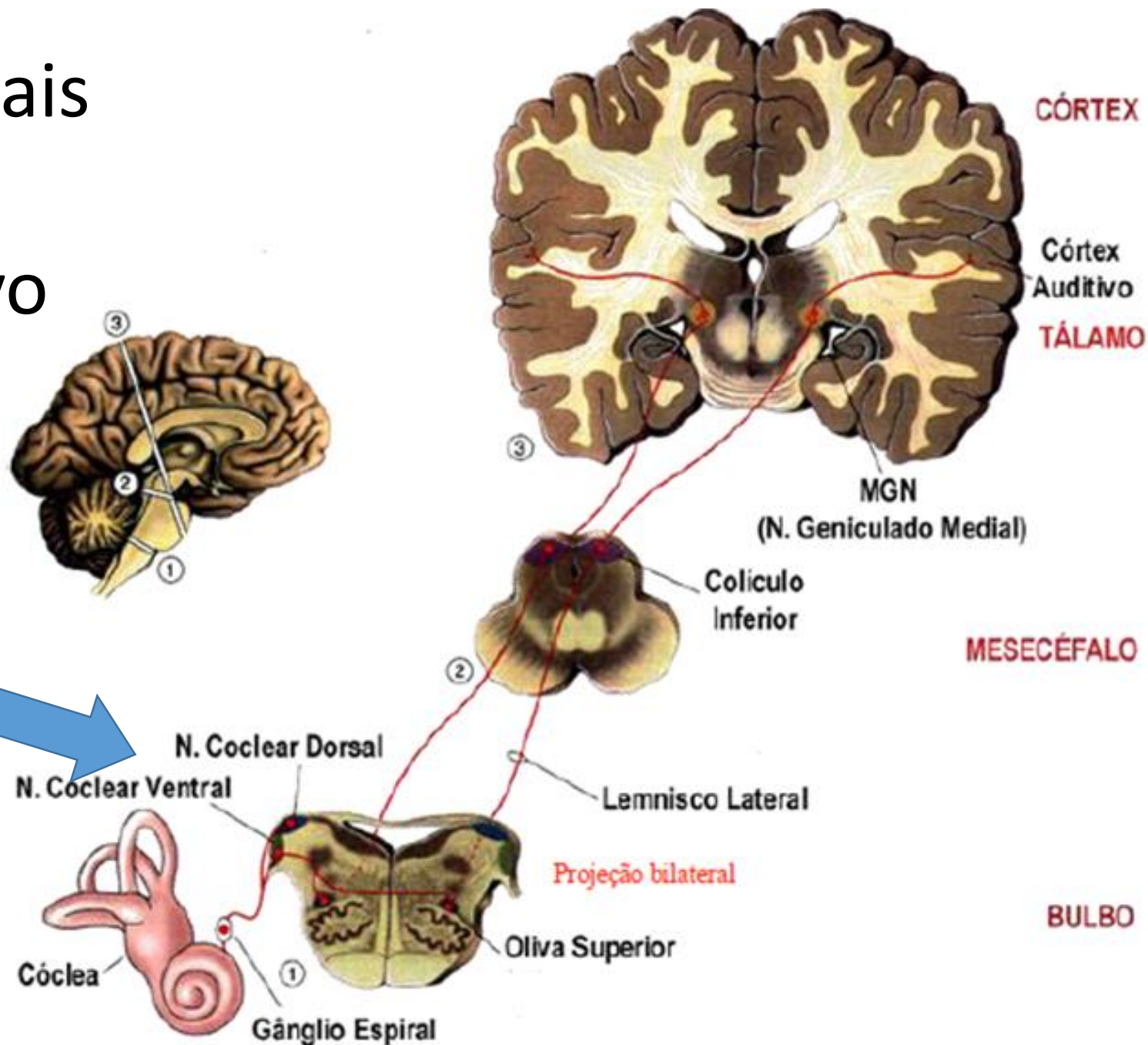
# Núcleos Integradores do Caminho Auditivo

- O primeiro centro integrador é comum à via primária e é formado pelos núcleos cocleares (bolbo raquidiano). Daqui, pequenas fibras convergem para a via reticular ascendente.
- O segundo centro integrador é o complexo olivar superior, também localizado no bolbo raquidiano: a maioria das fibras auditivas fazem sinapse a este nível após cruzarem a linha média
- O terceiro centro integrador é o colículo inferior, localizado no mesoencéfalo. O colículo inferior e o complexo olivar desempenham papel fundamental na localização do som.
- O quarto e último centro integrador, antes do córtex, é o corpo geniculado medial que se localiza-se no tálamo. Neste centro realiza-se importante trabalho de integração: preparação da resposta motora (por exemplo vocal).



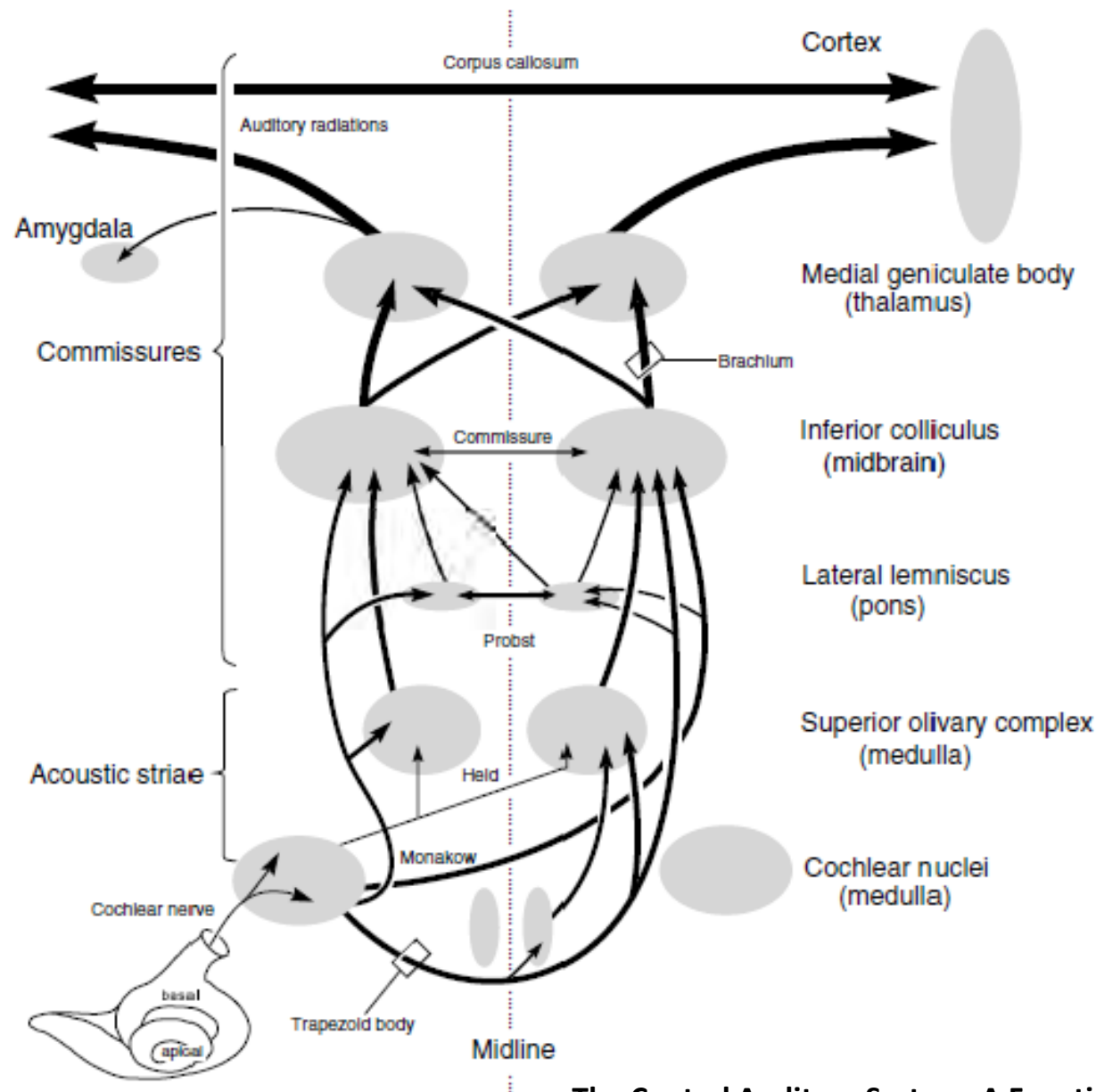
# Cortes Transversais Ilustrando o Caminho Auditivo

Existe um  
“cruzamento” de  
informações de um  
hemisfério ao outro



# Esquema das principais Vias ao Córtex

- vias “estéreo” em cada hemisfério garantem redundância
- Cruzamento Binaural difere na função:
  - **Localização Fonte Sonora**
  - **Integração de Sentidos**
  - **Segmentação de Fonemas**



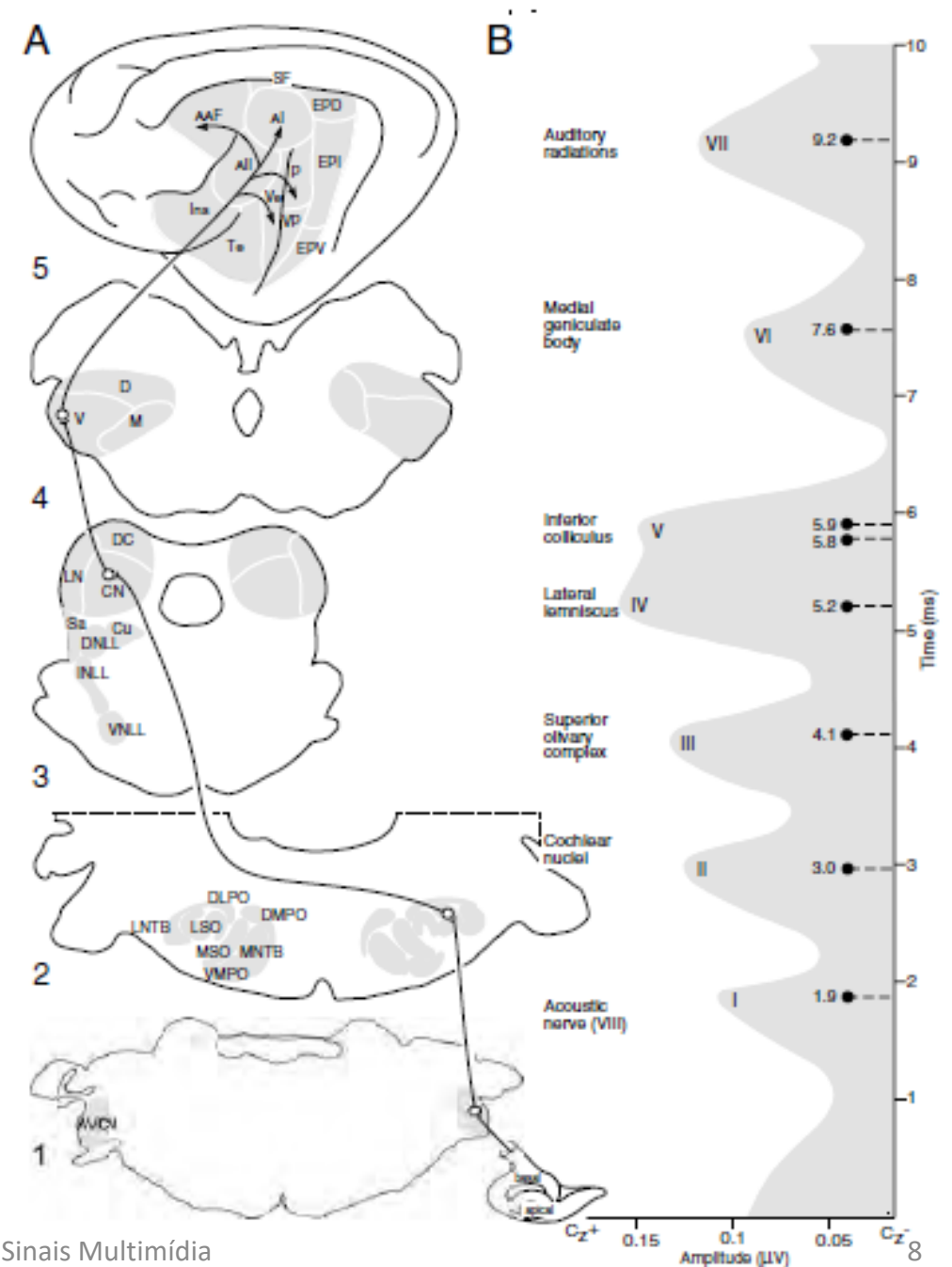
# Funcionalidades das diferentes regiões:

- ***Cochlear nucleus***: origem dos padrões de resposta básicos e paralelismo da informação.
- ***Olivary complex***: construção das vias binaurais e do tempo de resposta.
- ***Lateral lemniscal nuclei***: núcleos de diferenciação bioquímica.
- ***Inferior colliculus***: convergência e integração multisensorial.
- ***Medial geniculate body***: modulação da informação auditiva
- ***Auditory cortex***: interface da audição com as redes cerebrais de comunicação “superiores”

(A) Distribuição dos Diferentes Tipos de Neurônios (vide próx. Slide) e

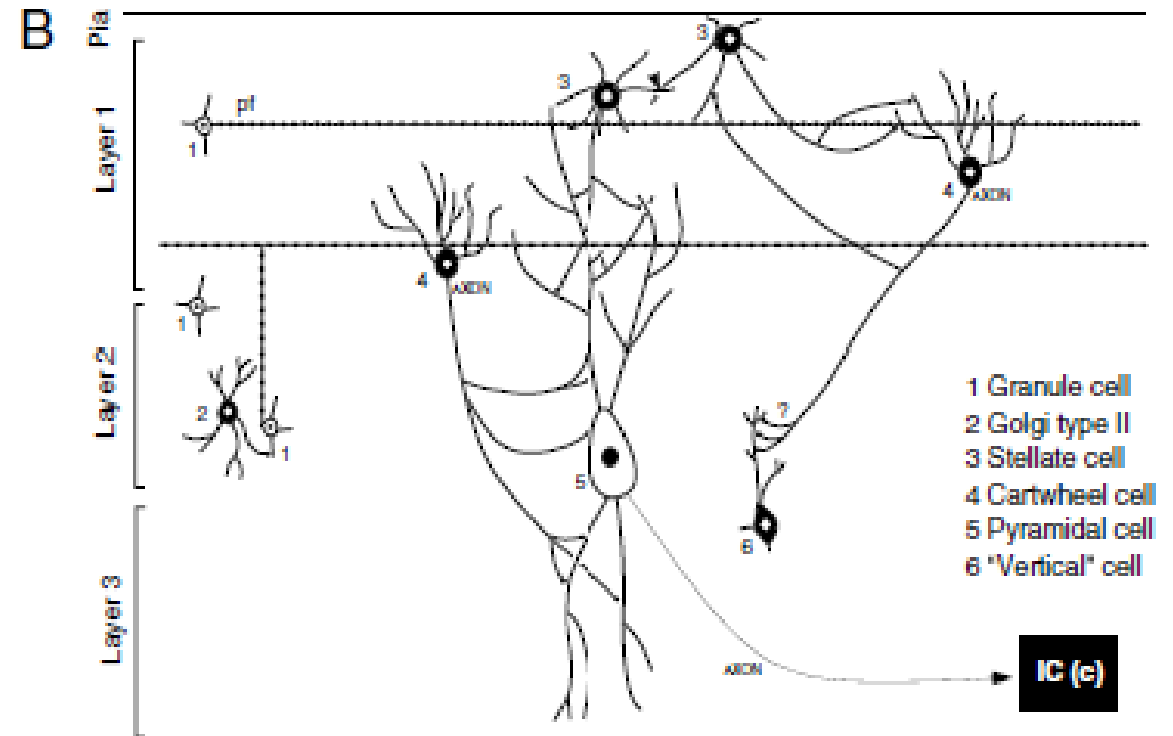
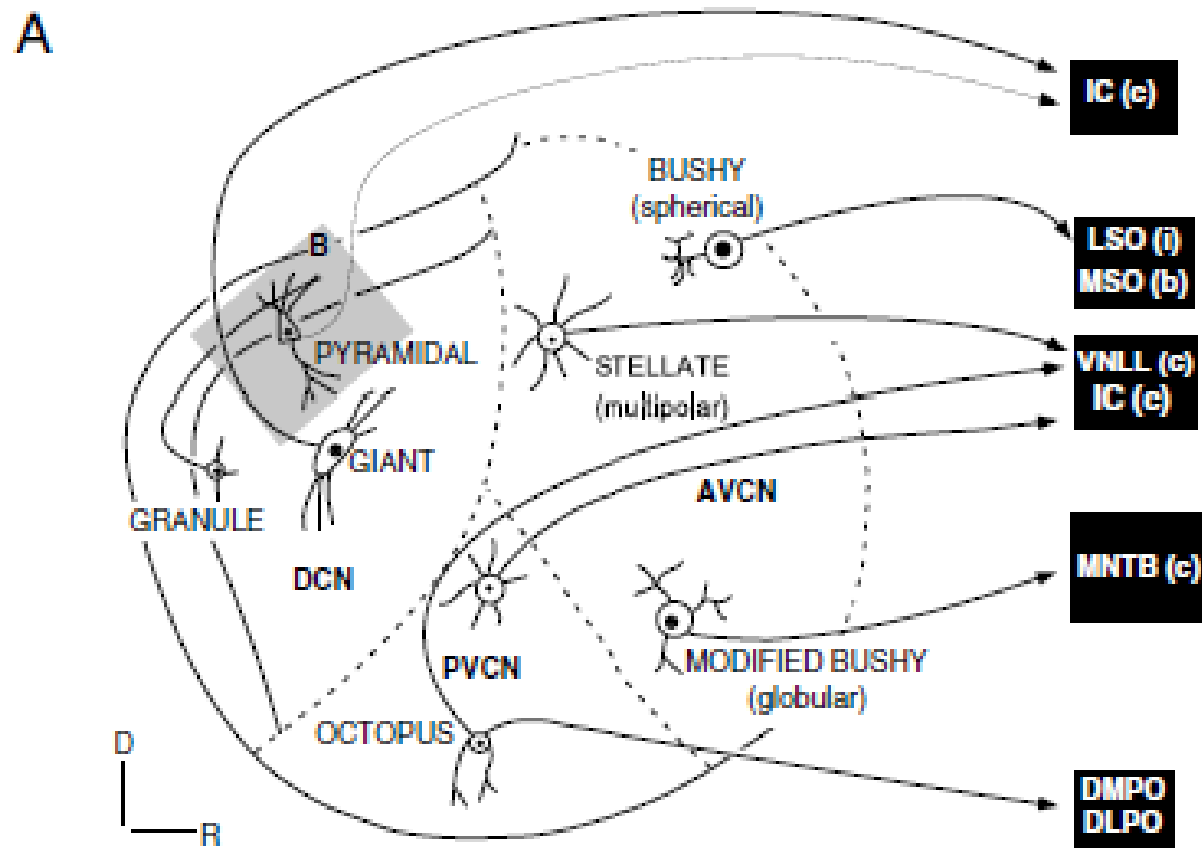
(B) Tempo de Resposta no Caminho Auditivo: só para propagar até o Córtex, cerca de 9,2 ms

**The Central Auditory System: A Functional Analysis**  
Jeffery A. Winer and Christoph E. Schreiner



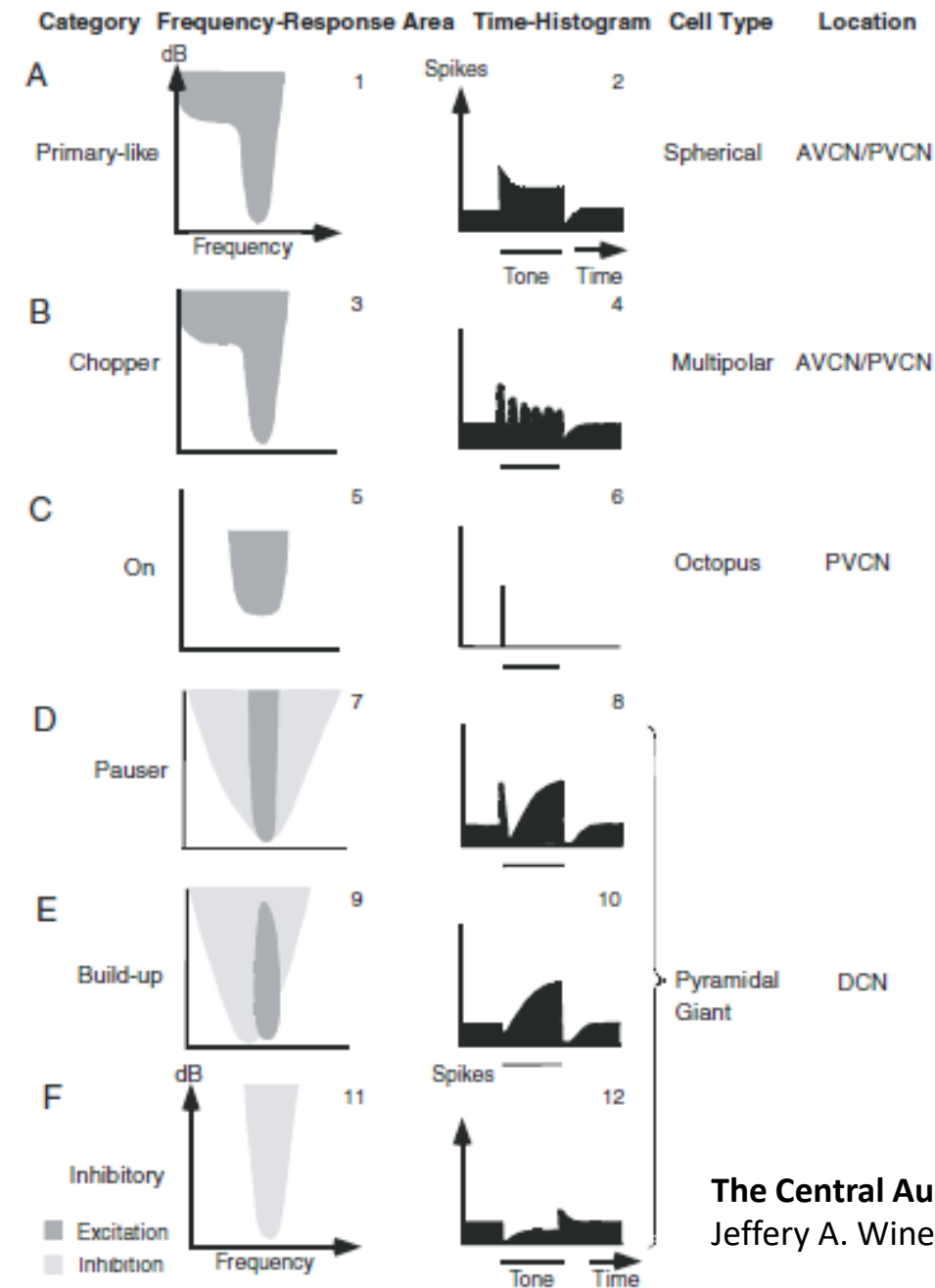


# Neurônios de Diferentes Tipos



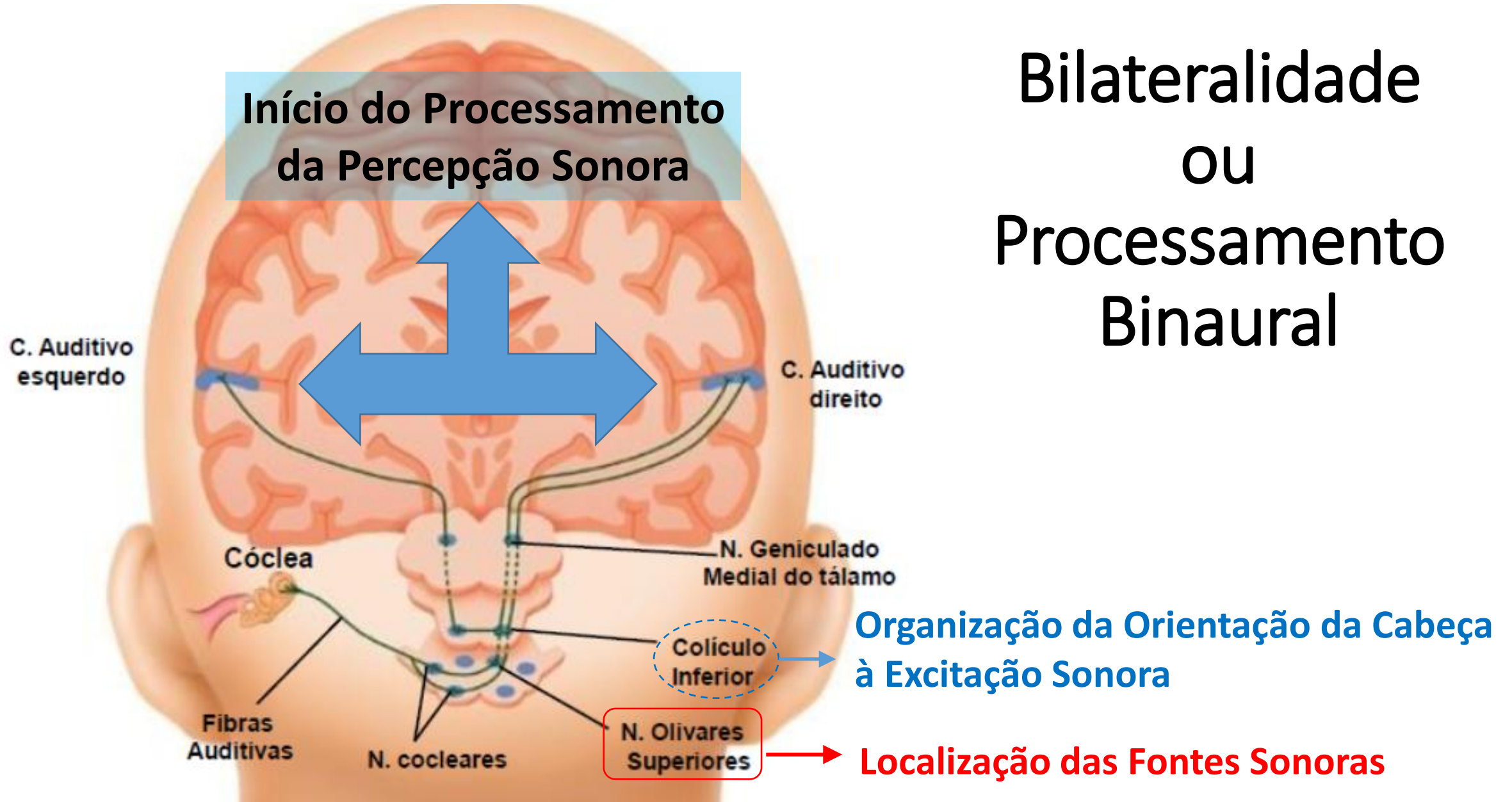
**The Central Auditory System: A Functional Analysis**  
 Jeffery A. Winer and Christoph E. Schreiner

# Exemplos de Respostas Neuronais

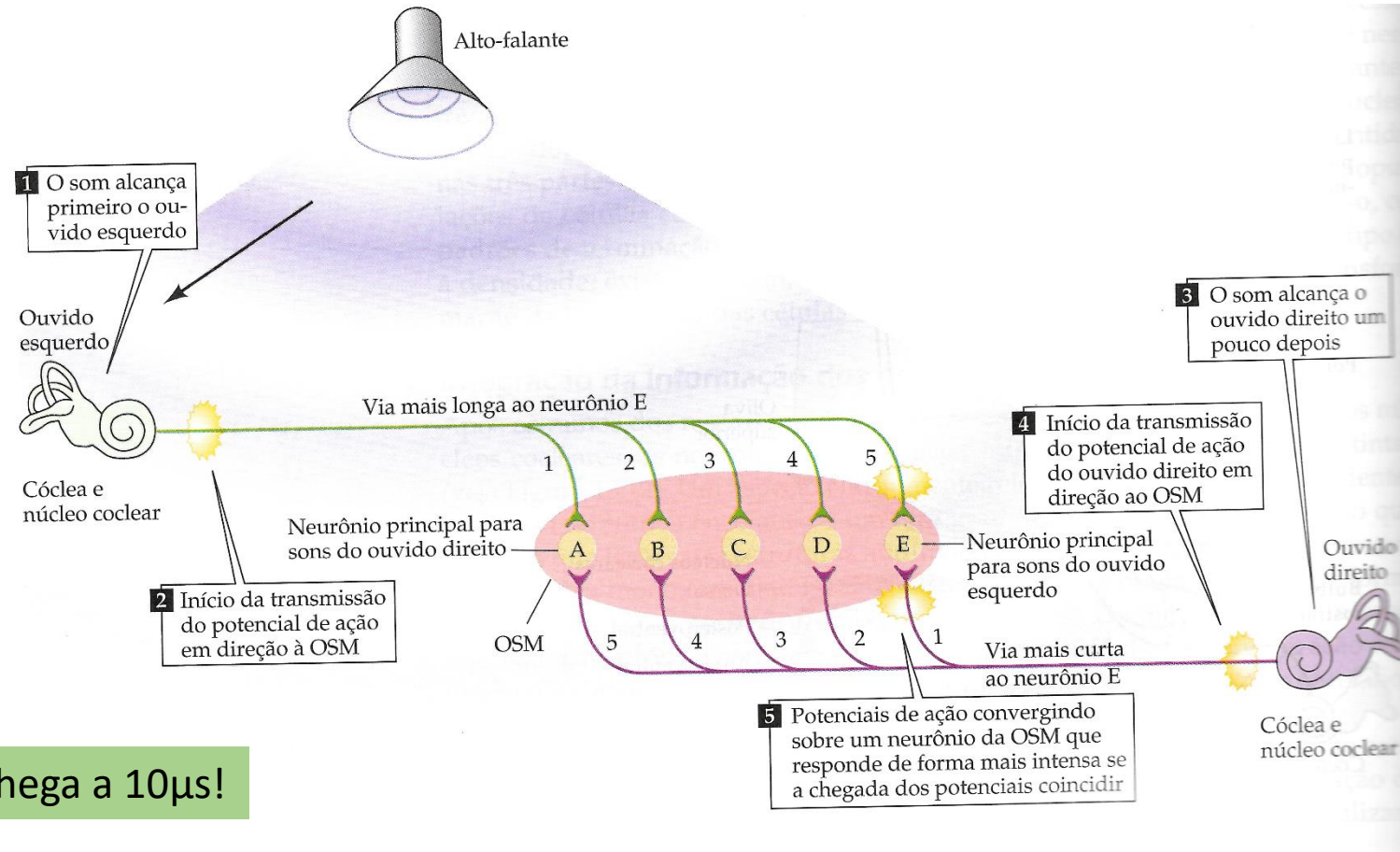


**The Central Auditory System: A Functional Analysis**  
 Jeffery A. Winer and Christoph E. Schreiner

# Bilateralidade OU Processamento Binaural

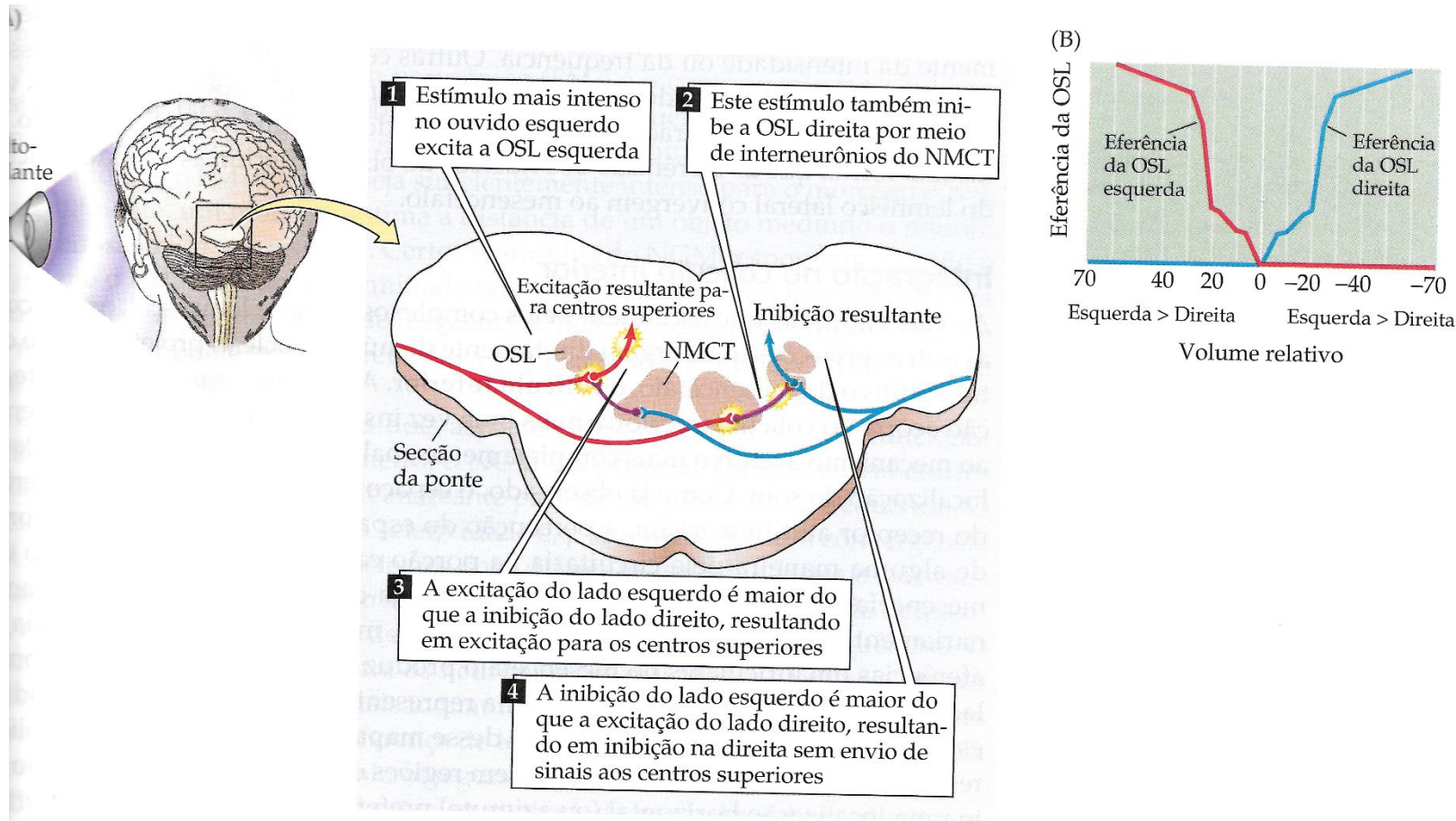


# Exemplo de Processamento da Oliva Superior Medial (OSM)



A resolução chega a  $10\mu\text{s}$ !

# Neurônios da Oliva Superior Lateral (OSL) e o Núcleo Medial do Corpo Trapezóide (NMCT)

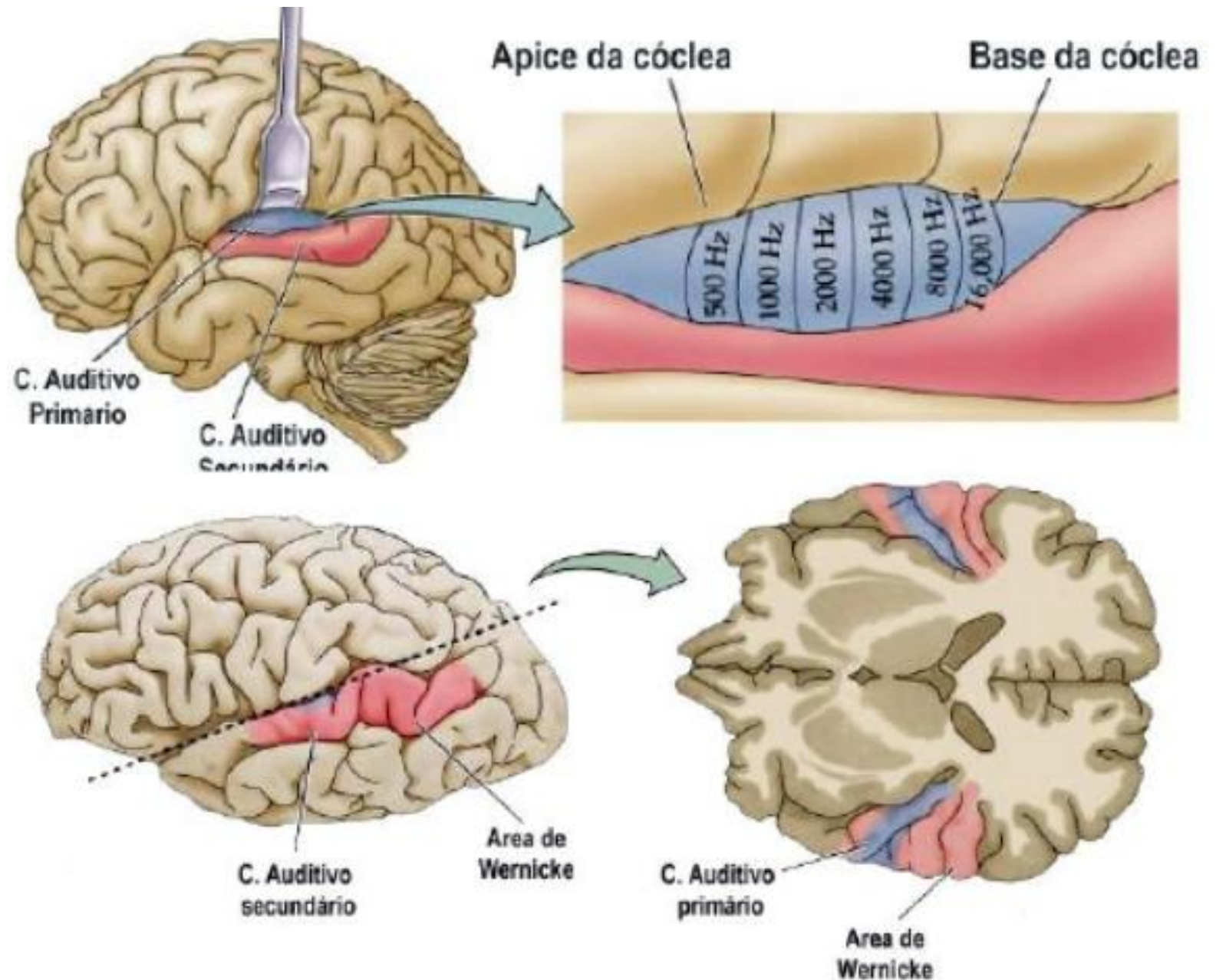




Córtex Primário (A1):  
\* Todos os Sons

Córtex Secundário (A2):  
\* Identificação Fonemas

Área de Wernicke:  
\* Compreensão da Voz  
(hemisfério esquerdo)

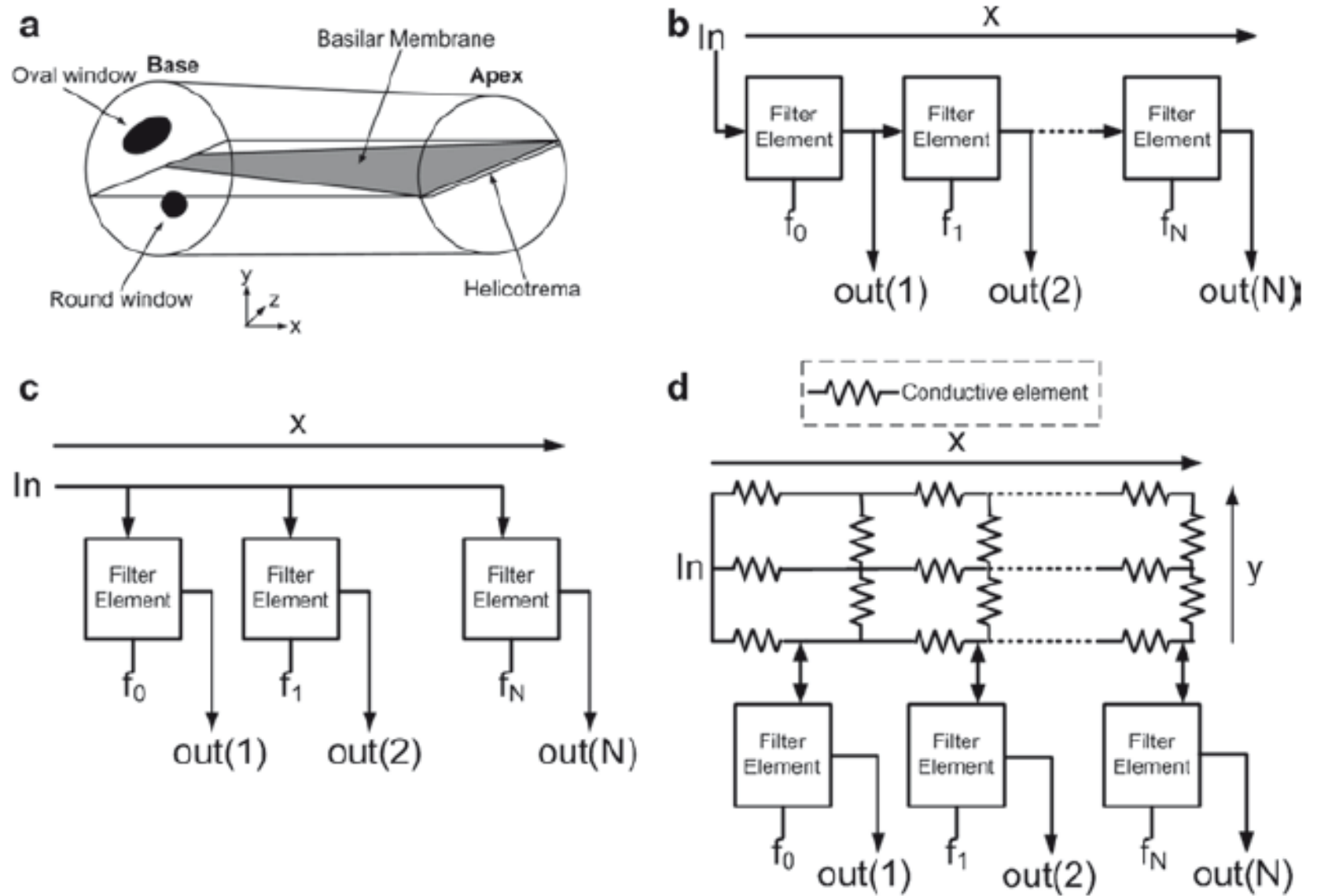


# Modelagem do Processamento Auditivo

- Baseada no funcionamento Bio-Neuronal: Engenharia Neuromórfica
  - Cócleas Artificiais em Silício
  - Tentativa de Reproduzir “exatamente” em algum tipo de circuito o funcionamento do sistema Auditivo
  - Aplicação:
    - Aparelhos de auxílio para degradação das funções auditivas
    - Próteses Artificiais: reabilitação de acidentes
- Processamento que só aproveita características Sistêmicas: apenas tenta reproduzir “processamento efetivo” no Sistema Auditivo

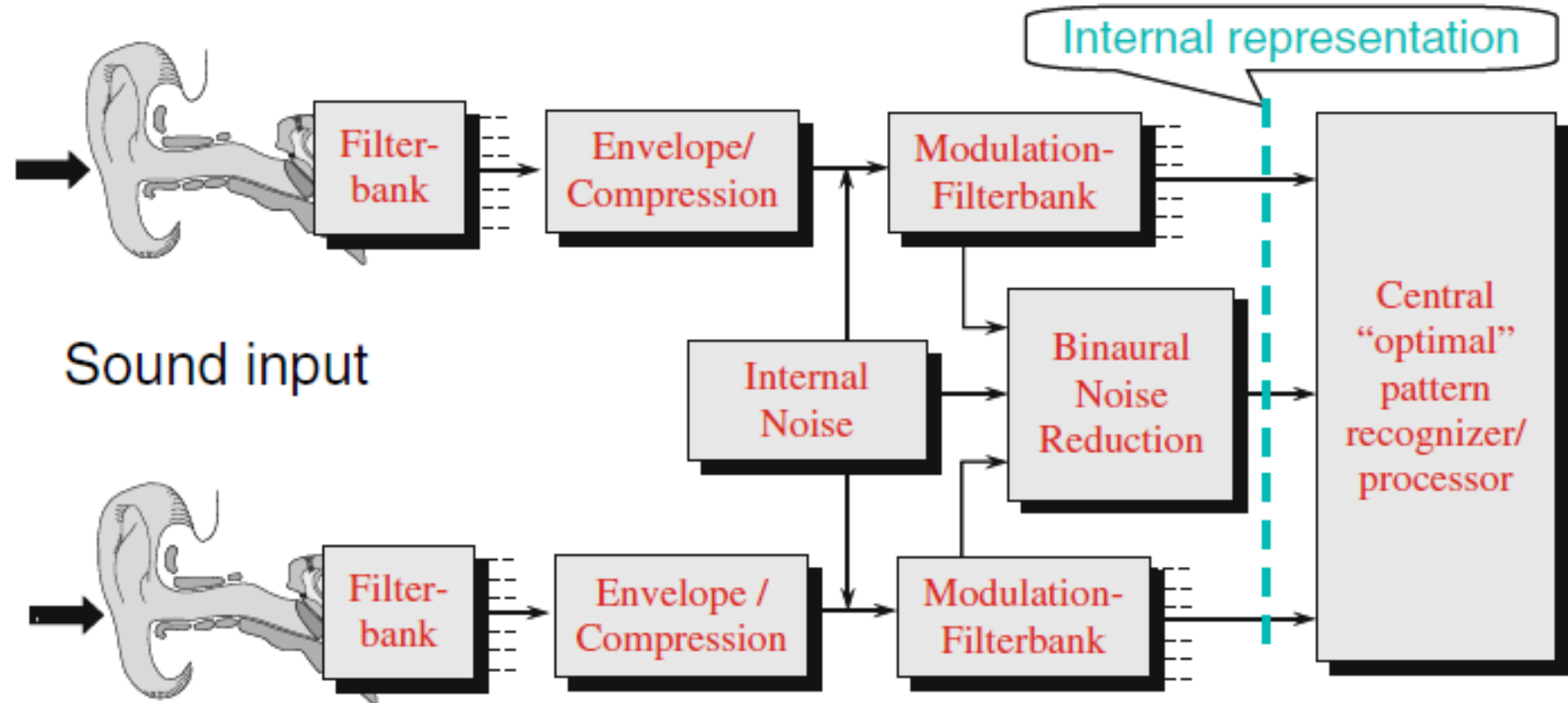
# Modelos de Cóclea em Silício

Aplicações em  
Implantes Auditivos  
e em Engenharia de  
Reabilitação



(a) The uncoiled cochlea. (b) The 1-D cochlea cascade structure. (c) The 1-D parallel structure. (d) The 2-D structure

# Modelo de Processamento “Efetivo” na Audição



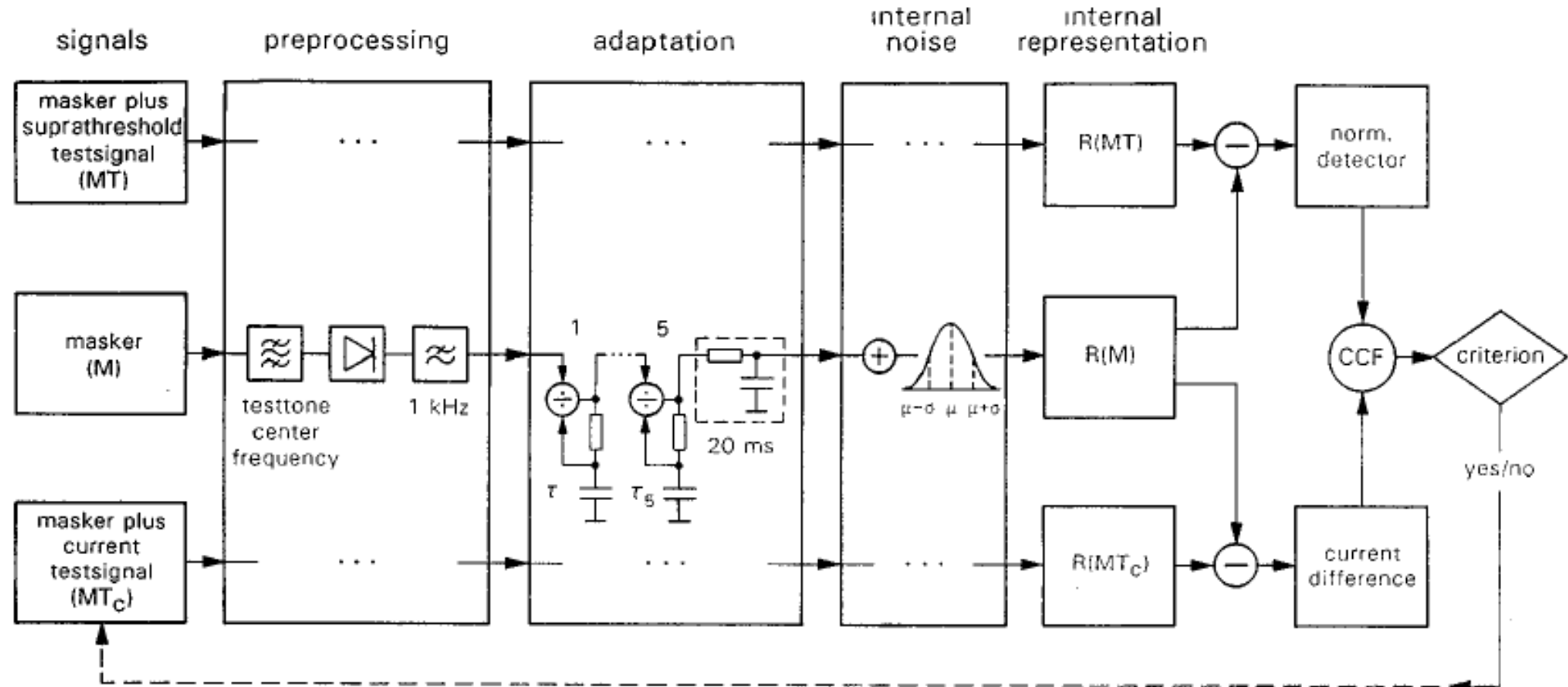
KOLLMEIER, B., *Anatomy, Physiology and Function of the Auditory System*, University of Oldenburg, Oldenburg, Germany.

# Funções Básicas dos Blocos no Modelo

- O Banco de Filtros logo após a Orelha Externa processa a resposta em frequência da Membrana Basilar, com a relação  $Q/\Delta f$  constante e resposta “triangular”.
- O bloco Envelope/Compressão modela a sensibilidade variável em relação ao nível sonoro do sinal de entrada.
- O bloco Ruído Interno modela as diferenças entre os sinais processados e os limiares perceptivos detectados.
- O bloco Modulação/Banco de Filtros modela o processamento de baixas frequências na percepção dos fonemas da voz humana.
- O bloco Redução de Ruído Binaural modela o ganho obtido no processamento estéreo.
- O processador Reconhecedor de Padrões efetua a decisão sobre o sinal de Fala percebido.



# Exemplo de Modelo “Efetivo”

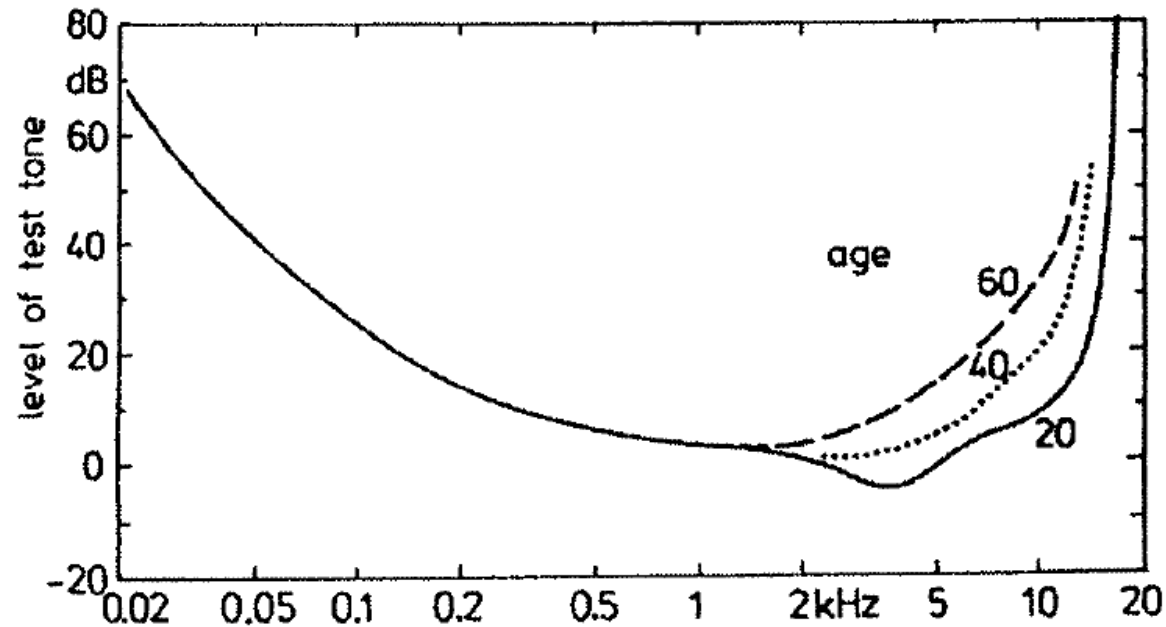


DAU, T.; PÜSCHEL, D.; **A quantitative model of the “effective” signal processing in the auditory system. I. Model structure**, J. Acoust. Soc. Am. **99** (6), June 1996, 3615-22.

# Estágios do Modelo

- Pré-Processamento: Na saída da Membrana Basilar o sinal é retificado em meia onda e filtrado por PB em 1 kHz. O filtro passa baixas preserva a informação na envoltória do sinal (um tipo de demodulação AM).
- Os Efeitos da adaptação são modelados pelos elos de realimentação nas redes de atraso RC (compressão logaritmica do sinal BF e linearização “AF”) que modelam os efeitos de mascaramento temporal.
- O ruído interno gaussiano modela a perda de informação na compresão logarítmica.
- O sinal de Máscara ajuda a estimar os efeitos dos limiares psicoperceptivos para os sinais a serem identificados.

# Limiar e Modelo



$$A(f) / \text{dB} = 3.64(f / \text{kHz})^{-0.8} - 6.5e^{-0.6(f / \text{kHz} - 3.3)^2} + 10^{-3}(f / \text{kHz})^4$$

Terhardt 1979

# Processamento no Córtex Auditivo da Voz Humana

# Importância do Processamento:

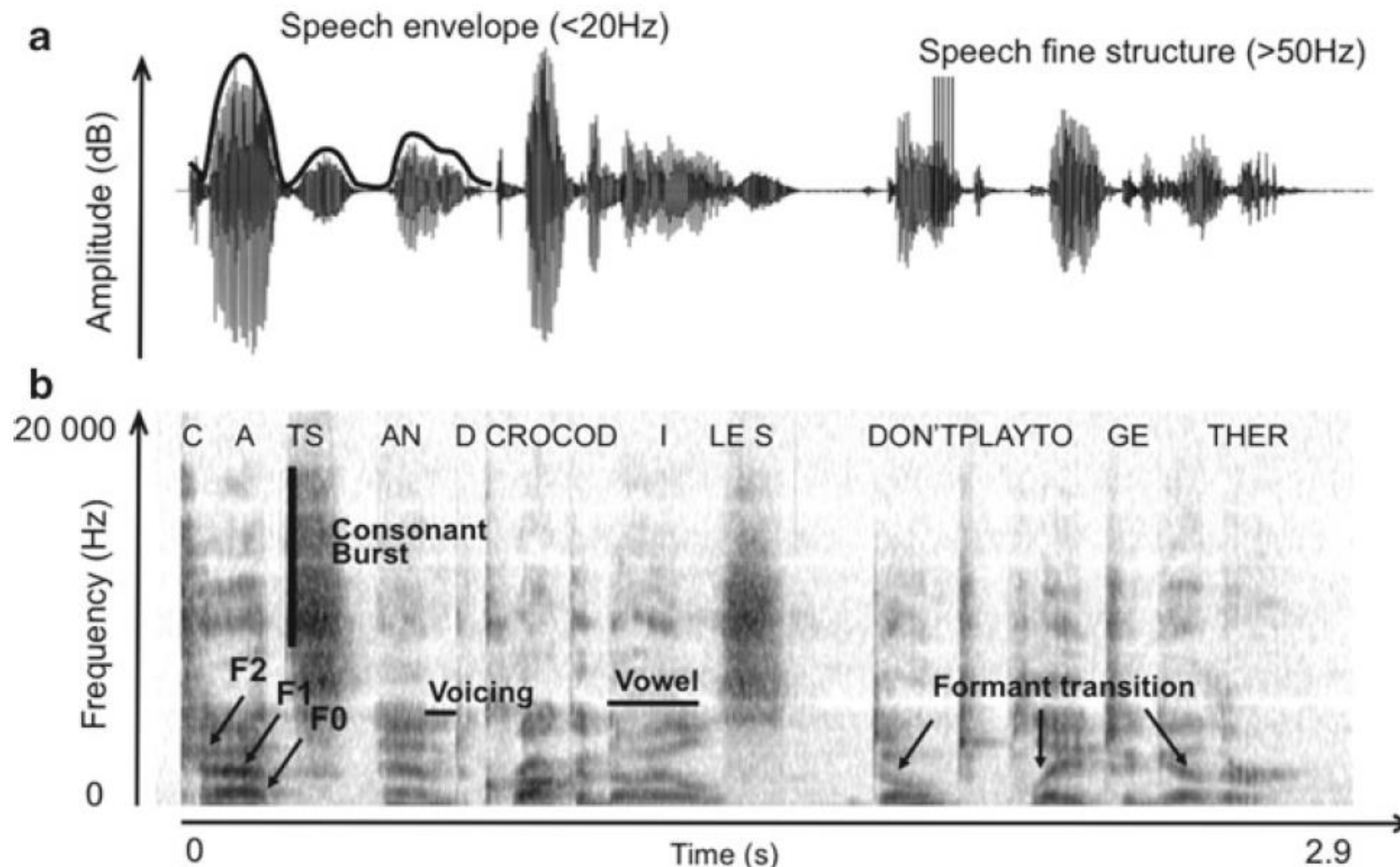
*O papel do Córtex Auditivo não é apenas de representar de forma eficiente a entrada ouvida, mas também e talvez primordialmente, **converter** a estrutura acústica de entrada **num código** que possa ser **buscado (comparado)** com outros tipos de representação importantes na comunicação*



# VOZ

- \* Envelope
- \* Estrutura Fina

- \* F0
- \* Formantes (F1, F2, F3, F4)
- \* Vogal
- \* Consoante
- \* Surdo
- \* Sonoro



GIRAULD, A-L.; POEPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.

## (a) Análise Temporal e (b) Tempo-Frequência (Mista)

(a1) Envelope da Forma de Onda, abaixo de 20 Hz, indica a segmentação (sonora/surda) e pausas (baixa energia entre palavras);

(a2) Acima de 50 Hz a Forma de Onda revela estrutura fina na periodicidade (*pitch*), especialmente na voz sonora (com pregas vocais).

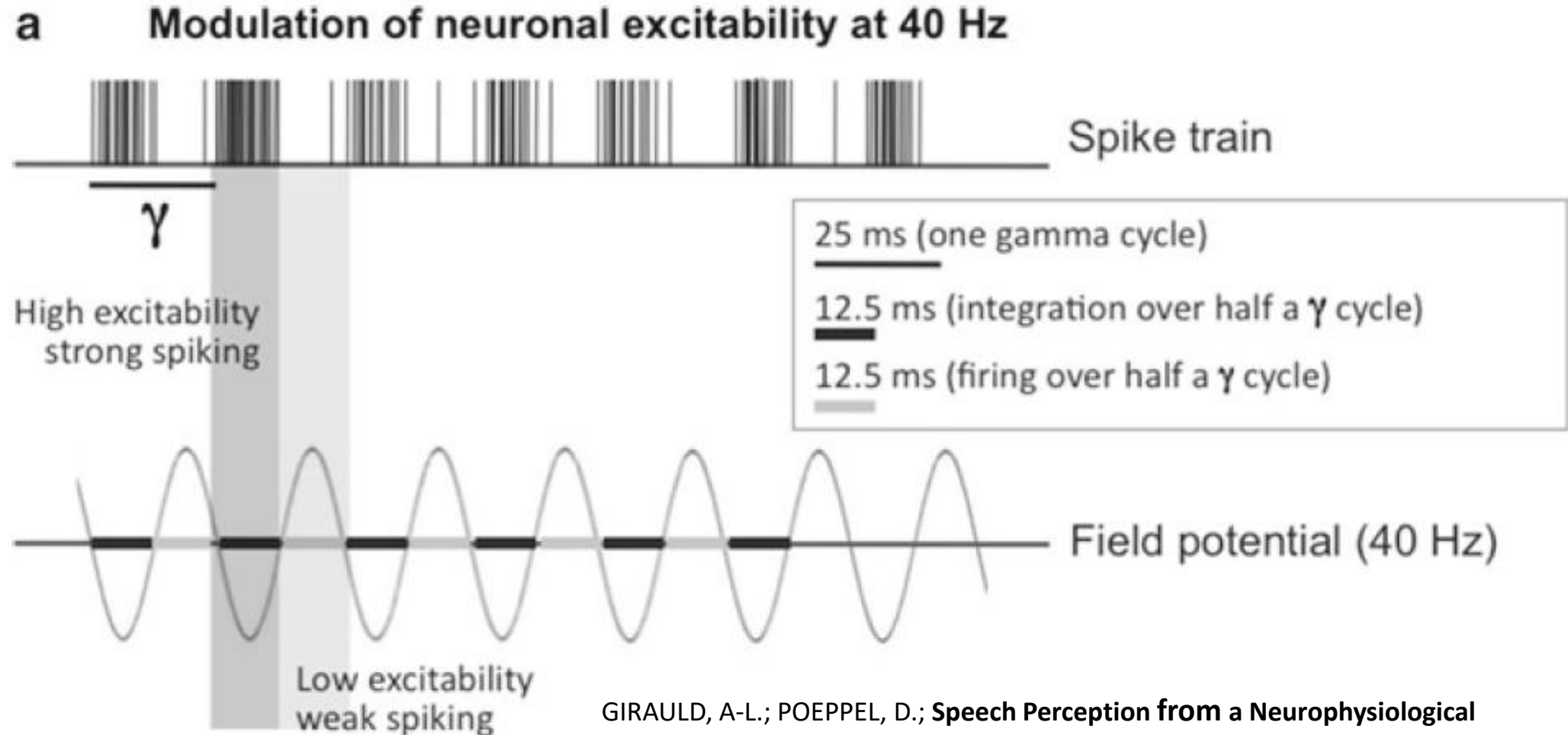
(b1) Espectrograma, ou análise Tempo-Frequência revela a frequência Fundamental (F0) e também a “Linha de Pitch” e as frequências formantes (F1, F2,...);

(b2) Faixa das Consoantes, transição entre as formantes e na prosódia contorno do *pitch*.

# Algumas Modulações Encontradas

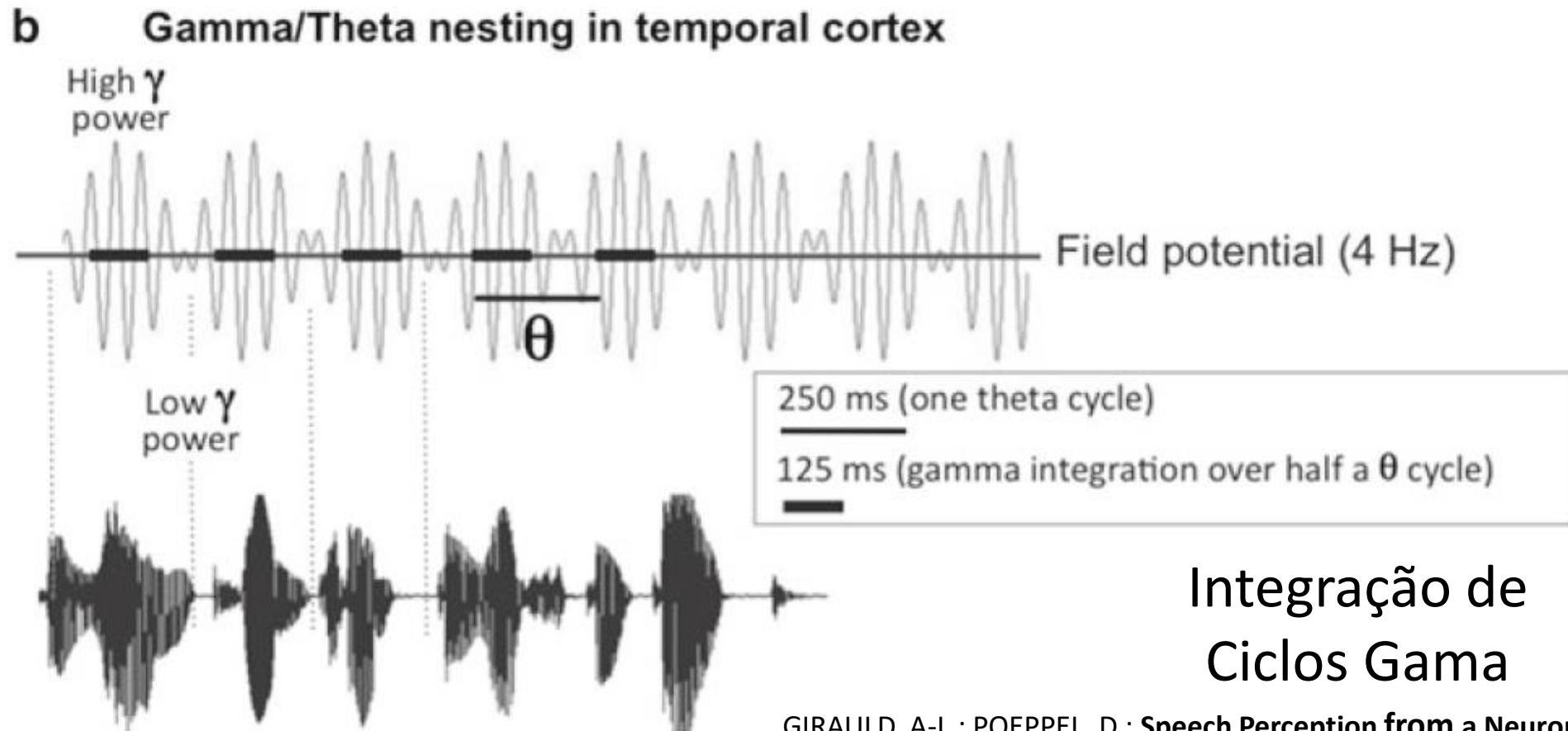
- Amplitude (AM)
- Amplitude rápida (fast AM)
- Frequência (FM)
- Frequência lenta (slow FM)
- Mistas (fast AM – FM), (AM – slow FM)
- Fase (PM)

# Ondas Gama: Modulação 40 Hz, fonemas/sílabas



GIRAULD, A-L.; POEPPPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.

# Ondas Theta: Modulação 4 Hz, palavras



Integração de  
Ciclos Gama

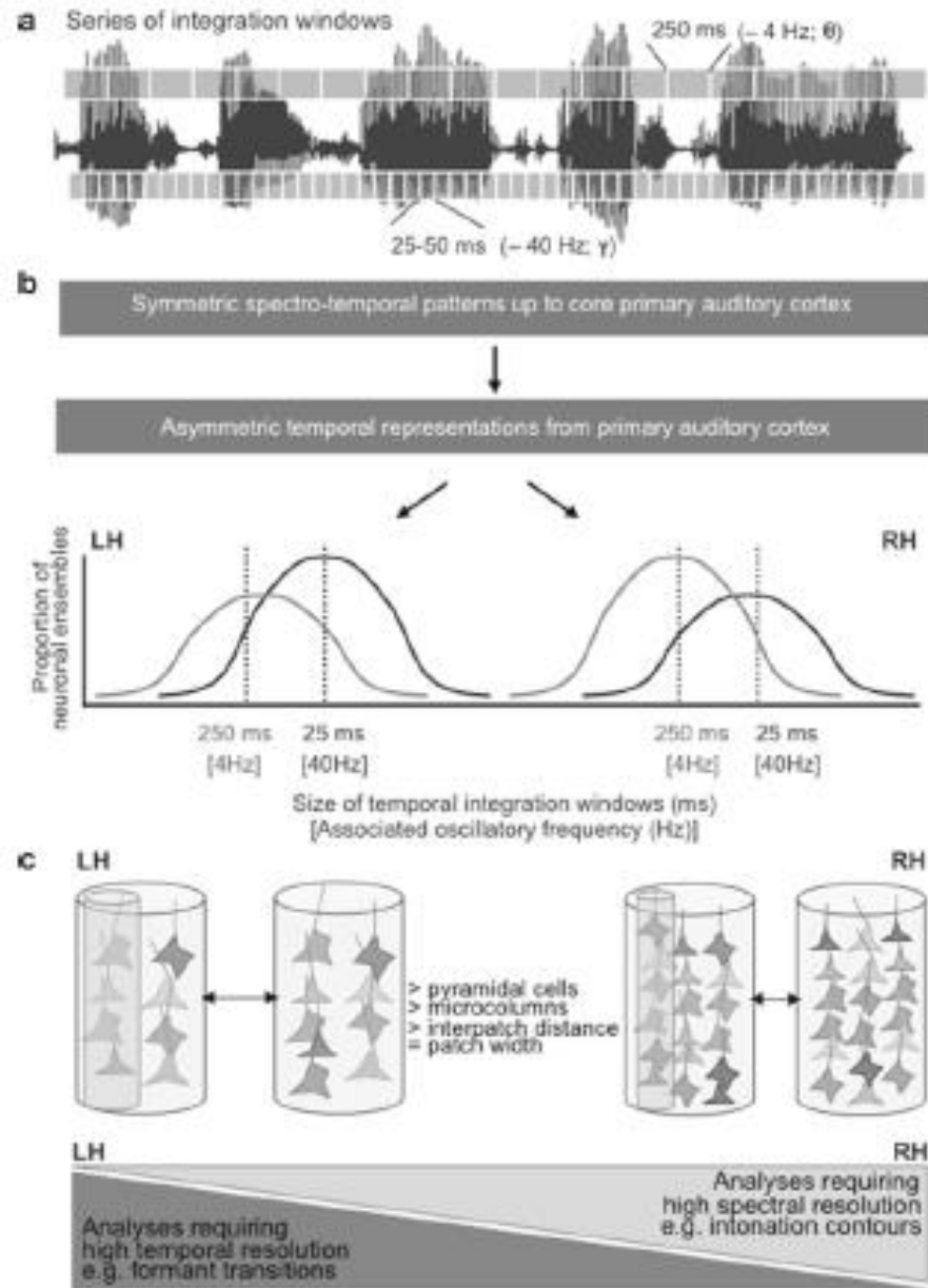
GIRAULD, A-L.; POEPPPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.



# Diferenças Hemisférios Direito e Esquerdo Voz

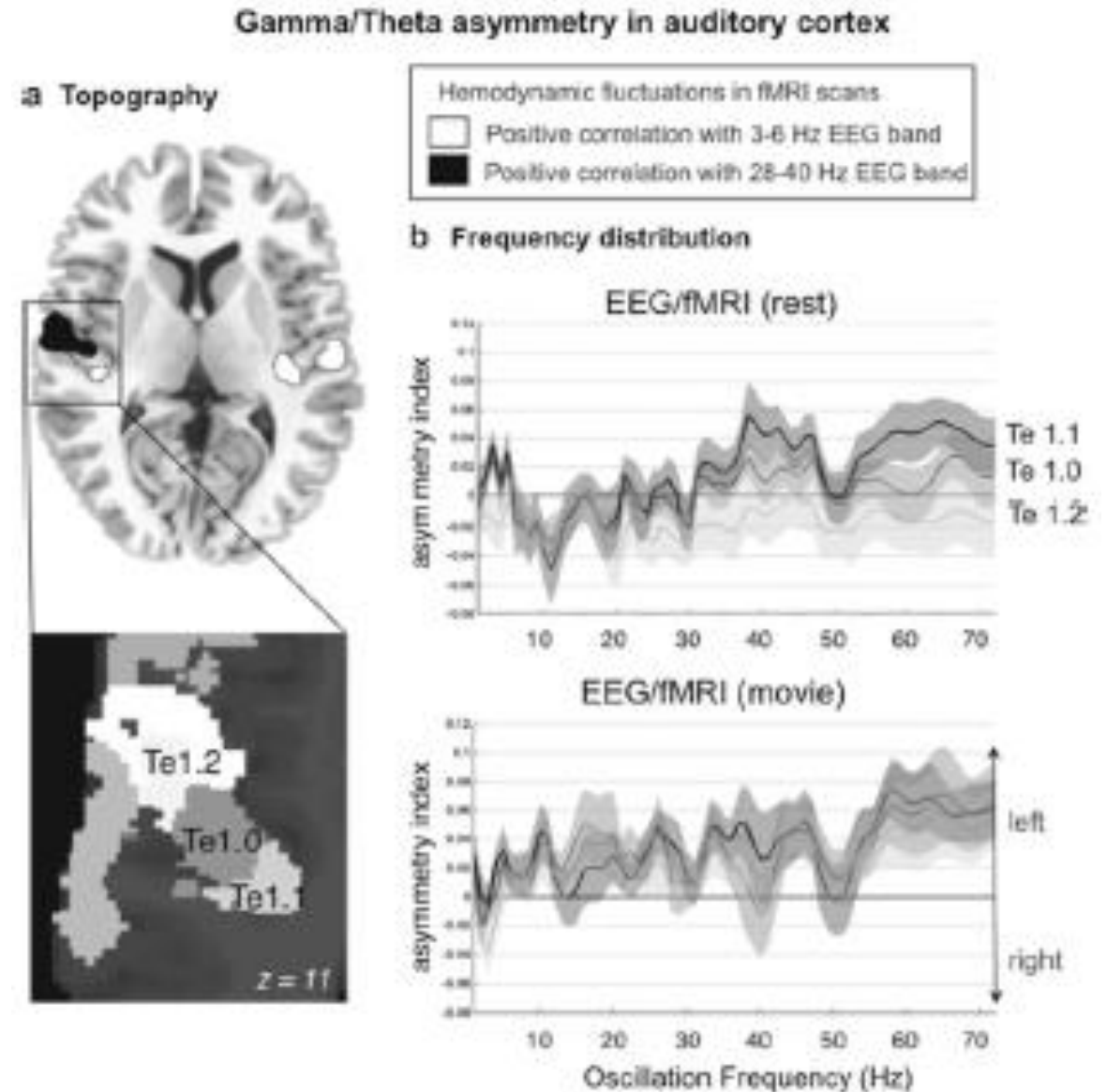
Neurônios  
Piramidais  
concentram-se  
hemisfério Direito

GIRAULD, A-L.; POEPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.



# Assimetria Gama Theta

- Pela fMRI pode-se acompanhar a diferença de incidência de ondas Gama e Theta:
- Maior concentração Gama no hemisfério Esquerdo
- Maior concentração Theta no hemisfério Direito

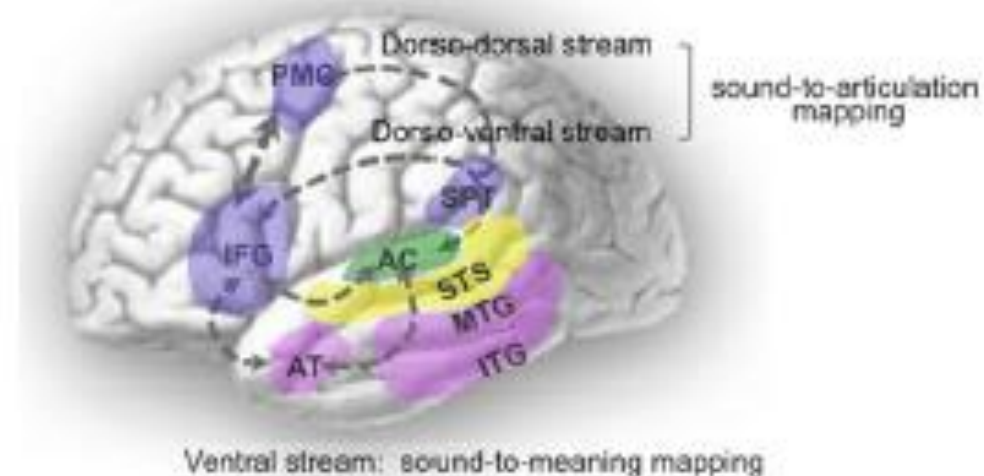


# Modelo Funcional

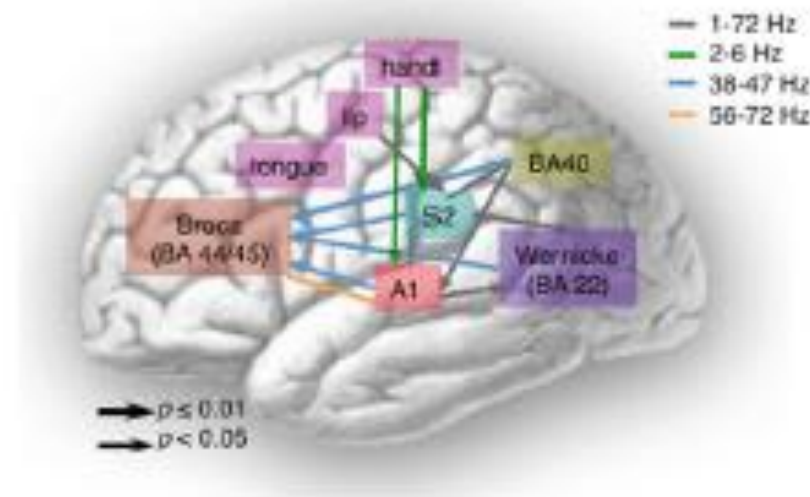
a) Ilustrações de fMRI sobre a dinâmica da informação no Córtex: sequência de regiões solicitadas

a) Ondas em cada região

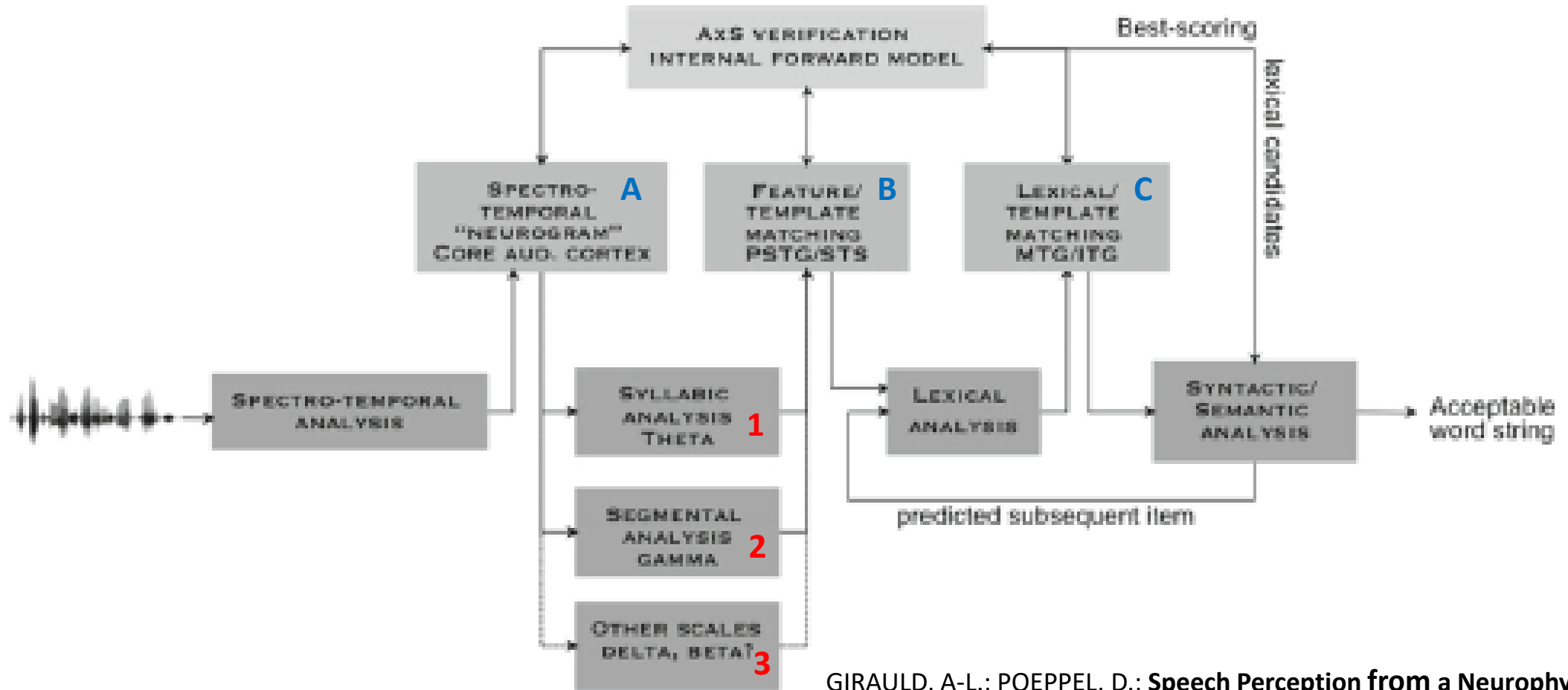
a Dual stream functional neuroanatomical model



b Oscillation-based functional model



# Modelo Análise-por-Síntese da Audição/Fala



GIRAULD, A-L.; POEPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.

# Modelo Análise-Por-Síntese

- **Matching** bidirecional:
  - (i) ascendente (dos fonemas às palavras) e
  - (ii) descendente, padrões prévios de palavras, conhecimento da linguagem, etc.
- Essas hipóteses correspondem a uma “pré-ativação” de representações, tais como padrões acústicos, fonológicos, léxicos, etc.
- Os Blocos (1),(2) e (3) representam possíveis propostas de análises para um reconhecimento satisfatório, de baixo p/ cima (**bottom-up**)
- Os Blocos (A),(B) e (C) representam possíveis propostas cognitivas de um **matching** adequado, de cima p/ baixo (**top-down**)

# Modelos de Processamento em Aplicações

- Mel Cepstrais
- Bancos de Filtros BARK – MPEG
- RASTA Processing

# MEL CEPSTRAIS

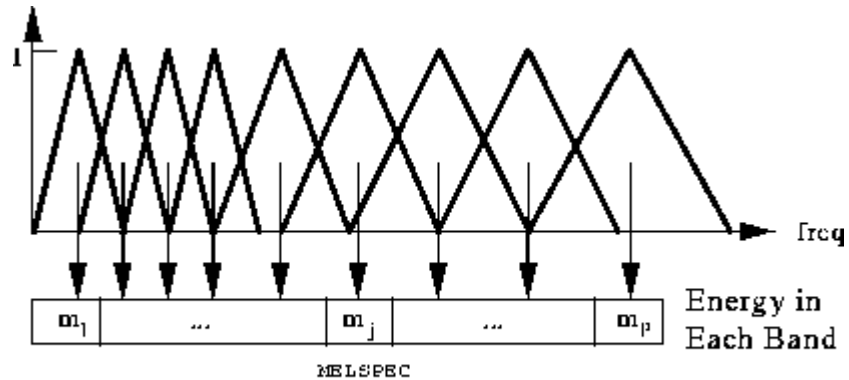
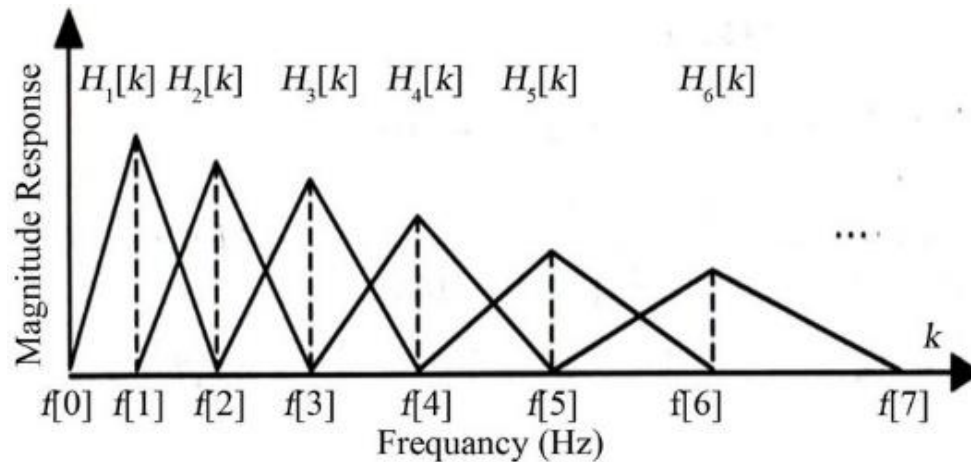


Fig. 5.3 Mel-Scale Filter Bank



- Bancos de Filtro Triangulares
- $f_c$  em escala log (em mels)

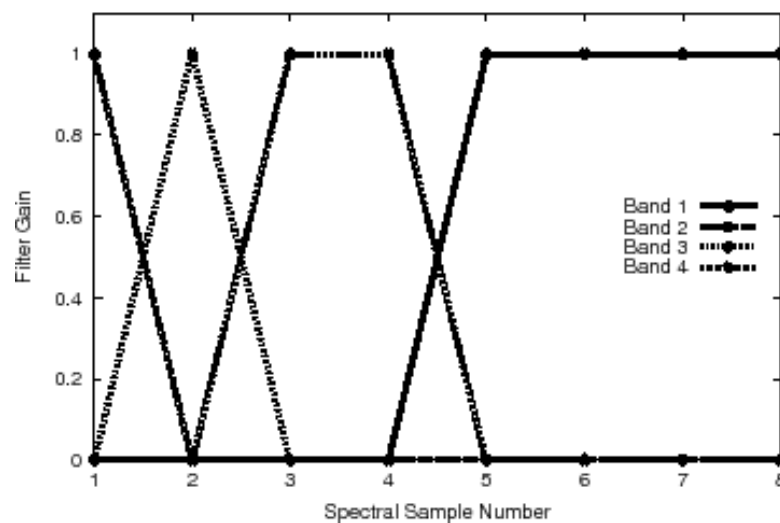
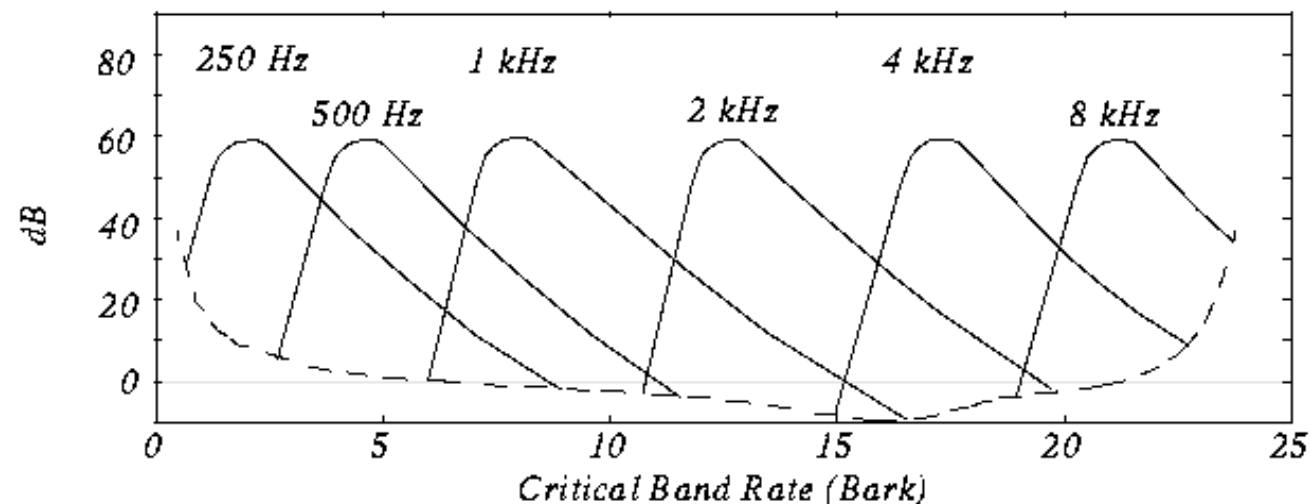
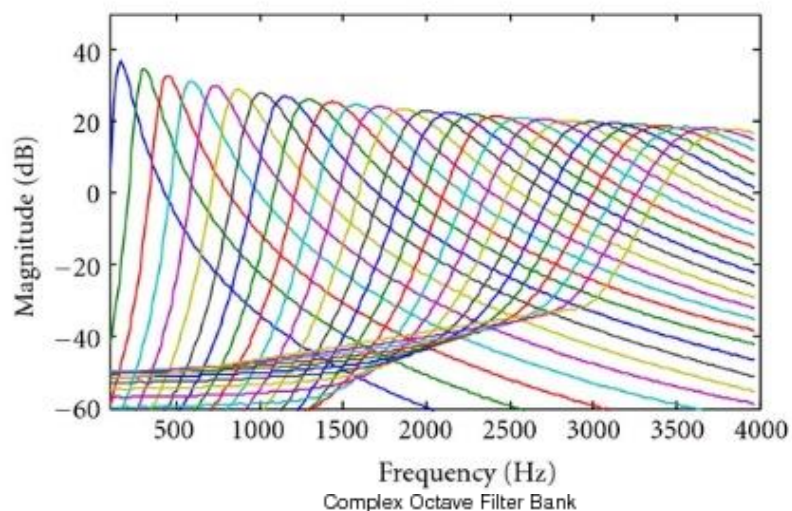
$$m = 2595 \log_{10} \left( 1 + \frac{f}{700} \right) = 1127 \ln \left( 1 + \frac{f}{700} \right)$$

- Relação  $Q/\Delta f$  constante
- Utilizado em sistemas “robustos” a ruído
- Bom desempenho para Voz

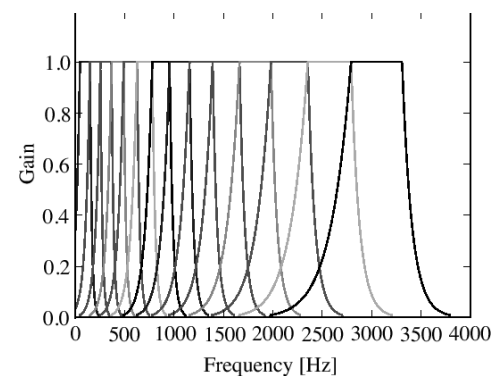


# Escala BARK

- “Shape dos Filtros Variável”,  $f_c$  em escala não linear (Bark)
- Bastante utilizada em Áudio (Graves e Agudos equalizados)



$$\text{Critical band number(Bark)} = \begin{cases} f/100, & \text{for } f < 500 \\ 9 + 4 \log_2(f/1000), & \text{for } f \geq 500 \end{cases}$$



MPEG

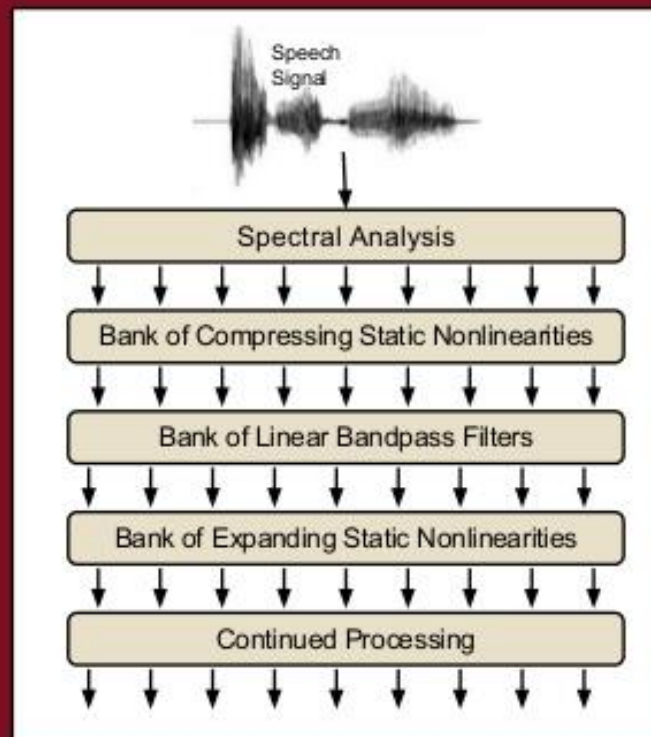
# RASTA Processing

## The Key

→ suppress constant factors in the auditory-like spectrum, prior to estimation of language model.

Research issues:

- What domain is filtering in?
- What filter to use?



- Processamento Altamente Paralelizado
- Redundância para tornar informação mais robusta

## Exercícios 6. Questões para entregar.

- 1) Descreva sumariamente os estágios de processamento no Córtex Auditivo Humano.
- 2) Como o cérebro humano consegue perceber som “espacial”?
- 3) Cite e descreva algumas regiões funcionais no Córtex Auditivo Humano.
- 4) O que é a análise da envoltória do som?
- 5) Dê exemplo de análise tempo-frequência sonora.
- 6) Qual a diferença entre as ondas Gama e Theta na análise da fala?
- 7) Justifique o termo análise-por-síntese na audição/fala humana.
- 8) Faça um plot comparativo entre as escalas mel e bark.

# Referências:

- BOSI, M. and GOLDBERG, R. E., **Introduction to Digital Coding and Standards**, Kluwe Academic Publishers, Hong Kong, 2003.
- DUTOIT, T. and MARQUES, F., **Applied Signal Processing – A Matlab Based Proof of Concept**, Springer, New York, 2009.
- PRESSNITZER, D. et al (Eds), **Auditory Signal Processing – Physiology, Psychoacoustics and Models**, Springer, Paris, 2004.
- YOU, Y., **Audio Coding – Theory and Applications**, Springer, Minnesota, 2010.
- KOLLMEIER, B., **Anatomy, Physiology and Function of Auditory System**, in: Handbook of Signal Processing in Acoustics, Havelock, D. et al (Eds), Chapter 10, Springer, New York, 2008.
- DAU, T., **Auditory Processing Models**, in: Handbook of Signal Processing in Acoustics, Havelock, D. et al (Eds), Chapter 10, Springer, New York, 2008.
- GIRAULD, A-L.; POEPPEL, D.; **Speech Perception from a Neurophysiological Perspective**, In: The Human Auditory Cortex, Ch.9, Springer, 2012.
- Douglas O'Shaughnessy (1987). [\*Speech communication: human and machine\*](#). Addison-Wesley. p. 150. [ISBN 978-0-201-16520-3](#).
- BEAR, M. F., CONNERS, B. W. e PARADISO, M. A., **Neurociências: Desvendando o Sistema Nervoso**, Artmed, São Paulo, 2017.