

---

**NOVAS TÉCNICAS DE  
PROCESSAMENTO DIGITAL PARA  
AJUDA A PORTADORES DE  
DEFICIÊNCIA AUDITIVA COM PERDAS  
EM ALTAS FREQUÊNCIAS**

**ALAN MENDES MAROTTA**

Dissertação apresentada ao Instituto Nacional de Telecomunicações, como parte dos requisitos para obtenção do Título de Mestre em Engenharia Elétrica.

**ORIENTADOR:**

Prof. Dr. Francisco José Fraga da Silva

Santa Rita do Sapucaí  
2004

---

## FOLHA DE APROVAÇÃO

Dissertação defendida e aprovada em \_\_\_\_ / \_\_\_\_ / \_\_\_\_ ,  
pela comissão julgadora:

---

(Prof. Dr. Francisco José Fraga da Silva/INATEL)

---

(Prof. Dr. Carlos Alberto Ynoguti/INATEL)

---

(Prof<sup>a</sup>. Dra. Kátia Miriam de Melo Silveira/UNIFENAS)

---

Prof. Dr. Adonias C. da Silveira  
**Coordenador do Curso de Mestrado**

---

Ao meu pai e minha mãe,  
meus primeiros e grandes professores.

---

## AGRADECIMENTOS

Ao meu filho Cauan, por me trazer um novo sentido à vida.

À minha esposa Nayara, pela força e companheirismo nessa empreitada.

Ao prof. Wander, a quem devo muito pela ajuda em momentos difíceis e também pelo incentivo e pela força que foram fundamentais para eu trilhar o caminho acadêmico.

À prof<sup>a</sup> Rosana, que me orientou durante quatro anos da graduação nos trabalhos com o laboratório de computação, pelas oportunidades e pelos incentivos que muito ajudaram.

Ao Professor Fraga, pela paciência e dedicação despendida durante a elaboração deste trabalho.

À Dra. Kátia Miriam pelo estímulo e ajuda prestadas.

A todos os colegas, professores e funcionários do Departamento de Telecomunicações do Instituto Nacional de Telecomunicações pela amizade.

---

## ÍNDICE

|  |     |
|--|-----|
| LISTA DE FIGURAS .....                                 | i   |
| LISTA DE TABELAS.....                                  | ii  |
| LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS .....         | iii |
| RESUMO .....   | iv  |
| ABSTRACT .....   | v   |
| CAPÍTULO I – Introdução .....                          | 01  |
| CAPÍTULO II – Audiologia – fundamentos acústicos ..... | 04  |
| II.1 Acústica .....                                    | 04  |
| II.1.1 – Ondas Sonoras .....                           | 05  |
| II.1.2 – Tipos de ondas sonoras .....                  | 06  |
| a) Quanto à propagação da onda no espaço .....         | 06  |
| b) Quanto à complexidade .....                         | 06  |
| c) Quanto à periodicidade .....                        | 07  |
| II.1.3 Dimensões da onda sonora .....                  | 08  |
| a) Frequência .....                                    | 08  |
| b) Frequência angular .....                            | 09  |
| c) Fase inicial .....                                  | 09  |
| d) Amplitude .....                                     | 09  |
| e) Comprimento de onda .....                           | 10  |

|   |    |
|---|----|
| II.1.4 – “Qualidades” da onda sonora – altura e intensidade ..... | 11 |
| II.1.5 – Níveis de audição .....                                  | 12 |
| II.1.6 – Decibel Nível de Audição (dB NA) .....                   | 14 |
| CAPÍTULO III – Anatomia e funcionamento do sistema auditivo ..... | 17 |
| III.1 – Introdução .....  | 17 |
| III.2 – Ouvido externo – energia sonora .....                     | 17 |
| III.3 – Ouvido médio – energia mecânica .....                     | 19 |
| III.4 – Ouvido externo – energia sonora .....                     | 20 |
| CAPÍTULO IV – Psicoacústica .....                                 | 23 |
| IV.1 – Métodos Psicofísicos .....                                 | 24 |
| IV.2 – Percepção da intensidade do som .....                      | 25 |
| IV.3 – Percepção da altura (ou frequência) .....                  | 26 |
| IV.4 – Percepção das Variações Temporais .....                    | 26 |
| CAPÍTULO V – Avaliação básica da audição .....                    | 27 |
| V.1 – Introdução .....  | 27 |
| V.2 – Anamnese .....  | 28 |
| V.3 – Inspeção do meato acústico .....                            | 28 |
| V.4 – Pesquisa dos limites tonais por via aérea .....             | 29 |
| V.5 – Audiograma .....  | 31 |
| V.6 – A importância da calibração .....                           | 32 |
| CAPÍTULO VI – Foco da pesquisa .....                              | 33 |
| CAPÍTULO VII – Métodos de rebaixamento de frequências .....       | 37 |
| VII.1 – Introdução .....  | 37 |
| VII.2 – Esquema de ajuda digital .....                            | 38 |

|   |    |
|---|----|
| VII.3 – Métodos de compressão de frequências .....                | 39 |
| VII.3.1 – Compressão linear de frequências .....                  | 39 |
| VII.3.2 – Compressão não linear de frequências .....              | 39 |
| VII.3.3 – Método de compressão de frequências implementado .....  | 41 |
| VII.4 – Método de transladação de frequências proposto .....      | 42 |
| VII.4.1 – Cálculo da “Frequência de Destino” .....                | 42 |
| VII.4.2 – Cálculo da “Frequência de Origem” .....                 | 43 |
| CAPÍTULO VIII – Esquema de processamento implementado .....       | 46 |
| VIII.1 –Processamento quadro-a-quadro .....                       | 46 |
| VIII.2 – Janelamento Hamming .....                                | 47 |
| VIII.3 – Critério discriminativo de processamento do quadro ..... | 48 |
| VIII.4 – Algoritmo implementado .....                             | 49 |
| VIII.5 – Visualização dos resultados do sinal processado .....    | 52 |
| CAPÍTULO IX – Testes de inteligibilidade .....                    | 54 |
| IX.1 – Avaliação Discriminativa .....                             | 54 |
| IX.2 – Discussão dos resultados .....                             | 61 |
| IX.3 Matriz de confusão .....                                     | 62 |
| CAPÍTULO X – Conclusões .....                                     | 64 |
| REFERÊNCIAS BIBLIOGRÁFICAS .....                                  | 66 |

---

## LISTA DE FIGURAS:

|  |    |
|--|----|
| Figura II.1 – Escala de frequências e sua classificação dentro da faixa audível .....                            | 05 |
| Figura II.2 – Frentes de ondas circulantes que divergem de uma fonte puntiforme .....                            | 05 |
| Figura II.3 – Onda estacionária em uma corda fixa em ambas as extremidades .....                                 | 06 |
| Figura II.4 – Forma de onda senoidal simples .....   | 07 |
| Figura II.5 – Formas de ondas senoidais complexas .....  | 07 |
| Figura II.6 – Forma de onda aperiódica .....   | 08 |
| Figura II.7 – Dimensões físicas da onda sonora: frequência, amplitude e comprimento de onda .....                | 09 |
| Figura II.8 – Sons nas diferentes frequências e intensidades audíveis para a maioria da população. ....          | 15 |
| Figura II.9 – Espectro de frequência para sons habituais [NORTH1989] .....                                       | 16 |
| Figura III.1 – Sistema auditivo: ouvido externo, médio e interno .....   | 18 |
| Figura III.2 – Ondas sonoras na forma de pressão .....   | 19 |
| Figura III.3 – Movimento do líquido na cóclea quando o estribo é impelido para frente, Imagem: [GUYTO1981] ..... | 21 |



|  |    |
|--|----|
| Figura III.2 – Movimento quadro a quadro do órgão de Corti .....   | 22 |
| Figura V.1 – O Diapasão como recurso na avaliação por via óssea (A) e<br>via aérea (B) .....   | 27 |
| Figura V.2 – Representação dos símbolos audiométricos .....  | 31 |
| Figura V.3 – Simbologia usada para indicar no audiograma a falta de estímulo<br>em uma dada frequência .....   | 32 |
| Figura VI.1 – Sistema Auditivo, cóclea normal (a) e cóclea prejudicada (b) .....   | 33 |
| Figura VI.2 – Classificações da audição quanto ao tipo de perda .....  | 34 |
| Figura VI.3 – Exemplo de um caso de perda em rampa.....  | 35 |
| Figura VI.4 – Audiograma de uma criança residente em Vitória/ES .....  | 36 |
| Figura VII.1 – Esquema em blocos básico da Ajuda Digital Auditiva .....  | 38 |
| Figura VII.2 – Relação entre a frequência de entrada e a frequência<br>comprimida.....   | 40 |
| Figura VII.3 – Espectro de um quadro de 25 ms do sinal de voz<br>original e comprimido .....   | 42 |
| Figura VII.4 – Ilustração do algoritmo de janelamento para determinação da<br>frequência de origem da porção do espectro de altas frequências<br>a ser transladado ..... | 44 |
| Figura VII.5 – Espectro do quadro do sinal de voz original e transladado .....   | 45 |
| Figura VIII.1 – Janelamento do sinal de áudio .....  | 47 |
| Figura VIII.2 – Janela Hamming .....   | 48 |
| Figura VIII.3 – Fluxograma geral do processamento .....  | 50 |
| Figura VIII.4 – Fluxograma do processamento discriminativo .....   | 51 |

|  |    |
|--|----|
| Figura VIII.5 – Espectogramas do sinal de áudio da palavra “pensando” .....    | 53 |
| Figura IX.1 – Tela de entrada do programa de teste de inteligibilidade .....   | 55 |
| Figura IX.2 – Tela de respostas do programa de teste de inteligibilidade ..... | 56 |

---

## LISTA DE TABELAS:

|                  |   |    |
|------------------|---|----|
| Tabela II.1 –    | Velocidade do som em diferentes meios .....   | 10 |
| Tabela II.2 –    | Escala logarítmica em decibels pressão e intensidade sonora   | 13 |
| Tabela V.1 –     | Valores obtidos para o zero dB audométrico nos padrões de<br>calibração ASA-51,ISO-64 e ANSI-69 ..... | 29 |
| Tabela IXI.1 –   | Resultados percentuais de reconhecimento do primeiro grupo<br>de ouvintes .....                       | 58 |
| Tabela IXI.2 –   | Resultados percentuais de reconhecimento do segundo grupo<br>de ouvintes .....                        | 59 |
| Tabela IXI.3 –   | Resultados percentuais de reconhecimento compilados dos<br>dois grupos de ouvintes .....              | 60 |
| Tabela IXI.4.a – | Matriz de Confusão - Proc: nenhum; freq. corte: 1500 .....  | 62 |
| Tabela IXI.4.b – | Matriz de Confusão - Proc:compressão; freq. corte: 1500 ....  | 62 |
| Tabela IXI.4.c – | Matriz de Confusão - Proc: transladação; freq. corte: 1500 ..   | 62 |
| Tabela IXI.4.d – | Matriz de Confusão - Proc: nenhum; freq. de corte: 2000 ....  | 62 |
| Tabela IXI.4.e – | Matriz de Confusão - Proc: compressão; freq. corte: 2000 ...  | 63 |
| Tabela IXI.4.f – | Matriz de Confusão - Proc: transladação; freq. corte: 2000 ..   | 63 |
| Tabela IXI.4.g – | Matriz de Confusão - Proc: nenhum; freq. de corte: 2500 ....  | 63 |
| Tabela IXI.4.h – | Matriz de Confusão - Proc: compressão; freq. corte: 2500 ...  | 63 |
| Tabela IXI.4.i – | Matriz de Confusão - Proc: transladação; freq. corte: 2500 ..   | 63 |

---

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS:

|           |   |
|-----------|---|
| NIS       | - Nível de Intensidade Sonora                                       |
| NPS       | - Nível de Pressão Sonora, em inglês: SPL – Sound Pressure Level    |
| dB NA     | - decibel nível de audição  |
| ANSI      | - American National Standards Institute                             |
| ISO       | - International Standards Organization                              |
| ASA       | - American Standards Association                                    |
| ASHA      | - American Speech Hearing and Association                           |
| FFT       | - Transformada Rápida de Fourier (Fast Fourier Transform)           |
| IPA       | - Alfabeto Fonético Internacional (International Phonetic Alphabet) |
| FIX       | - Função do MatLab, retorna o valor do próximo número inteiro       |
| F         | - Frequência sonora, dada em Hz ou kHz                              |
| T         | - Período da onda sonora, dado em segundos                          |
| A         | - Amplitude da onda sonora, dada em volts.                          |
| $\omega$  | - Frequência angular, dada em radianos por segundos                 |
| $\lambda$ | - Comprimento de onda, espaço percorrido pela onda, dado em metros  |
| v         | - Velocidade de propagação do som, dado em metros/segundo           |
| Pe        | - pressão efetiva, geralmente dada em $N/m^2$ ou Pa                 |
| $\rho_0$  | - densidade de fluxo  |
| c         | - velocidade de fase  |
| dB        | - decibel, medida de nível sonoro logarítmico                       |

|                        |  |
|------------------------|--|
| Hz                     | - Hertz  |
| Pa                     | - Pascal   |
| $f_{IN}$               | - frequência de entrada                              |
| $f_{OUT}$              | - frequência de saída                                |
| $f_S$                  | - taxa de amostragem                                 |
| $K$                    | - fator de compressão                                |
| $a$                    | - fator de distorção                                 |
| $f_{DESTINO}$          | - frequência de destino do quadro transladado        |
| $f_{MÁX.PRESERVADA}$   | - máxima frequência preservada                       |
| $QtAmostrasSinal$      | - número total de amostras contido no sinal de áudio |
| $TamQuadro$            | - tamanho do quadro (número de amostras por quadro)  |
| $TamSobP$              | - número de amostras sobrepostas                     |
| $REL_{Faltas/Fbaixas}$ | - relação entre altas e baixas frequências           |

---

## **RESUMO:**

Uma considerável porcentagem de ouvintes com perdas auditivas severas apresenta audiogramas onde as perdas são grandes para altas frequências e pequenas para baixas frequências. Para esses casos, o rebaixamento do espectro de frequências para uma faixa onde existe alguma audição residual pode ser uma boa solução para se implementar um aparelho auditivo digital. Neste trabalho são apresentados dois diferentes algoritmos de rebaixamento de frequências: compressão e transladação. Resultados de testes subjetivos de discriminação de consoantes mostraram que um desempenho melhor foi observado quando empregados os algoritmos de processamento propostos.

---

## **ABSTRACT:**

A considerable percentage of listeners with severe hearing loss have audiograms where the losses are high for high frequencies and low for low frequencies. For these patients, lowering the speech spectrum to the frequencies where there is some residual hearing could be a good solution to be implemented in a digital hearing aid device. In this work we have presented two different frequency-lowering algorithms: frequency compression and frequency shifting. Results of subjective tests of consonant discrimination showed an improvement when the proposed algorithms were used.

# CAPÍTULO I

## INTRODUÇÃO

Existem vários tipos de deficiências auditivas. A origem de perdas auditivas sensoneurais pode ser devido a defeitos na cóclea, nervo auditivo ou ambos. Estes problemas reduzem a faixa dinâmica de audição. O limiar de audição é elevado, mas o limiar de desconforto é quase o mesmo do que para ouvintes com audição normal ou até pouco inferior. Para algumas faixas de frequências, o limiar de audição é tão alto que é igual ao limiar de desconforto, i.e., é impossível para o ouvinte escutar algum som para essas frequências.

A perda auditiva que pode comprometer a inteligibilidade de fala é mais comum para sons com altas e médias frequências do que para baixas frequências. Frequentemente existem somente pequenas perdas para baixas frequências (até 1 kHz) mas as perdas auditivas podem ser mais acentuadas acima de 1.5 ou 2 kHz. Esses fatos levaram pesquisadores a rebaixar o conteúdo espectral da fala com o objetivo de aproveitar a menor perda auditiva existente nas baixas frequências. Reprodução da fala a uma taxa de mostragem mais lenta ou redução da taxa de cruzamento por zero são alguns dos métodos que têm sido empregados nas últimas décadas [HICKS1981]. Todos esses métodos envolvem distorção de sinal, mais ou menos perceptível, geralmente dependente do grau de alteração espectral realizada.

Muitos dos esquemas de rebaixamento de frequências têm alterado perceptivelmente importantes características da fala, como padrões rítmicos e temporais, pitch e duração de elementos segmentais.



Hicks *et al.* [HICKS1981] fizeram uma das mais notáveis investigações sobre rebaixamento de frequências. Sua técnica envolve a compressão monotônica do espectro de tempo curto sem alteração do pitch e ao mesmo tempo evitando alguns dos problemas observados em outros métodos.

Reed *et al.* [REED1983] realizou experimentos discriminativos de consoantes com ouvintes cuja audição é normal. Eles observaram que o esquema de rebaixamento de frequências de Hicks apresentou melhor desempenho para sons fricativos e sons africados se comparado com filtragem passa-baixas (simulando a perda auditiva) para uma largura de banda equivalente.

Por outro lado, o desempenho da filtragem passa-baixas foi melhor para vogais, semivogais e sons nasais. Para sons plosivos, ambos os métodos mostraram resultados similares. Em geral, o desempenho nas melhores condições de rebaixamento de frequências foi quase o mesmo que obtido na filtragem passa-baixas para uma largura de banda equivalente. Em uma pesquisa posterior, Reed *et al.* [REED1985] estenderam os resultados do sistema de Hicks *et al.* para deficientes com perdas em altas frequências. Em geral, o desempenho observado nos ouvintes portadores de deficiência foi inferior do que o obtido por ouvintes normais.

Alguns anos atrás, Nelson e Revoile [NELSO1998] descobriram que relativamente a ouvintes com audição normal, aqueles com perda auditiva entre moderada e severa requerem aproximadamente o dobro da relação pico-vale para detecção de picos espectrais na presença de ruído quando os sinais têm um alto número de picos por oitava. A relação pico-vale é determinada pela relação entre o máximo valor de densidade de uma dada frequência e o menor valor subsequente de densidade de frequência. Os resultados revelaram que a detecção de picos espectrais em ruído está significativamente relacionada com a habilidade de identificação de consoantes em ouvintes com perda auditiva de moderada para severa.

Todos esquemas de rebaixamento de frequências mencionados comprimem o espectro de fala para uma estreita banda de frequências, aumentando o número de picos por oitava, porém mantendo a mesma relação pico-vale. Conforme a

investigação de Nelson e Revoile, o aumento da relação pico vale pode permitir melhor detecção de picos espectrais e, melhor identificação de consoantes. Recentemente, Muñoz *et al.* [MUÑOZ1999] combinaram formatação espectral (i.e., aumento da relação pico-vale) e compressão de frequências. Eles demonstraram que a fala assim processada melhorou a identificação de sons fricativos e africados enquanto não proporcionou mudança significativa na identificação de vogais e outros sons por ouvintes com perda auditiva severa em altas frequências.

Baseado na investigação de Nelson, formulamos a hipótese de que o desempenho relativamente pobre do esquema de rebaixamento de frequências de Hicks e Muñoz foi devido ao aumento do número de picos por oitava, o qual é inerente ao método de compressão de frequências usado nestes dois sistemas.

Neste trabalho, propomos um novo algoritmo de rebaixamento de frequências que não aumenta o número de picos por oitava, pois está baseado em transladação de frequências ao invés de compressão de frequências. Além disso, a transladação de frequências é aplicada somente para sons fricativos e africados, deixando todos os outros tipos de sons intocados, porque é somente para essa classe de sons que a técnica traz reais benefícios, para fins comparativos, como têm sido demonstrado por todos os trabalhos mencionados anteriormente. Além do algoritmo de transladação, foi implementado também um algoritmo de compressão de frequências baseado nas idéias de Hicks e Muñoz.

O segundo capítulo, intitulado “Audiologia”, apresenta características físicas do som e grandezas relativas ao som. O terceiro capítulo é uma extensão do anterior, associando o som com as sensações que o mesmo produz nos indivíduos. Com esse embasamento, é possível compreender como foram processadas as informações e como foram conduzidos os testes subjetivos de discriminação auditiva. Estes procedimentos, bem como os métodos empregados e a discussão dos resultados, estão descritos nos demais capítulos.

## CAPÍTULO II

### AUDIOLOGIA – FUNDAMENTOS ACÚSTICOS

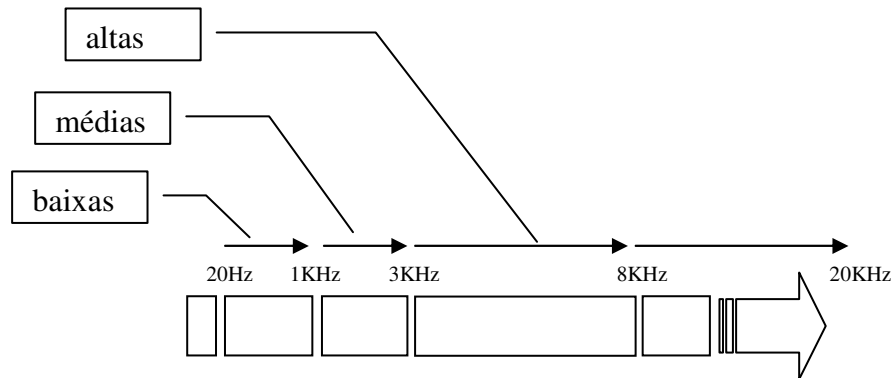
Para a compreensão das informações associadas ao processamento do som, é preciso conhecer alguns conceitos que envolvem a Audiologia.

“A Audiologia é a ciência que estuda a audição, e, logicamente, não há audição sem haver som” [FROTA2001]. Uma vez que o elemento básico de trabalho é o som, é de fundamental importância o estudo dos fenômenos acústicos e das sensações auditivas que os mesmos provocam. Assim, trataremos o assunto em três tópicos: a acústica, que nos trará conhecimentos físicos do som; a psicoacústica, responsável especificamente pelo estudo da relação entre os eventos acústicos (sons) e as sensações que os mesmos provocam nos indivíduos e finalmente a avaliação básica da audição.

#### II.1 – Acústica

A Fonoaudiologia é uma ciência nova, que comparada com outras, ainda está apenas engatinhando. Ela nos fornece poucos conceitos de ciências exatas que são, no entanto, extremamente importantes. Conforme [FROTA2001], a Acústica engloba a geração, transmissão e recepção de energia na forma de ondas vibracionais na matéria. Isso ocorre quando átomos e moléculas de um fluido ou sólido são deslocados de sua posição normal e uma força interna elástica de restauração atua no sentido de fazê-los voltar a essa posição.

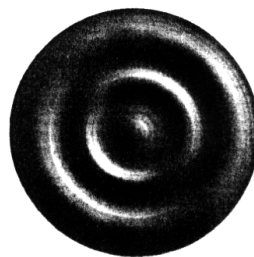
Considera-se som um distúrbio vibracional cuja frequência atinge uma faixa de 20 a 20.000 Hz em uma intensidade capaz de produzir sensação auditiva. Na **Figura II.1** pode ser vista a escala de frequências sonoras e suas respectivas classificações.



**Figura II.1** – Escala de frequências e sua classificação dentro da faixa audível

### II.1.1 – Ondas Sonoras

O movimento ondulatório pode ser imaginado como o transporte de energia, de um ponto para outro no espaço, sem o transporte simultâneo de matéria. Para as ondas sonoras, o transporte se dá através de uma oscilação ou vibração que se propaga em um meio elástico, provocando uma perturbação na pressão estática deste meio. Passada a perturbação, a pressão estática volta ao normal. Essa propagação se dá em todas as direções, e a perturbação das partículas do meio é paralela à direção de propagação da onda. Devido a essas duas características, a onda sonora é considerada tridimensional e longitudinal (**Figura II.2**).



**Figura II.2** – Frentes de ondas circulares que divergem de uma fonte puntiforme

Para ocorrer propagação das ondas sonoras, é necessário que o meio possua massa e elasticidade. Massa é a quantidade da matéria que está presente, e

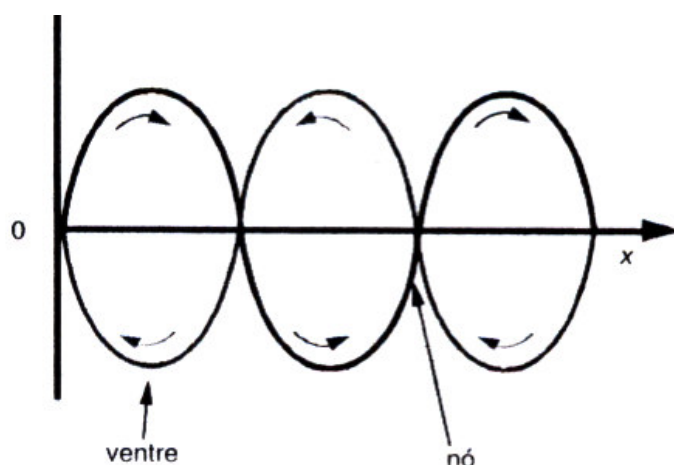
elasticidade é a propriedade de um corpo que possibilita a recuperação de uma distorção sofrida por uma força externa, tanto de forma como de volume.

## II.1.2 – Tipos de ondas sonoras

### a) Quanto à propagação da onda no espaço

As ondas sonoras quanto à propagação no espaço podem ser divididas em ondas sonoras progressivas (onde o movimento de todas as partículas têm a mesma amplitude) e ondas sonoras estacionárias (formadas quando uma onda bate em uma superfície e volta).

Nas ondas estacionárias, a amplitude de movimento de cada partícula depende da porção que a mesma ocupa em relação aos nós (**Figura II.3**): em um ventre a amplitude de vibração é máxima; em um nó a amplitude de vibração é mínima (nula).



**Figura II.3** – Onda estacionária em uma corda fixa em ambas as extremidades

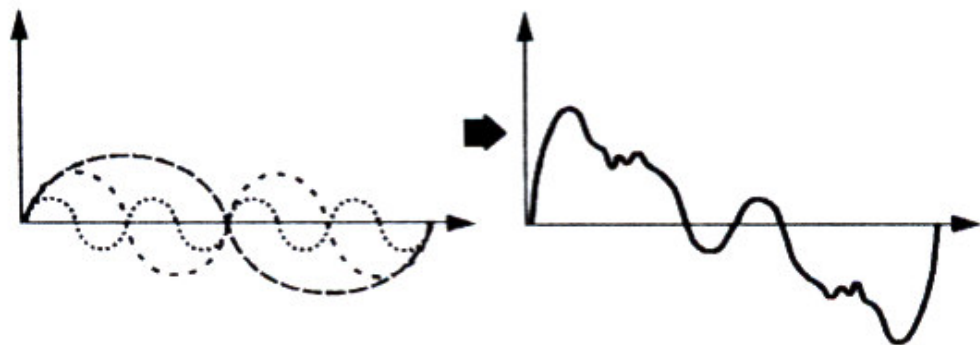
### b) Quanto à complexidade

As ondas sonoras quanto à complexidade podem ser divididas em ondas senoidais (que resultam de um movimento harmônico simples, que origina o tom puro, com uma única frequência, conforme pode ser visto na **Figura II.4**) e ondas complexas (que é qualquer onda sonora composta de uma série de senóides simples que podem diferir em amplitude, frequência e fase (**Figura II.5**)).



**Figura II.4** – *Forma de onda senoidal simples*

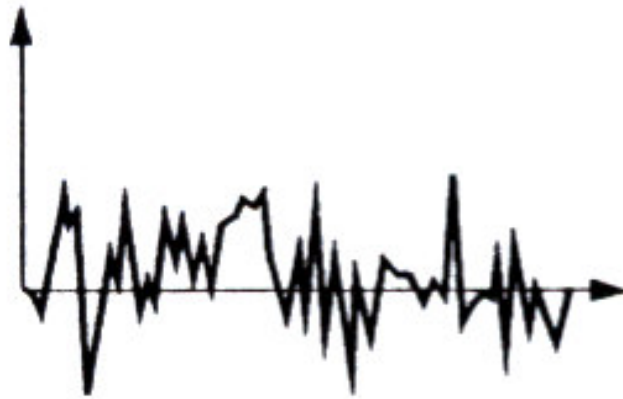
As ondas complexas podem ser analisadas quanto às senóides simples que as compõem através da chamada Análise de Fourier.



**Figura II.5** – *Formas de ondas senoidais complexas*

### c) Quanto à Periodicidade

As ondas sonoras quanto à periodicidade podem ser divididas em ondas sonoras periódicas (que se repetem em iguais intervalos de tempo) e ondas sonoras aperiódicas (nas quais não é possível identificar padrões de forma de onda em um intervalo de tempo, a partir do conhecimento prévio de suas características em outro intervalo de tempo (**Figura II.6**)).



**Figura II.6** – *Forma de onda aperiódica*

### **II.1.3 – Dimensões da onda sonora**

As ondas sonoras são determinadas por parâmetros físicos, tais como frequência, frequência angular, fase, amplitude e comprimento de onda (**Figura II.7**).

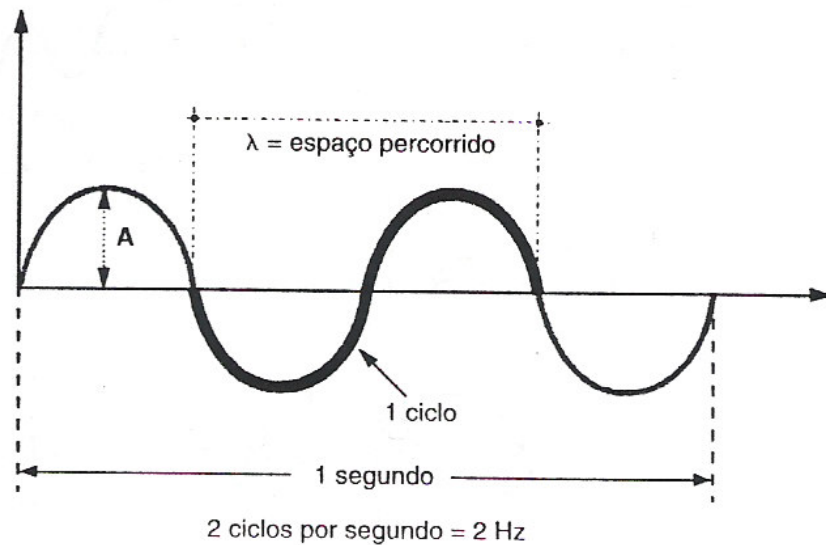
#### **a) Frequência**

A **frequência (f)** indica o número de movimentos completos que as partículas materiais realizam em um segundo. A unidade usada para determinar a quantidade de ciclos por segundo é o hertz (Hz). Assim como outras unidades de medida, apresenta também seus múltiplos, o kHz (1.000 Hz), o MHz (1.000.000 Hz) e o GHz (1.000.000.000 Hz).

O **período (T)**, é a quantidade de tempo, em segundos, requerida para as partículas materiais completarem 1 ciclo. A relação entre frequência e período é dada por:

$$f = \frac{1}{T} \text{ ou } T = \frac{1}{f} \quad (II.1)$$

A audição humana compreende uma faixa que vai de 20 Hz a 20.000 Hz. Outros animais têm faixa de percepção diferentes como, por exemplo, os gatos (10 a 60.000 Hz) e os golfinhos (de 10.000 a 240.000 Hz).



**Figura II.7** –Dimensões físicas da onda sonora: frequência, amplitude e comprimento de onda

### b) Frequência angular

A **frequência angular** ( $\omega$ ) é a frequência expressa em radianos por segundo. Desde que para completar um período é necessário que a fase cresça  $2\pi$ , a relação entre frequência e frequência angular é dada por:

$$\omega = 2\pi \cdot f \quad (II.2)$$

### c) Fase inicial

A **fase inicial** é o deslocamento, em graus, a partir de uma referência arbitrária que corresponde a  $0^\circ$ , no instante em que a vibração tem início. Duas ondas de igual frequência e amplitude podem diferir quanto à fase.

### d) Amplitude

A **amplitude** (**A**) é a medida do afastamento ou deslocamento das partículas materiais de sua posição de equilíbrio. Pode ser representada por um valor máximo (valor de pico) ou um valor eficaz (raiz da média quadrática). A amplitude relaciona-se à intensidade do som, podendo ser avaliada com o auxílio do osciloscópio, mas sua medida não é prática comum em Audiologia.



### e) Comprimento de Onda

O **comprimento de onda** ( $\lambda$ ) é o espaço percorrido pela oscilação, em metros, até que se repita o movimento em amplitude e fase. É denominado  $\lambda$ , e é dado pela relação:

$$\lambda = \frac{v}{f} \quad (II.3)$$

onde:  $v$  = velocidade de propagação do som

$f$  = frequência

No ar, a velocidade do som ao nível do mar e a 0°C (condições normais de temperatura e pressão – CNTP) é de 331,2 m/s. Geralmente adotamos para a velocidade do som o valor de 340 m/s, pois essa é a velocidade em uma temperatura de aproximadamente 14°C. A velocidade da onda sonora depende do meio em que ela se encontra. As velocidades do som em outros meios estão relacionados na **Tabela II.1**. Quanto maior a elasticidade e menor a densidade do meio, maior a velocidade do som.

**Tabela II.1** – Velocidade do som em diferentes meios:

| <b>Meio</b>          | <b>Temperatura<br/>(°C)</b> | <b>Velocidade<br/>(m/s)</b> |
|----------------------|-----------------------------|-----------------------------|
| ar                   | 0                           | 331,4                       |
| hidrogênio           | 0                           | 1286                        |
| oxigênio             | 0                           | 317,2                       |
| Água                 | 15                          | 1450                        |
| chumbo               | 20                          | 1230                        |
| alumínio             | 20                          | 5100                        |
| Cobre                | 20                          | 3560                        |
| Ferro                | 20                          | 5130                        |
| Granito              | 0                           | 6000                        |
| borracha vulcanizada | 0                           | 54                          |

#### II.1.4. – “Qualidades” da onda sonora – altura e intensidade

A percepção das dimensões físicas das ondas sonoras permite falar em “qualidades” dos sons. Entre estas “qualidades”, temos a altura e a intensidade.

A **altura** permite classificar os sons em graves ou agudos. Relaciona-se com a frequência: quanto maior a frequência, mais agudo é o som, e quanto menor a frequência, mais grave é o mesmo.

A **intensidade** permite classificar o som em forte e fraco. Relaciona-se com a amplitude da onda sonora e diz respeito à taxa do fluxo de energia através de uma unidade de área, normal à direção de propagação. Quanto maiores a amplitude, energia e pressão sonora, mais forte é o som. A intensidade relaciona-se com a pressão efetiva gerada por ondas planas progressivas e ondas esféricas da seguinte forma [FROTA2001]:

$$I = \frac{P_e^2}{\rho_0 c} \quad (II.4)$$

onde:  $P_e$  = pressão efetiva

$\rho_0$  = densidade de fluxo

$c$  = velocidade de fase

Assim, podemos analisar a intensidade de um som do ponto de vista da **intensidade sonora** (fluxo de energia) ou da **pressão sonora** (pressão que as moléculas exercem).

**Intensidade sonora** é a quantidade de energia transmitida por segundo sobre uma área de  $1\text{m}^2$ . Ela é dada em  $\text{W}/\text{m}^2$  ou  $\text{W}/\text{cm}^2$ , sendo o watt (W) uma unidade de fluxo de energia e o  $\text{m}^2$  e o  $\text{cm}^2$ , unidades de área.

Já a **pressão sonora** é definida como a força oscilante exercida pelas moléculas de ar sobre uma superfície na qual incidem. Ela costuma ser definida em  $\text{N}/\text{m}^2$  ou Pa. O Newton (N) é uma unidade de força e o pascal (Pa), uma unidade de pressão. Um pascal é igual a um Newton por metro quadrado.

## II.1.5 – Níveis de audição

É comum o uso de escalas logarítmicas para descrever pressões ou intensidades sonoras. Essas escalas são conhecidas como **níveis sonoros**. Como as pressões e intensidades sonoras presentes no ambiente compõem uma larga faixa, o uso de uma escala logarítmica condensa essa faixa extensa de números. A faixa de intensidades audíveis vai de aproximadamente  $10^{-16}$  até  $10^{-4}$  W/cm<sup>2</sup> ou de  $2 \times 10^{-5}$  a 20 Pa.

O uso de uma escala logarítmica dá-se também devido ao fato de que a percepção da sonoridade obedece a uma progressão logarítmica. A escala logarítmica mais utilizada para descrever níveis sonoros é a escala de decibel. O **decibel (dB)** significa 10 vezes o logaritmo de base 10 da razão entre duas grandezas, ou um décimo de um bel (o logaritmo de base 10 da razão entre duas grandezas, representado por B).

$$x \text{ decibels} = 10 \cdot \log \frac{G}{G_{ref}} \quad (II.5)$$

Assim, o decibel é uma escala que necessita de uma grandeza de referência. Conforme a grandeza de referência usada, a escala será diferente. Quando a escala do decibel é definida em função da intensidade sonora, tem-se o nível de intensidade sonora (NIS):

$$NIS = 10 \log \frac{I}{I_{ref}} \quad (II.6)$$

sendo:  $I_{ref}$  = intensidade de referência ( $10^{-12}$  W/m<sup>2</sup> ou  $10^{-16}$  W/cm<sup>2</sup>)

Substituindo (II.4) em (II.6), obtemos:

$$NPS = 10 \cdot \log \left( \frac{P_e}{P_{ref}} \right)^2 = 20 \cdot \log \frac{P_e}{P_{ref}} \quad (II.7)$$

sendo:  $P_{ref}$  = pressão efetiva de referência (20  $\mu$ Pa)

NPS = nível de pressão sonora expresso em decibels com referência a  $P_{ref}$  (dB re  $P_{ref}$ ). Em inglês: SPL – Sound Pressure Level.

O decibel é um valor relativo, de modo que sempre se deve saber qual foi o padrão usado para o dB. O padrão é a menor intensidade ou pressão possível de ser ouvida (**Figura II.8**).

**Tabela II.2** – Escala logarítmica em decibels pressão e intensidade sonora:

| <b>PRESSÃO SONORA</b>   | <b>INTENSIDADE SONORA</b>                       |
|---|---|
| decibel nível de pressão sonora   | decibel nível de intensidade sonora             |
| Referência: 20 $\mu$ Pa   | Referência: 10 <sup>-16</sup> W/cm <sup>2</sup> |
| dB NPS re 20 $\mu$ Pa ou<br>dB NPS re 2 x 10 <sup>-5</sup> N/m <sup>2</sup> | dB NIS re 10 <sup>-16</sup> W/cm <sup>2</sup>   |

É mais comum na aplicação em audiologia clínica o uso de medidas feitas com uma pressão de referência, sendo assim usado o NPS.

Quando a pressão de saída e de referência são iguais, o valor em dB é zero (o que pode ser verificado nas expressões, pois o logaritmo de 1 é zero). Isso não significa que exista silêncio: zero dB é um ponto de referência arbitrário.

Toda vez que um valor de decibel aumenta 10 unidades, a intensidade ou pressão aumenta 10 vezes. Por exemplo, a intensidade, quando se tem 50 dB, é 100 vezes maior que a intensidade quando se tem 30 dB, ao passo que a pressão é apenas 10 vezes maior.

Quando a intensidade é duplicada, o NIS aumenta em 3dB, quando a pressão é duplicada, o NPS aumenta em 6dB.

Em audiologia são utilizados também outros tipos de escalas de decibels, como por exemplo o dB NA, dB S, dB A, dB C e outros. Os exames audiológicos são normalmente relacionados numa escala chamada dB NA.

### II.1.6 – Decibel Nível de Audição (dB NA)

A sensibilidade do ouvido varia para as diferentes frequências. Para cada frequência, existe um nível mínimo de audição para a maioria da população. Experimentos psicoacústicos determinaram a faixa de audição humana (20 a 20.000 Hz) e o limiar mínimo de detecção ou audibilidade, isto é, a mais fraca intensidade ou pressão sonora capaz de impressionar o ouvido humano (para um tom puro em 50% das vezes em que o estímulo sonoro é apresentado).

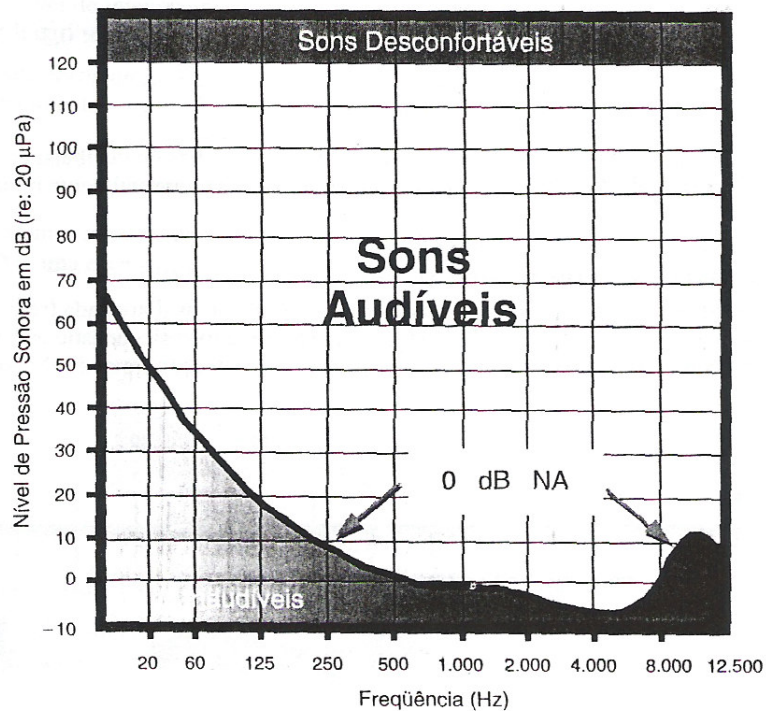
Para a frequência de 1.000 Hz este limiar equivaleria a uma pressão de 20 $\mu$ Pa. Este valor de pressão determina o estabelecimento do 0 dB NA para 1.000 Hz.

O procedimento foi empregado para as demais frequências, e foi construído um gráfico onde, na abscissa, eram dispostas as frequências e, na ordenada, os valores de pressão sonora. Também foram pesquisados os limiares de desconforto e dor, que para um tom de 1.000 Hz estão a aproximadamente 120 e 140 dB, respectivamente.

Assim, através de uma pesquisa realizada com indivíduos otologicamente normais, com idades entre 18 e 25 anos, criou-se um padrão para a audição das pessoas normais, sendo que o 0 dB neste padrão corresponde à menor intensidade ou pressão sonora por frequência ouvida pela maioria da população (**Figura II.8**).

O audiômetro de tons puros foi projetado como um instrumento eletrônico calibrado de tal forma que a leitura zero para cada frequência corresponde a este nível de audição modal para jovens adultos normais. Existem padrões para que a calibração seja realizada da forma adequada (por exemplo, ANSI-69, ASA-51 ou ISO-64).

Assim, ao realizar uma audiometria tonal, está-se na verdade verificando qual a diferença, em decibel para cada frequência, que a audição do indivíduo tem em relação à maioria da população (o 0 dB NA em cada frequência). Normalmente esta avaliação pode ser feita em intervalos de oitava ou meia oitava entre as frequências de 125 a 8.000 Hz e em intensidades que variam de -10 dB NA (audição 10 dB melhor que a maioria da população) a 120 dB NA.

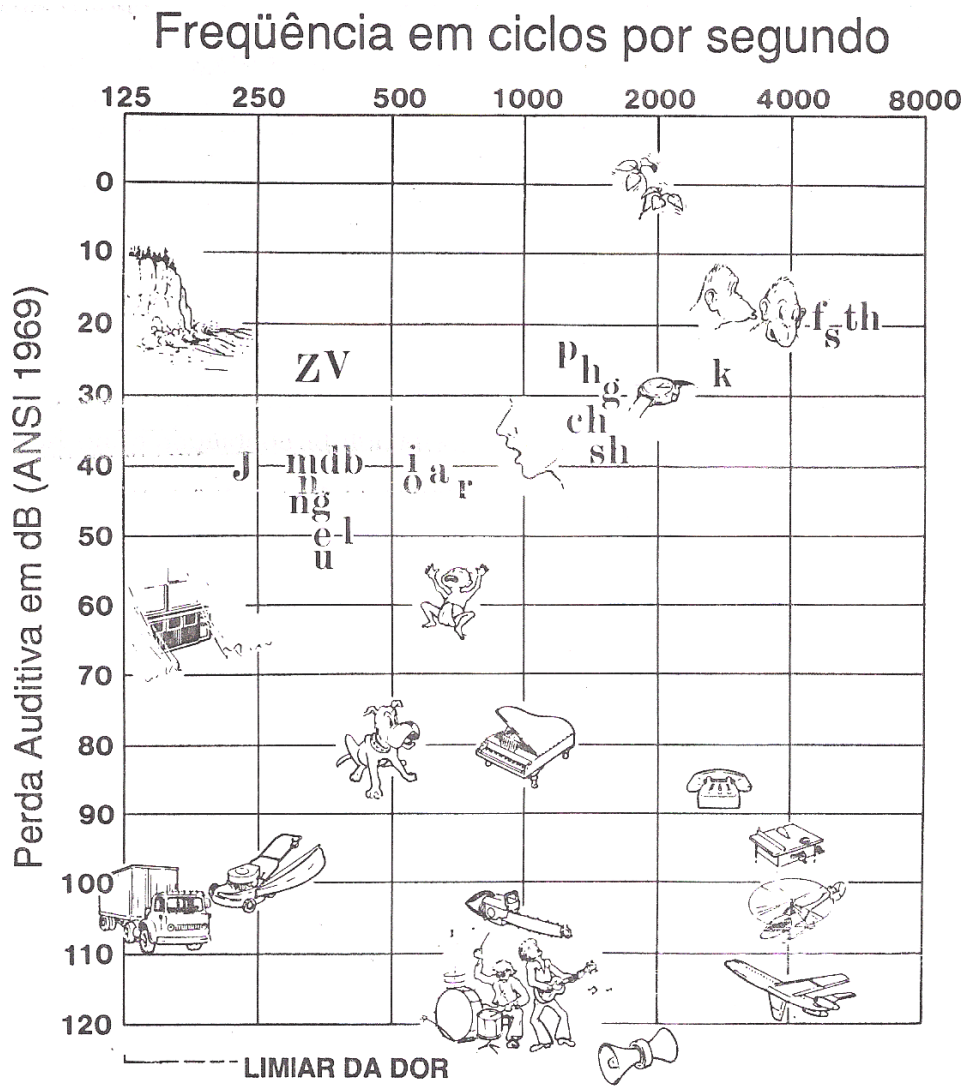


**Figura II.8** – Sons nas diferentes frequências e intensidades audíveis para a maioria da população.

A linha assinalada na **Figura II.8** corresponde ao 0 dB NA. Observa-se que, dependendo da frequência, uma mesma pressão sonora pode gerar ou não sensação auditiva.

Na **Figura II.9** estão ilustrados vários sons habituais, distribuídos conforme as suas respectivas características de frequência e intensidade.

Crianças com perdas auditivas de grau leve na fase de aquisição de fala e linguagem terão possivelmente dificuldades para identificar os sons de energia reduzida, como “f”, “s”, “th”, “z” e “v”.



**Figura II.9** – Espectro de frequência para sons habituais [NORTH1989]

# **CAPÍTULO III**

## **ANATOMIA E FUNCIONAMENTO DO SISTEMA AUDITIVO**

### **III.1 – Introdução**

O órgão responsável pela audição é o ouvido. O ouvido consiste em 3 partes básicas: o ouvido externo, o ouvido médio e o ouvido interno. Cada parte serve para uma função específica para interpretar o som.

O ouvido externo serve para coletar o som e o levar por um canal ao ouvido médio. O ouvido médio serve para transformar a energia de uma onda sonora em vibrações internas da estrutura óssea do ouvido médio e finalmente transformar estas vibrações em uma onda de compressão ao ouvido interno. O ouvido interno serve para transformar a energia da onda de compressão dentro de um fluido em impulsos nervosos que podem ser transmitidos ao cérebro. As três partes do ouvido estão ilustradas na **Figura III.1**.

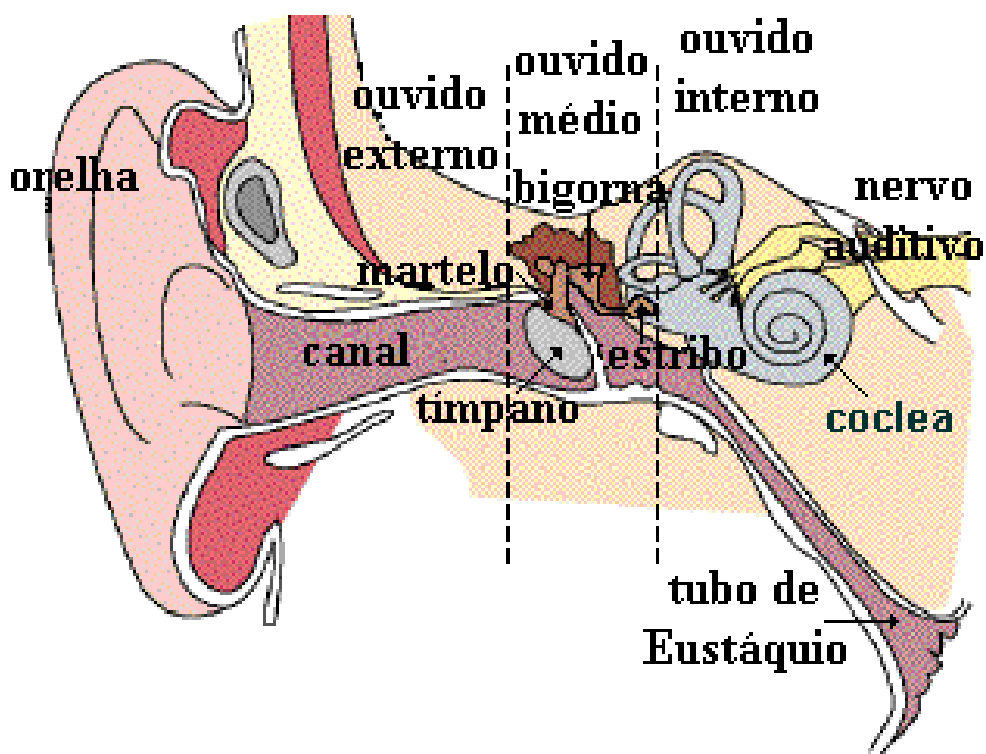
### **III.2 – Ouvido externo - energia sonora**

O ouvido externo é formado pelo pavilhão e pelo conduto auditivo externo. O pavilhão é comumente chamado de orelha, sendo formado por cartilagens cobertas por pele. Sua finalidade é coletar o som e conduzi-los para o meato externo. As principais funções desse meato auditivo são proteção e ressonância sonora. A



proteção é exercida pela cera produzida por glândulas ceraminosas da pele do meato. Esta cera forma uma película sobre a pele que impede a invasão de microorganismos. O meato auditivo externo tem 2,5 cm de comprimento e termina na membrana timpânica, amplificando bastante os sons na faixa de frequências de 1.500 Hz a 7.000 Hz.

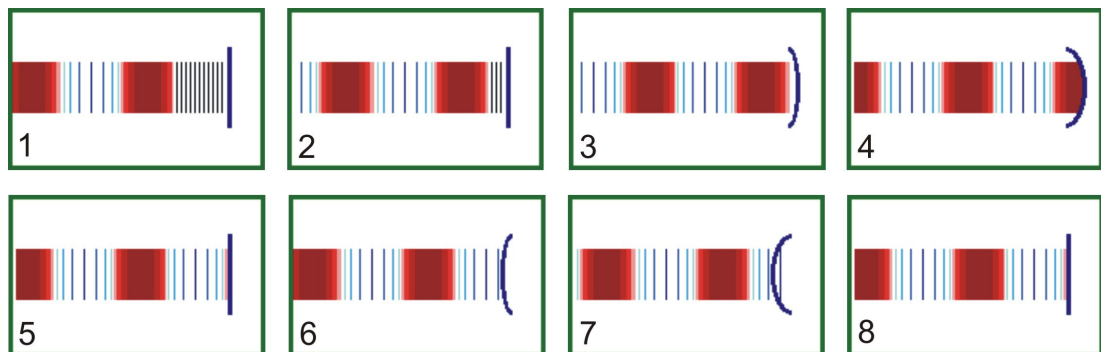
À medida que o som se propaga através do ouvido externo, o mesmo ainda se encontra na forma de uma onda de pressão, que é uma seqüência alternada de regiões de pressões mais baixas e mais altas. Somente quando o som alcança o tímpano, na separação entre o ouvido externo e médio, a energia da onda é convertida em vibrações na estrutura óssea do ouvido.



**Figura III.1** – Sistema auditivo: ouvido externo, médio e interno

### III.3 – Ouvido médio – energia mecânica

É a chamada caixa do tímpano, limitada externamente pela membrana timpânica e internamente pelo promontório, onde estão as janelas oval e redonda. Superiormente, encontra-se atígio e antro com células do mastóide e, inferiormente, a abertura da tuba auditiva. No ouvido médio, há três elementos importantes para a transmissão sonora: a trompa de Eustáquio, os músculos timpânicos e o sistema tímpano-ossicular, composto de 3 pequenos ossos interconectados - o **martelo**, a **bigorna** e o **estribo**. O tímpano é uma membrana muito durável e bem esticada que vibra quando a onda a alcança. Como mostrado na **Figura III.2**, uma compressão força o tímpano para dentro e a rarefação o força para fora. Logo, o tímpano vibra com a mesma frequência da onda. Como ela está conectada ao martelo, os movimentos do tímpano colocam o martelo, a bigorna e o estribo em movimento com a mesma frequência da onda. O estribo é conectado ao ouvido interno. Assim, as vibrações do estribo são transmitidas ao fluido do ouvido médio e criam uma onda de compressão dentro do fluido.



**Figura III.2** – Ondas sonoras na forma de pressão, adaptado da animação original do conteúdo de [www.if.ufrj.br/teaching/fis2/ondas2/ouvido/ouvido.html](http://www.if.ufrj.br/teaching/fis2/ondas2/ouvido/ouvido.html)

Os 3 pequenos ossos do ouvido médio agem como amplificadores das vibrações da onda sonora. Devido à vantagem mecânica, os deslocamentos da bigorna são maiores do que os do martelo. Além disso, como a onda de pressão que atinge uma grande área do tímpano é concentrada em uma área menor na bigorna, a força da bigorna vibrante é aproximadamente 15 vezes maior do que aquela do

tímpano. Esta característica aumenta nossa possibilidade de ouvir o mais fraco dos sons.

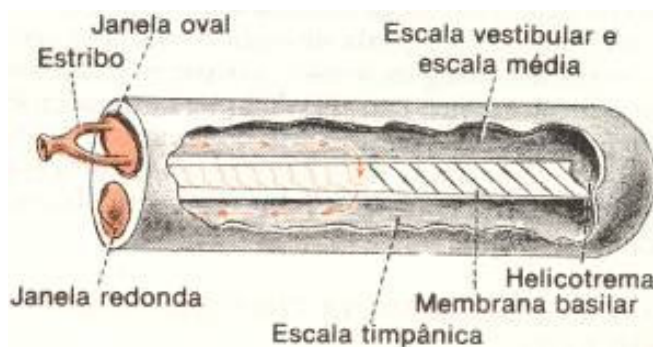
### **III.4 – Ouvido interno, energia hidráulica**

O ouvido interno consiste de uma **cóclea**, canais semicirculares, e do nervo auditivo. A cóclea e os canais semicirculares são cheios de líquido. O líquido e as células nervosas dos canais semicirculares não têm função na audição; eles simplesmente servem como acelerômetros para detectar movimentos acelerados e na manutenção do equilíbrio do corpo.

A cóclea é um órgão em forma de um caramujo que pode esticar até 3 cm. Além de estar cheio de fluido, a superfície interna da cóclea está alinhada com cerca de 20.000 células nervosas responsáveis pelas funções mais críticas na nossa capacidade de ouvir. Estas células nervosas possuem comprimentos diferentes, por diferenças minúsculas; elas também possuem diferentes graus de elasticidade no fluido que passa sobre elas. À medida que uma onda de compressão se move da interface entre o tímpano e o martelo no ouvido médio para a *janela oval* no ouvido interno, através da cóclea, as células nervosas na forma de cabelos entram em movimento. Cada célula capilar possui uma sensibilidade natural a uma frequência de vibração particular. Quando a frequência da onda de compressão “casa” com a frequência natural da célula nervosa, a célula irá ressoar com uma grande amplitude de vibração. Esta vibração ressonante induz a célula a liberar um impulso elétrico que passa ao longo do nervo auditivo para o cérebro. Em um processo que ainda não é compreendido inteiramente, o cérebro é capaz de interpretar as qualidades do som pela reação dos impulsos nervosos.

À medida que cada vibração sonora penetra na cóclea, a janela oval move-se para dentro, lançando o líquido da escala vestibular numa profundidade maior dentro da cóclea. A pressão aumentada na escala vestibular desloca a membrana basilar para dentro da escala timpânica; isso faz com que o líquido dessa câmara seja empurrado na direção da janela oval, provocando, por sua vez, o arqueamento da mesma para fora. Assim, quando as vibrações sonoras provocam

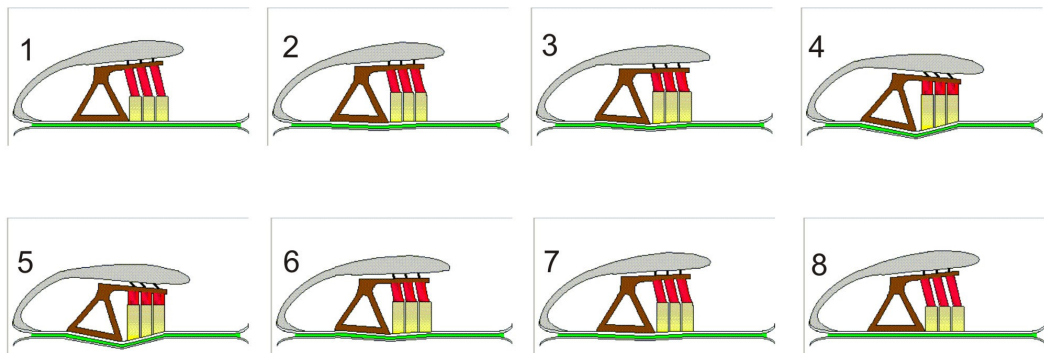
a movimentação do estribo para trás, o processo é invertido, e o líquido, então, move-se na direção oposta através do mesmo caminho, e a membrana basilar desloca-se para dentro da escala vestibular.



**Figura III.3** – Movimento do líquido na cóclea quando o estribo é impelido para frente,

Imagem: [GUYTO1981]

A vibração da membrana basilar faz com que as células ciliares do órgão de Corti se agitem para frente e para trás; isso flexiona os cílios nos pontos de contato com a membrana tectórica (tectorial). Na **Figura III.4** está ilustrado quadro-a-quadro o movimento do órgão de Corti. A flexão dos cílios excita as células sensoriais e gera impulsos nas pequenas terminações nervosas filamentosas da cóclea, que enlaçam essas células. Esses impulsos são então transmitidos através do nervo coclear até os centros auditivos do tronco encefálico e córtex cerebral. Dessa forma, a energia hidráulica é convertida em energia elétrica.



**Figura III.4** – Movimento quadro a quadro do órgão de Corti, adaptado da animação original de Fábio Mammano, do conteúdo de [www.afh.bio.br/sentidos/sentidos4.asp](http://www.afh.bio.br/sentidos/sentidos4.asp)

## CAPÍTULO IV

# PSICOACÚSTICA

A **psicoacústica** é parte de uma ciência chamada **psicofísica**, que estuda a relação entre os eventos físicos e as sensações que os mesmos produzem nos indivíduos. A **psicoacústica**, especificamente, estuda a relação entre os eventos acústicos (sons) e as sensações que os mesmos provocam nos indivíduos, traduzidas por reações comportamentais.

Como existe uma relação entre os eventos físicos e as sensações, cada uma das características físicas do som (intensidade, frequência e outras) pode ser analisada separadamente, no que se refere às sensações dos indivíduos.

Para se avaliar a sensação que um estímulo acústico provoca em um indivíduo, duas medidas são especialmente importantes: o limiar absoluto e a diferença de limiar.

O **limiar absoluto** identifica a transição da sensação para a não sensação e vice-versa (permitindo identificar o que é ouvido e o que não é ouvido), e a **diferença de limiar** indica uma mudança perceptível na sensação (demonstrando quais estímulos parecem iguais e quais estímulos parecem diferentes). Na Audiometria Tonal Liminar interessa a obtenção de um limiar de audição.

Muito se tem evoluído no âmbito da psicoacústica, mas os conhecimentos mais tradicionais não se tornaram, em sua maioria, inválidos. Eles apenas foram complementados, de forma que hoje se possui uma compreensão mais ampla do modo como os eventos sonoros são percebidos pelos seres humanos.

#### IV.1 – Métodos psicofísicos

Para avaliar quais as sensações que determinados sons provocam em um indivíduo, normalmente se lança mão de métodos de medida já estabelecidos. Estes métodos são normalmente emprestados da psicofísica e adaptados à realidade psicoacústica. Na psicofísica, chama-se *examinador* o indivíduo que conduz o teste e *observador* o indivíduo que está sendo testado.

São três os métodos psicofísicos clássicos (desenvolvidos por Fechner em 1860): o método dos limites, o método dos ajustes e o método dos estímulos constantes [HIRSH1952]. Os mesmos são usados da seguinte forma para análise dos estímulos auditivos, por exemplo na obtenção de um limiar de audibilidade:

- No **método dos limites** é apresentado um sinal em uma intensidade pré-determinada, e o observador responde se ouve ou não. Se ouve, a intensidade é reduzida até que ele deixe de ouvir (série descendente); se não ouve, a intensidade é aumentada até que ele ouça (série ascendente). As séries ascendentes ou descendentes são então alternadas, em um número mínimo de 10 séries. A menor intensidade ouvida 50% das vezes define o limiar de sensação.
- No **método dos ajustes**, a partir de uma intensidade inicial preestabelecida, o observador aumenta ou diminui ele próprio a intensidade do som até que o mesmo se torne audível/inaudível. Quando o som inicial é audível, o observador deve reduzi-lo até que se torne inaudível (série descendente), ou vice-versa (série ascendente). As séries ascendentes e descendentes são alternadas de forma que se possa estabelecer o limiar de sensação da mesma forma que no método anterior.
- No **método dos estímulos constantes**, diversas intensidades são apresentadas aleatoriamente por diversas vezes e, a cada uma delas, o observador responde se ouve ou não. A intensidade à qual o observador refere ouvir em 50% das apresentações é definida como o limiar de sensação.

Várias avaliações podem ser realizadas com os mesmos métodos, fazendo-se as adaptações necessárias. Estes e outros métodos psicofísicos, como os *métodos de escala* ou *método das quantidades*, são utilizados em testes na Audiologia. Para a audiometria tonal liminar utiliza-se habitualmente uma adaptação do método dos limites, chamado de **método adaptativo**.

Existem ainda basicamente três modos de apresentar os estímulos acústicos a um indivíduo que está sendo testado auditivamente: a estimulação monótica, a estimulação diótica e a estimulação dicótica.

A estimulação monótica é aquela em que o estímulo é apresentado exclusivamente em uma orelha. É o que acontece normalmente durante a realização de uma audiometria tonal ou logoaudiométrica.

A estimulação diótica, por sua vez, é aquela em que um mesmo estímulo é apresentado simultaneamente às duas orelhas. Finalmente, a estimulação dicótica acontece quando estímulos diferentes são apresentados simultaneamente às duas orelhas, como ocorre em alguns testes para avaliar processamento auditivo central.

#### **IV.2 – Percepção da intensidade do som**

A intensidade do som percebida pelos seres humanos pode ultrapassar os 120 dB, com sons que possuem uma pressão sonora que varia entre 20  $\mu\text{Pa}$  e 20 Pa ou mais. Entretanto, todos esses sons não são percebidos da mesma forma: alguns parecem excepcionalmente fracos, enquanto outros provocam dor por serem fortes demais.

Enquanto sons com uma pressão sonora de 20  $\mu\text{Pa}$  podem ser ouvidos em determinadas frequências, em outras podem não ser ouvidos, mesmo por um indivíduo que apresente uma audição normal. A percepção da intensidade do som pelos indivíduos depende de diversos fatores, alguns comuns a toda a população e outros, particulares.



### **IV.3 – Percepção da altura (ou frequência)**

Fundamental na capacidade humana de reconhecimento de diferentes fontes sonoras, a percepção e a diferenciação de diferentes frequências é uma das habilidades auditivas mais importantes do ser humano.

Do mesmo modo que o *loudness* é a dimensão psicoacústica que corresponde à sensação de intensidade, o *pitch* é a dimensão que corresponde à **sensação de frequência ou altura**.

O *pitch* de um som é influenciado basicamente por sua frequência ou, no caso de sons complexos, pelas suas frequências formantes mais importantes. A percepção do *pitch* ainda é influenciada pela duração, intensidade e complexidade da onda sonora.

O *pitch* está intimamente ligado à vibração das cordas vocais. Para sons surdos, onde não há vibração das cordas vocais, não existe *pitch*. Por isto a dificuldade de portadores de deficiência auditiva em falar sons surdos como “f” e “v”.

### **IV.4 – Percepção das variações temporais**

A fala compõe-se de uma série de seqüências de padrões sonoros que, ao longo do tempo, podem ser interpretadas. Assim, a percepção de padrões temporais é uma característica fundamental na audição, porque os eventos acústicos têm habitualmente uma duração definida e um intervalo entre suas ocorrências.

A audição é um sentido essencialmente ligado ao tempo.

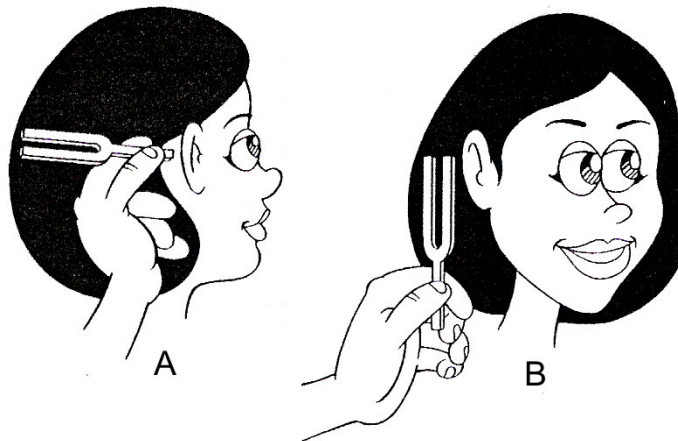
## CAPÍTULO V

# AVALIAÇÃO BÁSICA DA AUDIÇÃO

### V.1 – Introdução

Com o advento da válvula, surgiu a possibilidade de controlar e medir a intensidade e a altura do som. Assim, em 1922, surgiu o primeiro audiômetro. Antes disso, o recurso utilizado para testes auditivos era o diapasão. Os diapasões são barras metálicas em forma de U que, quando vibradas, executam movimentos de vaivém em direções opostas (**Figura V.1**). O diapasão pode ser empregado tanto para testes por via óssea (**Figura V.1.A**), quanto por via aérea (**Figura V.1.B**). O diapasão ainda continua sendo utilizado para confirmar diagnóstico.

Com o surgimento do audiômetro, nasceu conseqüentemente a Audiometria. Atualmente, existe no mercado uma infinidade de equipamentos para exames auditivos, possibilitando uma avaliação auditiva mais fidedigna.



**Figura V.1** – O Diapasão como recurso na avaliação por via óssea (A) e via aérea (B)

## **V.2 – Anamnese**

Antes da avaliação auditiva, deve-se sempre realizar uma conversa com o paciente a fim de obter importantes informações através da anamnese. É o primeiro contato com o paciente. Consiste de uma entrevista que, quando bem feita, traz informações que possibilitam uma idéia quantitativa da audição e hipóteses diagnósticas. Deve sempre ser valorizada a principal queixa do paciente, a partir da qual são selecionados os aspectos de maior relevância a serem questionados.

Segundo FROTA [FROTA2001], após ter sido feita a identificação completa do indivíduo, devem ser feitas perguntas sobre:

- Início e evolução do problema;
- Os sinais e sintomas (dor, presença de otorréia, prurido, diminuição da audição, dificuldade de entender as palavras, zumbido, vertigem, etc)
- Os antecedentes familiares
- Exposição a ruídos intensos
- Passado otológico
- Uso de medicamentos ototóxicos.

Durante a anamnese, enquanto se conversa, pode-se avaliar a capacidade auditiva do indivíduo, usando várias intensidades de voz a fim de se observar como o paciente se comporta e omitindo pistas visuais (não permitindo a leitura labial).

## **V.3. – Inspeção do meato acústico**

Após feita a anamnese, deverá ser realizada uma inspeção no meato acústico. O objetivo é uma inspeção cuidadosa do meato acústico externo e visualização da membrana timpânica, antecedendo a avaliação audiológica. É de grande importância, pois permite identificar a presença de corpos estranhos e de rolha de cera, que pode prejudicar a correta obtenção dos limiares tonais. Essa obstrução do canal auditivo pode causar perdas de até 40 dB NA.

#### V.4 – Pesquisa dos limites tonais por via aérea

Desde 1973, a padronização do American National Standardization Institute (ANSI) define o limiar de audição como o mínimo de pressão sonora efetiva para produzir uma sensação auditiva. Uma vez que essa sensação auditiva varia de pessoa para pessoa, o limiar deve ser descrito em termos de balanço de probabilidade. Considera-se então o limiar como sendo a menor intensidade do sinal necessária para provocar resposta em 50% das apresentações.

É bom lembrar que o valor medido pelo teste audiométrico não é em nível de pressão sonora – NPS, mas sim em nível de audição – NA. Isso significa que as respostas do paciente são o mínimo de pressão sonora necessário para de gerar sensação auditiva. Como os padrões levam em conta que as pessoas são diferentes e a audição também está relacionada à percepção, esse NA não é igual em todos os padrões, variando em menos de 5dB entre o padrão ISO-64 e o ANSI-69, conforme mostrado na *Tabela V.1* a seguir:

**Tabela V.1** – Valores obtidos para o zero dB audométrico nos padrões de calibração ASA-51, ISO-64 e ANSI-69

| <b>Frequência</b> | <b>125</b> | <b>250</b> | <b>500</b> | <b>1.000</b> | <b>2.000</b> | <b>4.000</b> | <b>8.000</b> | <b>Hz</b> |     |
|-------------------|------------|------------|------------|--------------|--------------|--------------|--------------|-----------|-----|
| ASA-51            | 51,8       | 39,5       | 24,1       | 17,2         | 18,0         | 14,3         | 26,8         | dB        | NPS |
| ISO-64            | 42,8       | 24,5       | 10,1       | 7,2          | 9,5          | 8,3          | 15,3         | dB        | NPS |
| ANSI-69           | 45,0       | 25,5       | 11,5       | 7,0          | 9,0          | 9,5          | 13,0         | dB        | NPS |

A audiometria necessita de um ambiente acusticamente tratado para que o ruído ambiental não mascare o teste, acusando perdas auditivas inexistentes. Assim, a sala a ser escolhida para realização dos exames deve ficar longe do barulho de carros e pessoas.

O aparelho eletroacústico utilizado para gerar os tons puros é chamado de audiômetro, que deverá ter os seguintes componentes básicos:

- Gerador de frequência: produz tons puros, nas frequências de 125, 250, 500, 750, 1.000, 1.500, 2.000, 3.000, 4.000, 6.000 e 8.000 Hz (escala básica de frequências, existem audiômetros com frequências superiores).
- Ajustes de intensidade: permite calibrar o nível do sinal (dB NA) em cada frequência. As escalas de intensidade são graduadas de forma diferente para a via aérea e óssea, com intervalos de 5dB. Existe um segundo atenuador que gradua a intensidade do ruído.
- Gerador de ruído: gera ruídos utilizados no mascaramento dos tons puros e testes logaudiométricos.
- Par de fones: utilizado nos testes por via aérea, cuja função é transformar a corrente elétrica em pressão sonora, em cada frequência e intensidade correspondentes.
- Vibrador ósseo: utilizado nos testes por via óssea, sua função é transformar a corrente elétrica em energia vibratória.
- Circuito para a logaudiometria: possibilita a realização de testes de fala.

A pesquisa dos limiares tonais por via aérea reflete o funcionamento de todo o sistema auditivo periférico.

É fundamental para o sucesso das respostas esperadas que seja feita uma **explicação detalhada do exame**. Deve ficar claro para o paciente que o objetivo do exame é descobrir a menor intensidade de som que ele é capaz de perceber, e que cada vez que ele escutar deve avisar imediatamente.

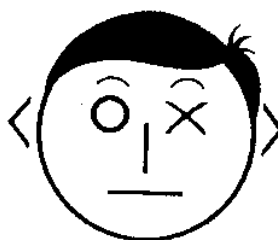
É recomendado iniciar o teste na frequência de 1.000 Hz por ser este um tom facilmente percebido. A seguir, pesquisa-se as frequências de 2.000, 4.000, 8.000, 500 e 250 Hz, nessa ordem. Para portadores de deficiência auditiva que apresentam perdas auditivas severas e profundas (\*), essa recomendação deve ser revista, iniciando o teste pelas frequências mais graves. Além das frequências citadas, é de fundamental importância a pesquisa das intermediárias (ou  $\frac{1}{2}$  de oitavas – 750,

1.500, 3.000 e 6.000 Hz), sempre que houver uma diferença maior ou igual a 20dB entre duas frequências vizinhas (oitavas).<sup>(\*)</sup>

### V.5 – Audiograma

O audiograma é a representação gráfica mundial das informações do exame audiométrico, contendo os limiares tonais no eixo das ordenadas. No eixo das abscissas fica disposta a escala de frequências para as quais é conduzida a pesquisa dos limiares tonais.

Com o audiograma, torna-se mais fácil a visualização e interpretação das informações referentes às perdas auditivas. No mesmo audiograma são representados os limiares tonais das duas orelhas. A simbologia utilizada para sinalizar a perda auditiva por via aérea da orelha direita é um círculo e da orelha esquerda um “X”. Para indicar os resultados por via óssea, utiliza-se o sinal de menor para a orelha direita e de maior para a orelha esquerda. Estes símbolos estão ilustrados na **Figura V.2**.

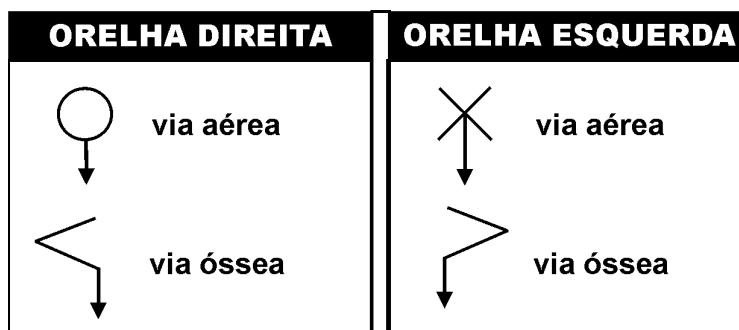


**Figura V.2** – Representação dos símbolos audiométricos

Quando em uma dada frequência não existe um nível de estímulo capaz de sensibilizar a audição, é indicado no audiograma o símbolo referente à orelha em questão e uma flecha para baixo. Esta simbologia é mostrada na **Figura V.3**.

---

<sup>(\*)</sup> Essa nomenclatura será esclarecida no capítulo seguinte



**Figura V.3** – Simbologia usada para indicar no audiograma a falta de estímulo em uma dada frequência

### V.6 – A importância da calibração

É de fundamental importância a calibração do audiômetro a fim de proporcionar maior confiabilidade aos resultados dos testes audiométricos realizados.

O profissional que manuseia o equipamento tem condições de, por si só, executar vários testes básicos do sistema, detectando alguns possíveis problemas nos transdutores: fones, alto-falantes e vibrador ósseo. Conforme explicado por IEDA [RUSSO1999], não é preciso ter um ouvido especial ou absoluto para executar esse tipo de teste, chamado de calibração biológica, por ser feita pelo próprio ouvido humano. Com isso é possível verificar possíveis fontes de mal funcionamento, tais como plugs, tomadas, fios partidos, botões e interruptores quebrados e ruídos provenientes da manipulação do equipamento.

A calibração que é válida tecnicamente, faz uso de um microfone de condensador ligado a um acoplador de 6ml, também chamando de “ouvido artificial”, proporcionando volume semelhante ao ouvido humano quando o fone está colocado ([CORL1953 ] apud [RUSSO1999]).

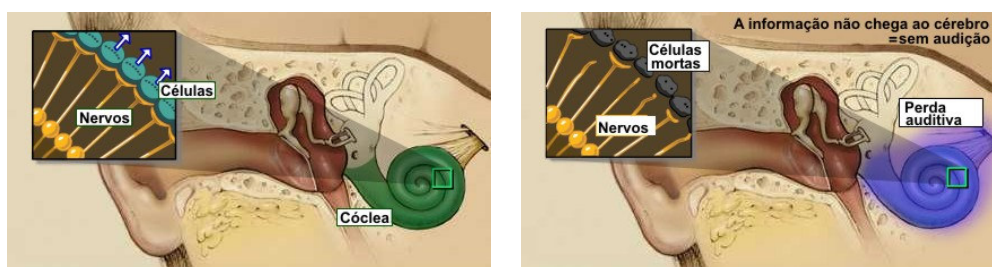
Finalmente, é importante observar que, ao anotar no audiograma os níveis de audição de um indivíduo, deve-se especificar o padrão de calibração ao qual o audiômetro foi submetido, uma vez que os valores de referência para zero dB diferem de padrão para padrão.

## CAPÍTULO VI

### FOCO DA PESQUISA

As perdas auditivas podem apresentar-se em níveis mais leves até mais profundas. Essas perdas podem ser sensoriais ou no canal auditivo. As perdas sensoriais se dão devido a defeitos da cóclea, nervo auditivo ou ambos.

Na **Figura VI.1** está ilustrado uma cóclea normal (a) e uma cóclea com problemas (b). A cóclea é responsável por transformar pressão sonora os estímulos elétricos, ao passo que o nervo auditivo é responsável por conduzir os estímulos elétricos que contém as informações sonoras ao cérebro.

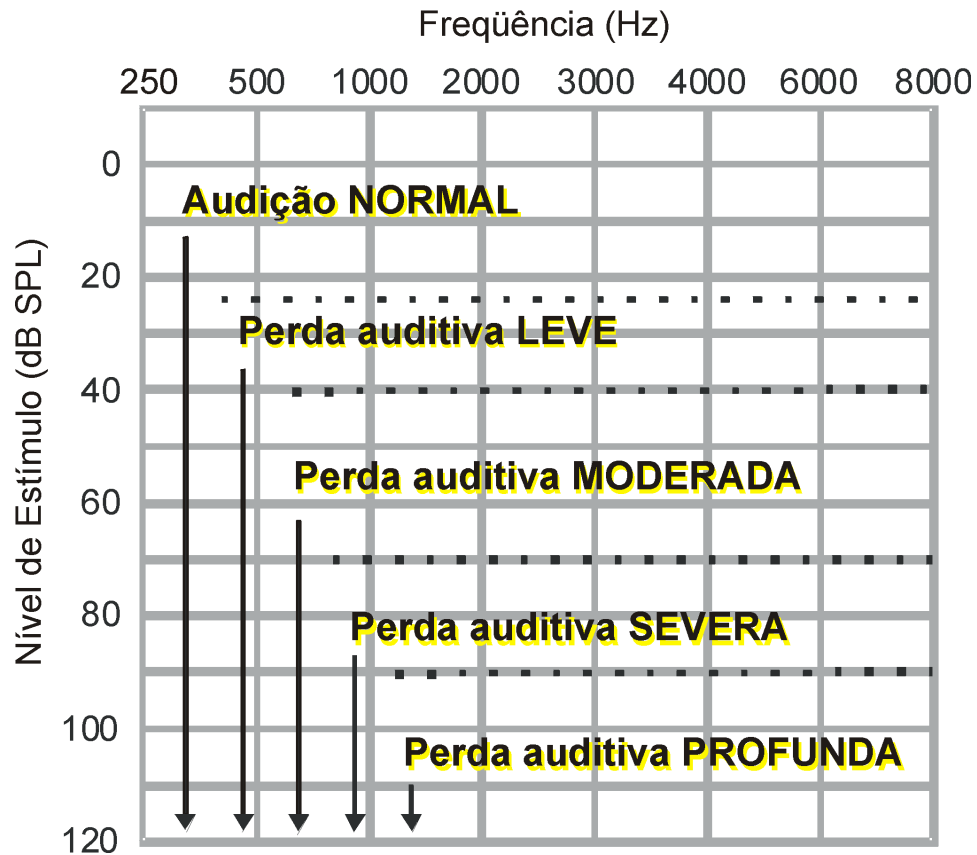


**Figura VI.1** – Sistema Auditivo, cóclea normal (a) e cóclea prejudicada (b)

As perdas auditivas são classificadas conforme o máximo nível de estímulo sonoro necessário para sensibilizar a audição, ou seja, a partir dos limiares auditivos. A classificação do tipo de perda é feita pela média dos limites tonais nas frequências de 500 Hz, 1 kHz e 2 kHz. Assim, se os resultados forem inferiores a 20 dB NA, considera-se a audição como “normal”. Se a média em 500 Hz, 1 kHz e 2 kHz for maior que 20 dB NA e menor que 40 dB NA, a perda é classificada como sendo



“leve”. Para médias entre 40 dB NA e 70 dB NA, o tipo de perda é “moderada”. Entre 70 dB NA e 90 dB NA a perda é do tipo “severa” e médias maiores que 90 dB NA são consideradas como perdas profundas. A classificação das perdas está ilustrada na **Figura VI.2**.

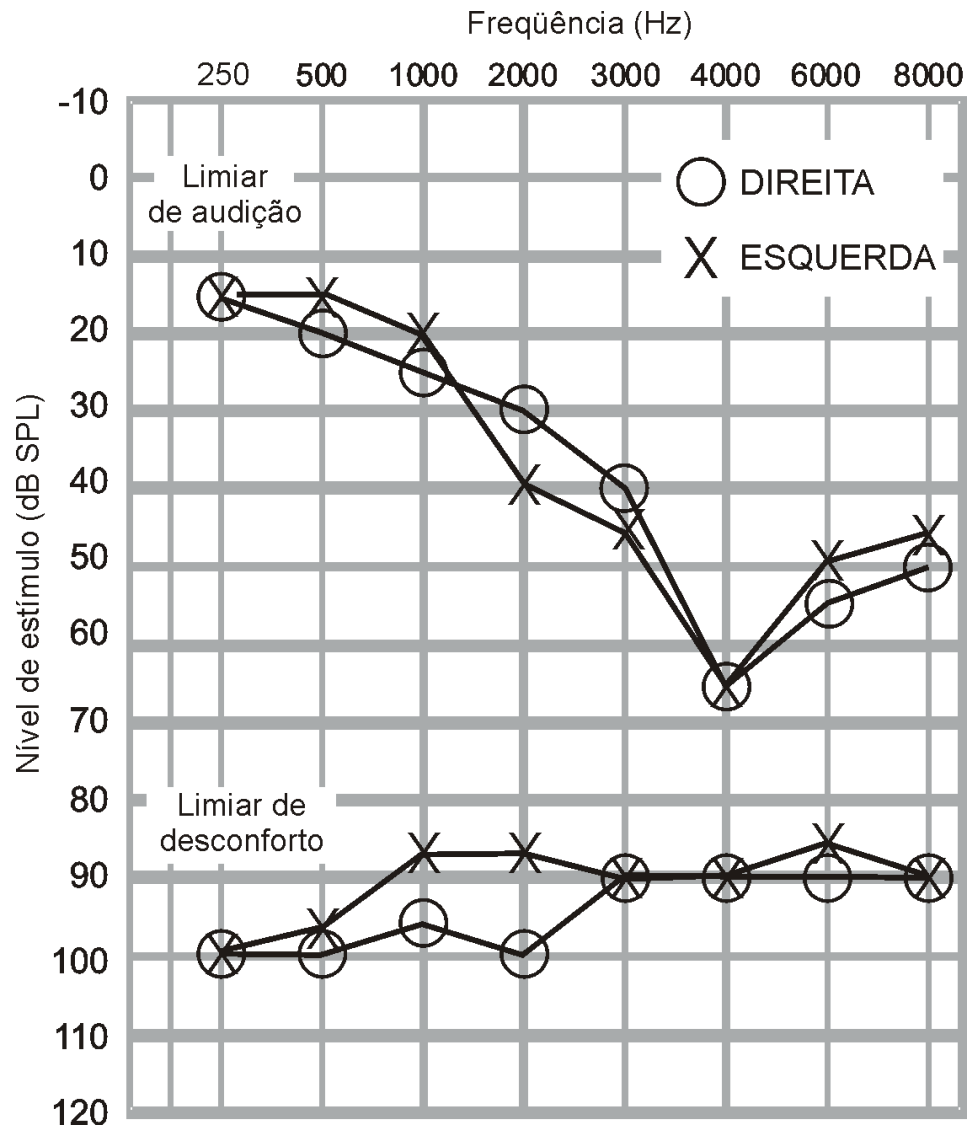


**Figura VI.2** – Classificações da audição quanto ao tipo de perda

Existe um tipo de perda especial denominado de “perda em rampa”, cuja característica é apresentar um resíduo auditivo em baixas frequências, com perdas maiores à medida que a frequência aumenta, e perdas profundas ou até totais em altas frequências.

Na **Figura VI.3** é ilustrado o caso do audiograma de um paciente onde pode ser facilmente visualizado o tipo de perda em rampa, onde as perdas são menores para as baixas frequências e vão aumentando à medida que a frequência é maior. Ainda nesta figura pode ser visto também o “limiar de desconforto”, que são os

níveis de estímulo que causam sensação de desconforto, que antecede por sua vez a sensação de dor (se o nível do estímulo aumentar ainda mais).



**Figura VI.3** – Exemplo de um caso de perda em rampa, adaptado de [ALSAK1996]

Na **Figura VI.4** está ilustrado o audiograma de uma criança residente em Vitória-ES. Foi a partir do conhecimento deste audiograma que teve início o trabalho de pesquisa para usar o processamento digital de áudio na ajuda a portadores de deficiência física que apresentem audição residual em baixas frequências. Este caso

motivou as pesquisas e o desenvolvimento de processamentos capazes de trazer características da fala contidas nas altas frequências para a faixa de frequências com alguma audição residual ainda preservada.

NOME D. V. S. N. DATA 10.2.00

END. \_\_\_\_\_ IDADE \_\_\_\_\_

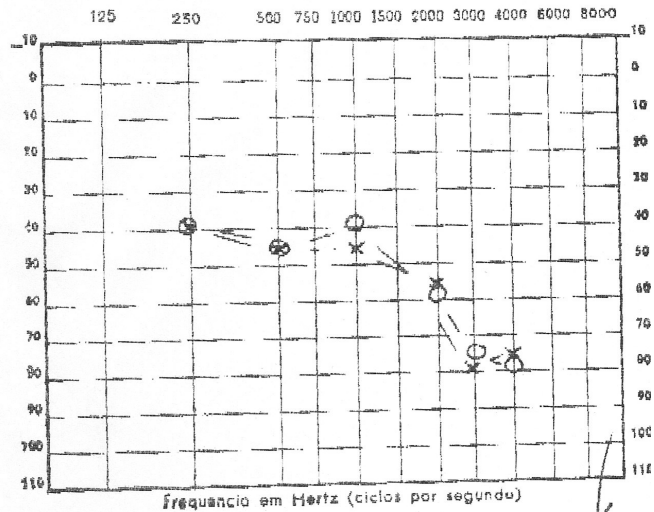
EXAME EM CAMPO LIVRE, COM PRÓTESES DANAVOX GENIUS

**CÓCLEA**

**DR. AZIZ LASMAR**  
 AV. N. S. COPACABANA, 1183 - Sl. 903/904  
 TELG. 522-4092 78/-7448

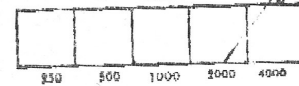
**AUDIOGRAMA**

| TESTE                           | O. E.<br>AZUL | O. D.<br>VERMELHO |
|---------------------------------|---------------|-------------------|
| AÉREA                           | X             | O                 |
| ÓSSEA                           | V             | Λ                 |
| NÃO RE-<br>PONDEU               | ←X            | →O                |
| ÓSSEA MAS<br>CARADA             | ∇             | Δ                 |
| LUSCHER                         | ▲             | ▴                 |
| IFRGER                          | △             | ▽                 |
| ACUPLINO-<br>METRIA             | ▽             | △                 |
| LIMAR DE<br>IRKILABILI-<br>DADE | △             | ▽                 |
| LIMAR DO<br>LOROSO              | ▴             | ▲                 |



Médias: O. E. 48dB O. D. 48dB

Discriminação



**Figura VI.4 – Audiograma de uma criança residente em Vitória/ES**

## CAPÍTULO VII

# MÉTODOS DE REBAIXAMENTO DE FREQUÊNCIAS

### VII.1 – Introdução

Entre as perdas auditivas podemos encontrar uma grande quantidade de casos de perdas em rampa. Muitas pesquisas foram realizadas com o intuito de tentar de alguma forma trazer as informações sonoras contida nas altas frequências para a faixa de frequências onde a perda auditiva ainda pode ser compensada por meio de amplificação.

Em um primeiro momento, poderia-se pensar que qualquer tipo de perda, por mais profunda que ela seja, poderia ser compensada por uma adequada amplificação da potência sonora nas frequências correspondentes. No entanto, esse raciocínio simples não pode ser aplicado na prática, devido à existência do limiar de desconforto.

Para exemplificar, consideremos o caso apresentando na **Figura V.3**. Para a frequência de 4.000 Hz, a perda auditiva é de 65 dB, ao passo que o limiar de desconforto para esta mesma frequência é de 90 dB. Isso significa que a máxima amplificação possível para esta frequência é de 25 dB, o que resultaria em uma perda residual de 40 dB. Existem casos mais extremos, onde o limiar de audição coincide com o limiar de desconforto ou, pior ainda, a perda auditiva é completa nas altas frequências.

Durante as últimas cinco décadas, muitas tentativas de rebaixar a frequência foram tentadas, dentre elas podemos citar:

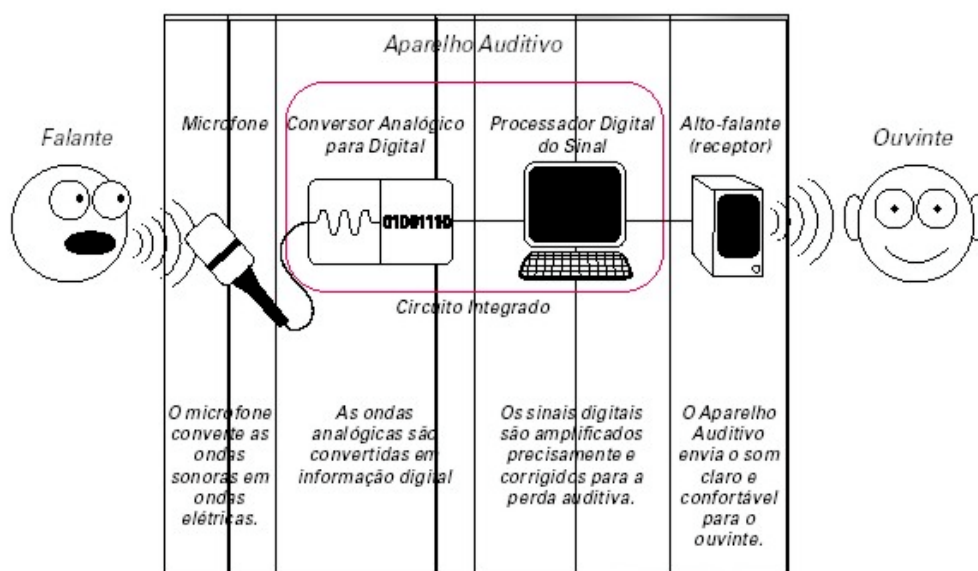
- *Slow playback*: reprodução do som gravado com velocidade menor;
- *Zero-crossing rate division*: divisão da taxa de cruzamento por zero;
- Compressão do espectro de tempo curto do sinal de fala.

Todos esses métodos envolvem algum tipo de distorção devido ao processamento. Dentre eles, foram pesquisados e implementados os métodos que se mostraram mais eficientes, a saber: o método de compressão e a proposição de um novo método que denominamos transladação de frequências e que será detalhado mais adiante.

## VII.2 – Esquema de ajuda digital

A ajuda auditiva digital é composta basicamente dos seguintes estágios: captação da fala, conversão do sinal analógico para digital, processamento digital do sinal e reprodução do sinal processado.

A **Figura VII.1**, a seguir, mostra o “Esquema de Ajuda”. Estão elucidados os estágios descritos acima.



**Figura VII.1** – Esquema em blocos básico da Ajuda Digital Auditiva

### **VII.3 – Métodos de Compressão de Frequências**

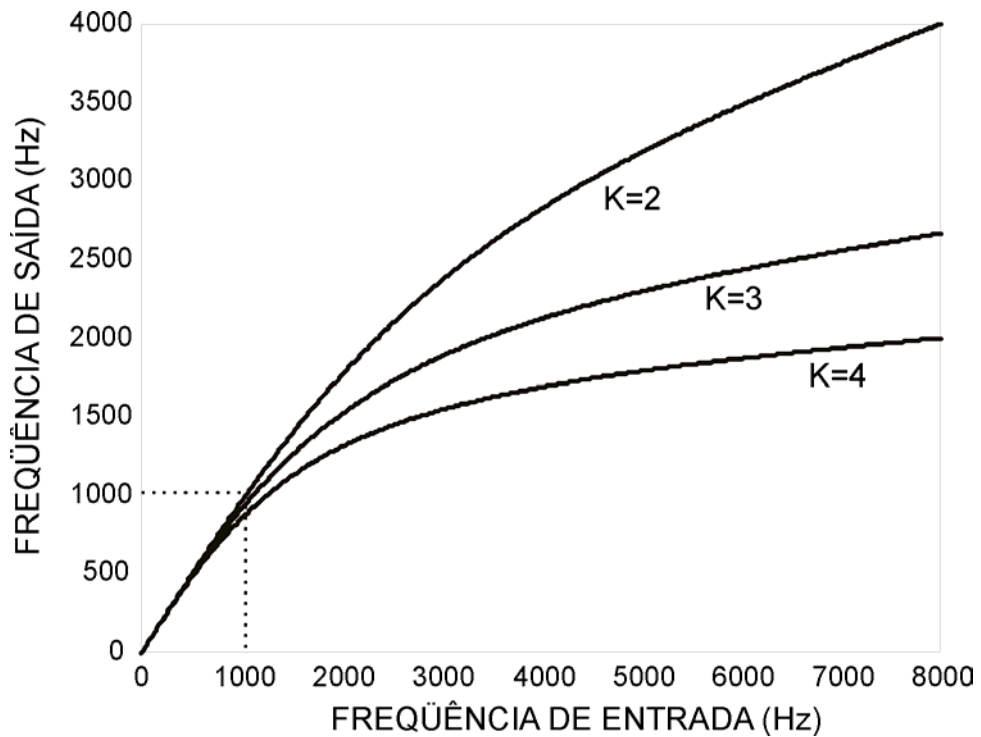
A compressão de frequências é um método de rebaixamento executado por um algoritmo que comprime o espectro de tempo curto do sinal de fala. A compressão pode se dar de duas maneiras diferentes: linear e não linear.

#### **VII.3.1 – Compressão Linear de Frequências**

Este método de compressão mantém a relação entre as frequências de entrada e saída de forma linear. Por exemplo, se o fator de compressão for  $K=2$ , todo o espectro de frequências será reduzido à metade. Assim, se a frequência na entrada for 2 kHz, na saída será 1 kHz. Da mesma forma, se a frequência máxima na entrada for 8 kHz, na saída será 4 kHz, sendo todo o espectro de frequências comprimido da mesma forma.

#### **VII.3.2 – Compressão não-linear de Frequências**

O método de compressão não-linear, também conhecido por método de compressão logarítmica, comprime o espectro de frequências de um quadro (de curta duração) do sinal de fala de forma não linear, realizando menor compressão nas baixas frequências e comprimindo mais as altas frequências. A relação entre a frequência de entrada e a de saída pode ser representada por intermédio de uma curva que relaciona ambas as frequências. Na **Figura VII.2** pode ser visto esta curva de relação para um fator de compressão  $K=2$ ,  $K=3$  e  $K=4$ .



**Figura VII.2** – Relação entre a frequência de entrada e a frequência comprimida

Para determinar a relação entre a frequência de entrada e a frequência de resultante comprimida, a seguinte fórmula foi empregada [REED1983]:

$$\frac{f_{IN}}{f_s} = \frac{1}{\pi} \tan^{-1} \left[ \left( \frac{1-a}{1+a} \right) \tan \left( K\pi \frac{f_{OUT}}{f_s} \right) \right] \quad (VII.1)$$

Onde,  $f_{IN}$  é a frequência de entrada

$f_{OUT}$  é a frequência de saída

$f_s$  é a taxa de amostragem

$K$  é o fator de compressão

$a$  é o fator de distorção.

A fim de obter máxima linearidade nas baixas frequências (abaixo de 1.000 Hz), o fator de distorção foi calculado pela seguinte fórmula:

$$a = \left( \frac{K - 1}{K + 1} \right) \quad (VII.2)$$

Portanto, (VII.1) se reduz a:

$$\frac{f_{IN}}{f_s} = \frac{1}{\pi} \cdot \tan^{-1} \left[ \frac{1}{K} \cdot \tan \left( K \cdot \pi \cdot \frac{f_{OUT}}{f_s} \right) \right] \quad (VII.3)$$

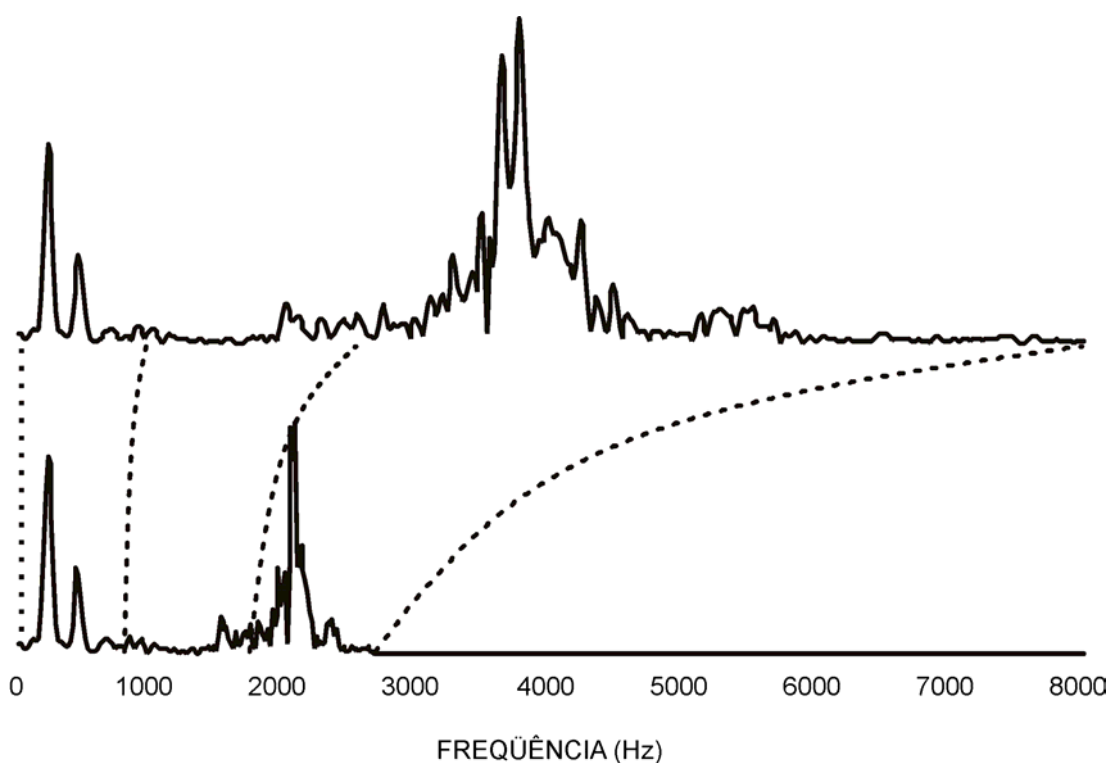
### VII.3.3 – Método de compressão de frequências implementado

O método de compressão de frequências foi implementado relacionando a frequência de saída do quadro de fala com a frequência de entrada de acordo com a *Fórmula VII.3*, que relaciona ambas as frequências. Isolando-se a  $f_{OUT}$  no primeiro membro, temos a relação da  $f_{OUT}$  para uma dada  $f_{IN}$  :

$$f_{OUT} = \left( \frac{f_s}{K \cdot \pi} \right) \cdot \tan^{-1} \left[ K \cdot \tan \left( \frac{\pi \cdot f_{IN}}{f_s} \right) \right] \quad (VII.4)$$

A **Figura VII.3** ilustra um quadro de 25 ms do sinal de fala com seu espectro de frequências original e comprimido.





**Figura VII.3** – Espectro de um quadro de 25 ms do sinal de voz original e comprimido

#### **VII.4 – Método de Transladação de Frequências proposto**

O método original de transladação de frequências proposto nesta dissertação baseia-se na transposição de uma porção da banda de altas frequências para uma banda do espectro de baixas/médias frequências ainda preservada, conforme a perda do deficiente em questão.

##### **VII.4.1 – Cálculo da “Frequência de Destino”**

A frequência de destino para a qual a porção do espectro de altas frequências será transladada, depende tão somente do exame audiométrico apresentado pelo portador de deficiência auditiva para o qual se fará o processamento. Isso porque a faixa de frequências que o portador de deficiências auditivas tem maior habilidade em ouvir varia para cada caso individual.

A partir da maior frequência onde ainda existe audição residual, calcula-se a frequência de destino para a qual a banda de altas frequências do espectro do quadro de fala será transladada.

Primeiramente determina-se a frequência máxima para a qual o limiar auditivo é inferior a 70dB ( $f_{MÁX.PRESERVADA}$ ). Assim, a frequência de destino é empiricamente determinada como sendo a média geométrica dessa “frequência máxima preservada” e 900 Hz, conforme a fórmula a seguir:

$$f_{DESTINO} = \sqrt{900 \cdot f_{MÁX.PRESERVADA}} \quad (\text{VII.5})$$

Onde,

$f_{DESTINO}$  - frequência inicial para a qual o quadro contendo as informações de altas frequências será transladado.

$f_{MÁX.PRESERVADA}$  - é a máxima frequência detectada através do exame audométrico onde a perda é menor ou igual a 70 dB.

É importante frisar que este método de transladação requer a priori que o audiograma considerado apresente perda em rampa, caso contrário não será possível aplicar este tipo de processamento.

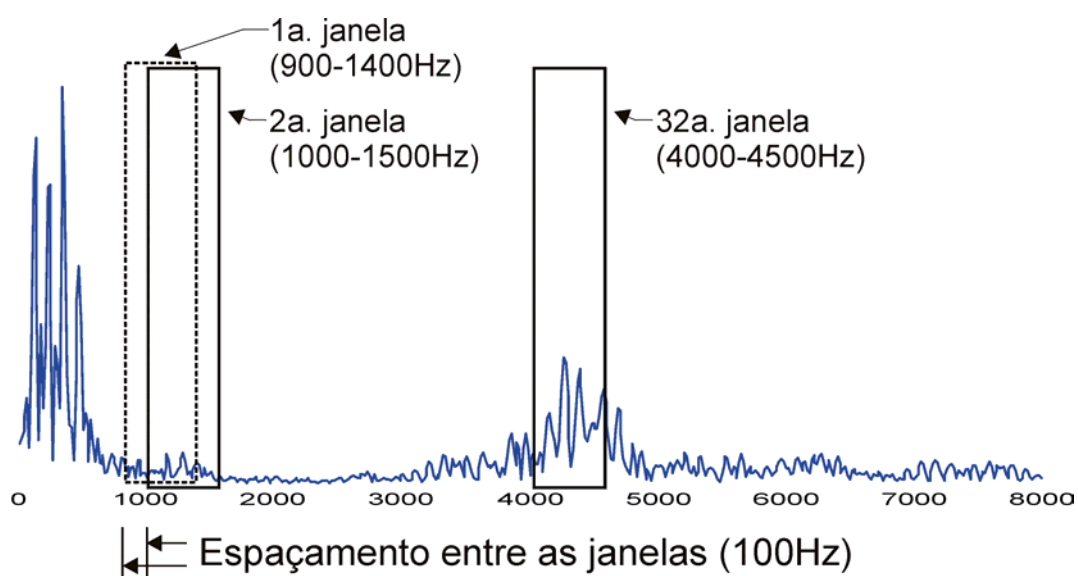
#### VII.4.2 – Cálculo da “Frequência de Origem”

A porção do espectro de frequências a ser transladada depende do quadro do sinal de fala em análise.

O algoritmo implementado faz uma análise do espectro de frequências a fim de determinar qual porção tem maior energia dentro do quadro. Para tal, o algoritmo divide o espectro de frequências do quadro em sub-janelas de 500 Hz cada, espaçadas umas das outras de 100 Hz. Para cada sub-janela é calculada a somatória da energia espectral e o quadro transladado será aquele cuja somatória for maior.

Na **Figura VII.4** está ilustrado o espectro de frequências digital de um quadro de 30 ms de um sinal de fala. A primeira janela começa em 900 Hz, a segunda em

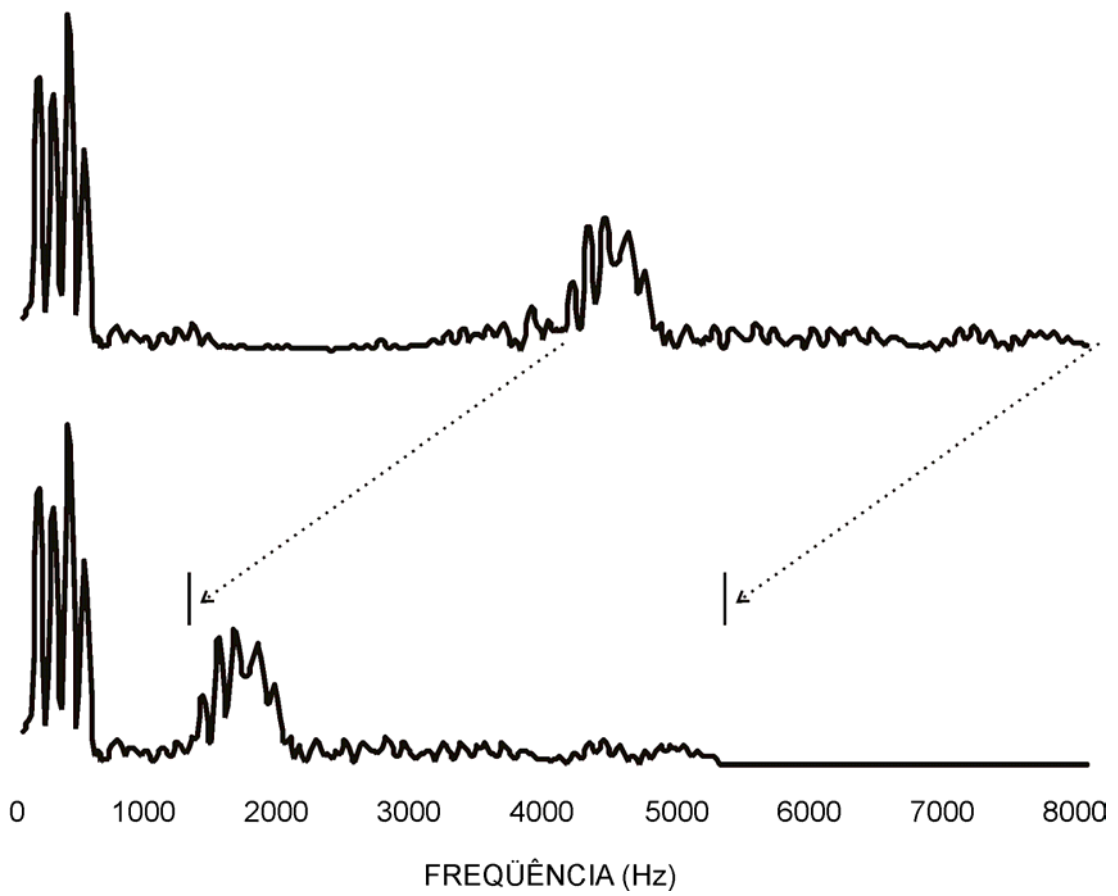
1000 Hz e assim sucessivamente. A somatória de energia é calculada para cada sub-janela de 500 Hz, espaçada da janela seguinte de 100 Hz, conforme descrito. Para o exemplo ilustrado na **Figura VII.4**, o algoritmo detectou que a 32ª janela apresentou a maior somatória de energia espectral, assim determinou-se a frequência de origem como sendo 4000 Hz.



**Figura VII.4** – Ilustração do algoritmo de janelamento para determinação da frequência de origem da porção do espectro de altas frequências a ser transladado

Após determinada a frequência de origem da porção que será transladada, as informações de frequência do espectro são efetivamente transladadas, sendo sobrepostas a partir da frequência de destino previamente calculada com base nos dados audiométricos do paciente.

Na **Figura VII.5** pode ser visualizado o espectro de frequências do quadro do sinal de voz original e o espectro do quadro com as informações de altas frequências transladadas.



**Figura VII.5** – *Espectro do quadro do sinal de voz original e transladado*

## CAPÍTULO VIII

### ESQUEMA DE PROCESSAMENTO IMPLEMENTADO

#### VIII.1 – Processamento quadro-a-quadro

O esquema de processamento implementado executa uma análise quadro-a-quadro. O sinal de áudio a ser processado é dividido em quadros de 25 ms cada. Cada quadro é então analisado no domínio da frequência, por meio da Transformada Rápida de Fourier (FFT – Fast Fourier Transform).

Os quadros são sobrepostos na razão de 50%. A taxa de amostragem utilizada foi de 16.000 amostras por segundo, o que de acordo com o teorema de Nyquist, permite uma frequência máxima de 8 kHz. Com essa taxa de amostragem, os quadros de 25ms têm 400 amostras cada, com uma sobreposição de 200 amostras, conforme ilustrado na **Figura VIII.1**.

O número total de quadros do sinal de áudio depende do tempo de fala desse sinal, que está relacionado com o número total de amostras contida no sinal. A fórmula a seguir (VIII.1), foi empregada para calcular o número de quadros gerados a partir do sinal de áudio a ser processado.

$$NumQuadros = \left\lceil \frac{QtAmostrasSinal - TamQuadro + TamSobP}{TamSobP} \right\rceil \quad (VIII.1)$$

Sendo que,

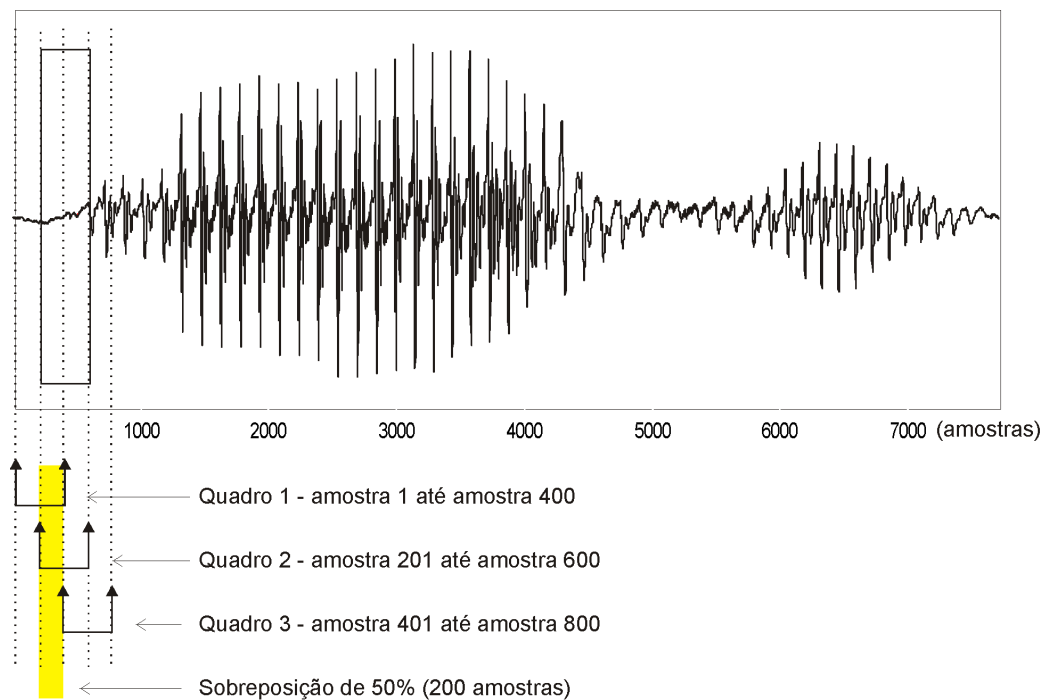
*NumQuadros* – número total de quadros.

$QtAmostrasSinal$  é o número total de amostras contido no sinal de áudio em processamento.

$TamQuadro$  é o tamanho do quadro, ou seja, o número de amostras por quadro, que como citado, foi usado um quadro contendo 400 amostras.

$TamSobP$  é o número de amostras sobrepostas, como a sobreposição foi de 50%, foi empregado 200 amostras de sobreposição.

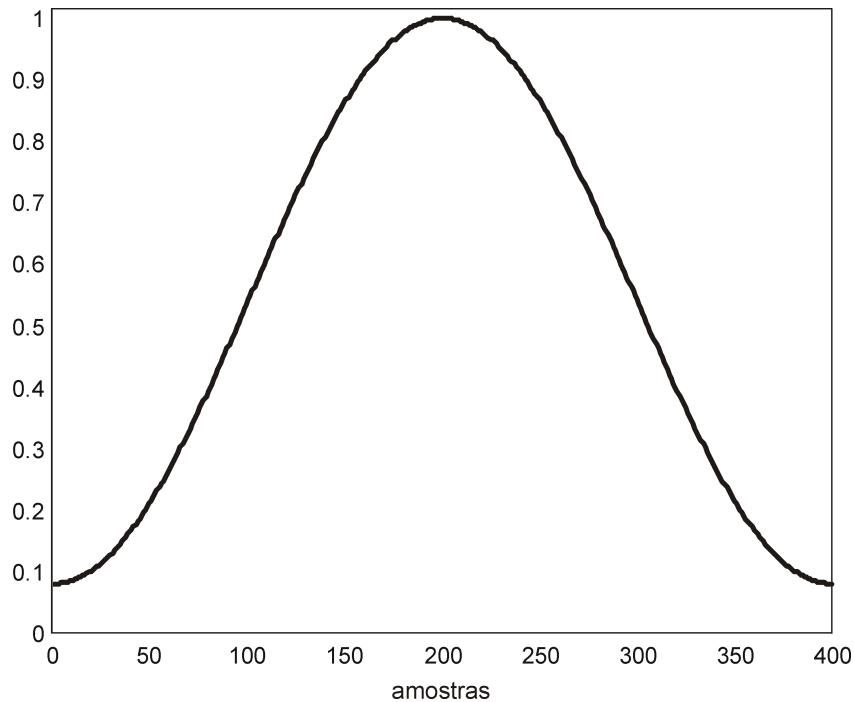
É importante observar também que o valor calculado para o número total de quadros será sempre o próximo valor inteiro natural, conforme a anotação matemática utilizada ( $\lfloor \rfloor$ ).



**Figura VIII.1** – Janelamento do sinal de áudio

### VIII.2 – Janelamento Hamming

Para minimizar o efeito de janelamento de tempo curto, multiplicou-se cada quadro por uma janela Hamming (ilustrada na **Figura VIII.2**) antes da passagem para o domínio da frequência por meio da FFT.



**Figura VIII.2** – *Janela Hamming*

### VIII.3 – Critério discriminativo de processamento do quadro

Conforme relatado anteriormente, o processamento, tanto de compressão quanto de transladação, não produz resultados satisfatórios para os quadros que não apresentam informações relevantes em altas frequências. Portanto, foi criado um critério para discriminar os quadros a serem ou não processados.

Esse critério levou em conta a energia contida nas altas frequências, processando apenas os quadros cuja relação entre a somatória da energia das amostras digitais superiores a 3,5 kHz e a somatória da energia das amostras digitais inferiores a 3,5 kHz fosse superior a um certo limiar. O cálculo desta relação está na *Fórmula (VIII.2)* a seguir:

$$REL_{Faltas/Fbaixas} = \frac{\sum_{3500}^{8000} E(amostras\_quadro)}{\sum_1^{3500} E(amostras\_quadro)} \quad (VIII.2)$$

Analisando-se os fonemas da língua portuguesa que contém informações em altas frequências, determinou-se um limiar capaz de distingui-los. Para determinar este limiar, foram analisados os quadros de vários sinais de fala contendo os fonemas [f] ( f ), [v] ( v ), [ʒ] ( j ), [ʃ] ( x ), [s] ( s ), [z] ( z ), [x] ( r )<sup>(\*)</sup>. A conclusão que se chegou foi que este limiar ficou em torno de 0,3. Assim, os quadros cuja relação entre altas e baixas frequências (VIII.2) foi igual ou superior a 0,3 foram processados, não sofrendo processamento os demais quadros.

#### VIII.4 – Algoritmo implementado

O algoritmo proposto foi implementado usando a ferramenta de processamento matemático MatLab, versão 5.3.

Na implementação do algoritmo de processamento, dois passos principais devem ser considerados: o primeiro consiste na análise dos dados audiométricos próprios de cada deficiente e o segundo passo na análise e processamento quadro-a-quadro (no domínio da frequência) do sinal de áudio. Na **Figura VIII.3** é apresentado o fluxograma geral do sistema de processamento proposto.

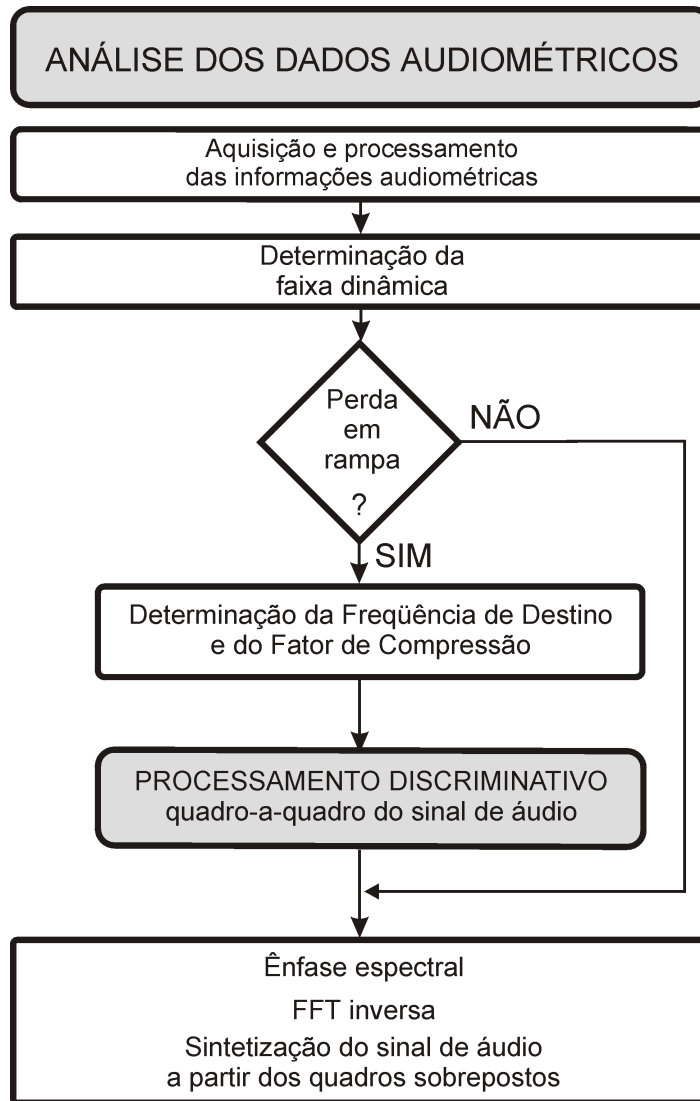
Como mostrado na **Figura VIII.3**, uma análise é feita para determinar se os dados audiométricos indicam uma perda em rampa. Se não for esse o caso, o sinal de áudio não passa pelo processamento de compressão ou transladação, apenas passa pela ênfase espectral (compensação das perdas em cada frequência).

Caso seja detectada a perda em rampa, a frequência de destino para transladação e o fator de compressão K são determinados, pois dependem apenas dos dados audiométricos e independem do sinal de áudio em processamento.

---

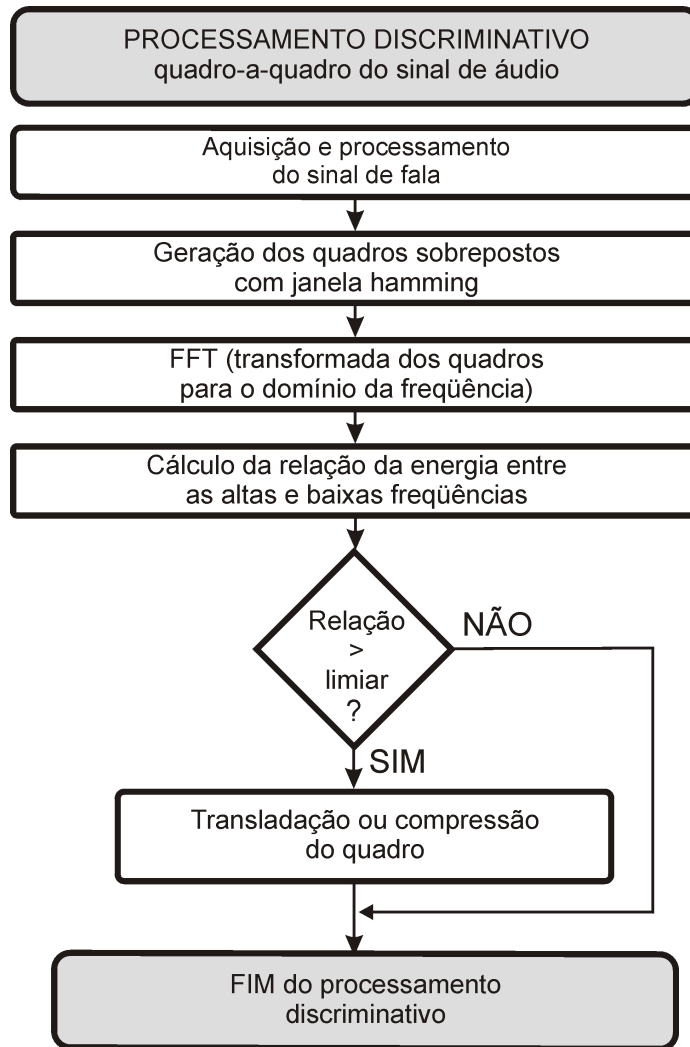
<sup>(\*)</sup> A notação usada para representar os sons segue a nomenclatura do Alfabeto Fonético Internacional (IPA – International Phonetic Alphabet)





**Figura VIII.3** – Fluxograma geral do processamento

A segunda parte do algoritmo é o processamento quadro-a-quadro, que executa a compressão ou transladação propriamente dita, conforme descrito no capítulo anterior. A **Figura VIII.4** mostra o fluxograma do processamento discriminativo quadro-a-quadro.



**Figura VIII.4** – Fluxograma do processamento discriminativo

Segundo o fluxograma da **Figura VIII.4**, o sinal de fala a ser processado é carregado na memória de processamento e dividido em quadros de 25ms cada, conforme explicado na seção VIII.1. Os quadros são multiplicados por janelas hamming para minimizar o efeito do janelamento de tempo curto. A seguir é a FFT dos quadros, gerando novos quadros contendo as amostras digitais no domínio da frequência.

Já no domínio da frequência, é calculada a relação entre as altas e baixas frequências. A partir dessa relação o algoritmo discriminativo determina se o quadro

deve sofrer processamento ou não. O processamento acontece se a relação for superior ao limiar, conforme descrito na seção VIII.3.

Depois do processamento discriminativo, é feita uma ênfase espectral do sinal, com o objetivo de compensar as perdas auditivas amplificando o sinal de áudio até o limite permitido, que é determinado pelo limiar de desconforto<sup>(\*)</sup>.

Após a ênfase espectral, os quadros no domínio da frequência são convertidos de volta para o domínio do tempo. A síntese para recompor o sinal de áudio processado é feita por meio de janelas triangulares de 400 amostras com sobreposição de 50%.

### **VIII.5 – Visualização dos resultados do sinal processado**

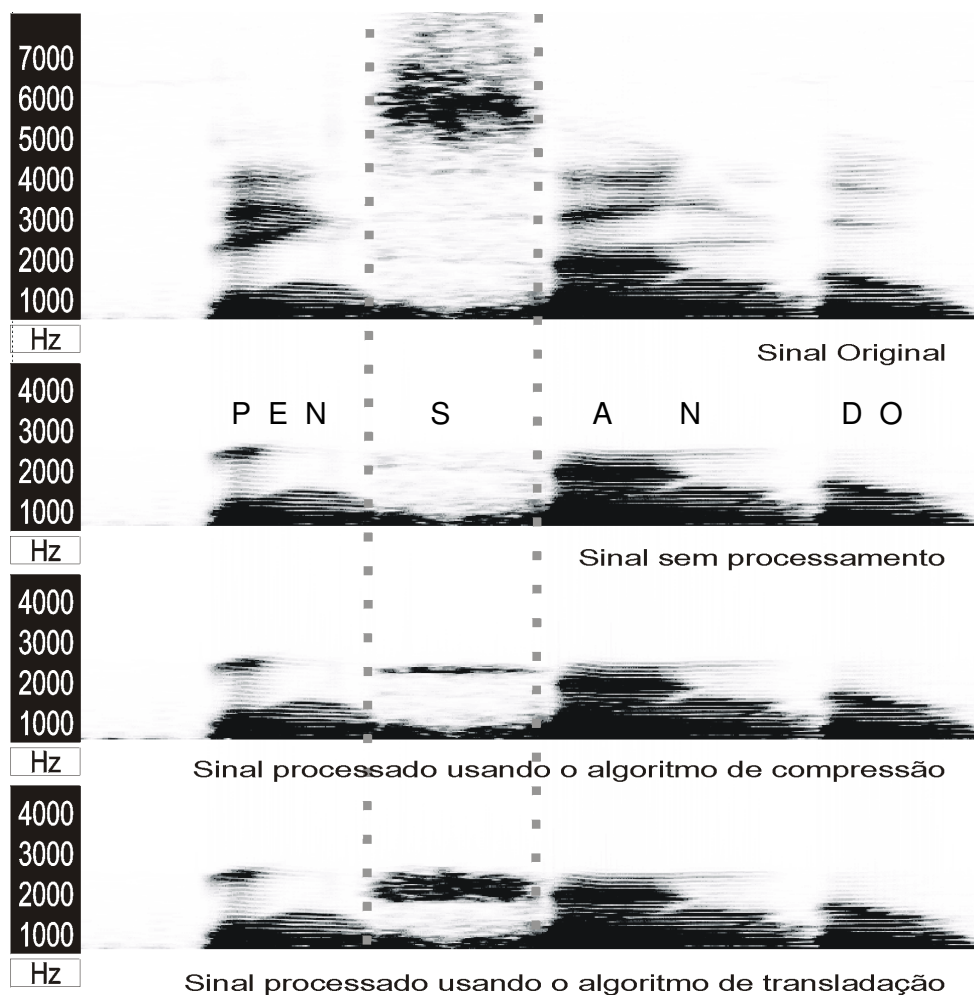
O sinal de áudio submetido ao processamento apresenta as características esperadas no domínio da frequência. Para visualizar esse resultado, faz-se uso do espectograma, que é uma representação gráfica “frequência X tempo”. Pelo espectograma observa-se facilmente a composição de frequências ao longo do sinal de áudio.

Para ilustrar o resultado do processamento, submetemos ao esquema de processamento o sinal de áudio gravado da palavra “pensando”, pronunciada por um orador masculino. Na **Figura VIII.5** estão dispostos os espectogramas do sinal original, do sinal sem processamento, do sinal processado usando o algoritmo de compressão e do sinal processado usando o algoritmo de transladação proposto.

Para o sinal de áudio original, vê-se distribuídas as informações de frequência de 0 a 8 kHz, pois o mesmo foi amostrado a uma taxa de 16 kHz. O sinal sem processamento, na verdade foi submetido a um filtro passa-baixas, para simular as perdas do portador de deficiência. Neste exemplo considerou-se uma frequência máxima preservada de 2,5 kHz, assim, a máxima frequência contida nos sinais processados não pode ser superior a este valor, pois após o processamento os sinais são igualmente filtrados para simular as perdas do deficiente.

---

<sup>(\*)</sup> Caso o limiar de desconforto para cada frequência não conste no audiograma do portador de deficiência, considera-se um limiar de desconforto de 120 dB (NA) para todas as frequências.



**Figura VIII.5** – Espectogramas do sinal de áudio da palavra “pensando”

Entre as linhas pontilhadas está localizado o fonema [s] (s), onde o algoritmo discriminativo atuou. Percebe-se que para o sinal sem processamento, não existem informações de frequência entre 500 e 2500 Hz. Já para os sinais processados, vê-se claramente uma “nuvem” de informação menor para o sinal comprimido e maior para o sinal transladado. Isso comprova que os sinais submetidos ao processamento possuem características dentro da faixa de médias frequências oriundas das altas frequências, o que não acontece no sinal sem processamento.

## CAPÍTULO IX

### TESTES DE INTELIGIBILIDADE

#### IX.1 – Avaliação Discriminativa

O teste de inteligibilidade foi realizado por meio de uma avaliação subjetiva, feita por pessoas com audição normal. As perdas auditivas foram simuladas por meio de filtragem passa-baixas.

O esquema básico do teste consiste na reprodução acústica de duas sílabas fonéticas (usando fones de ouvido). O avaliador, após ter ouvido 3 vezes cada conjunto de sílabas, deve optar por uma dentre as 7 possibilidades que lhe são apresentadas. As sílabas são escolhidas aleatoriamente de um banco de dados formado por 21 diferentes sílabas fonéticas. Cada uma dessas sílabas é composta por um dos 7 sons fricativos mais usados na língua portuguesa ( [ f ], [ v ], [ ʒ ], [ ʃ ], [ s ], [ z ], [ x ] ) e por uma das 3 vogais: [ a ], [ i ] ou [ u ] <sup>(\*)</sup>. Estas sílabas foram pronunciadas uma vez por cada um dos 6 oradores (3 masculinos e 3 femininos), ou seja, o banco de dados original foi formado por 126 locuções. Cada uma destas locuções gerou 9 diferentes arquivos *WAVE* processados: sílaba original, sílaba com frequências comprimidas e sílaba com frequências transladadas, passadas por 3 diferentes filtros passa-baixas, com frequências de corte em 1,5 2 e 2,5 kHz, formando a base de dados de fala final (composta por 1134 arquivos *WAVE*).

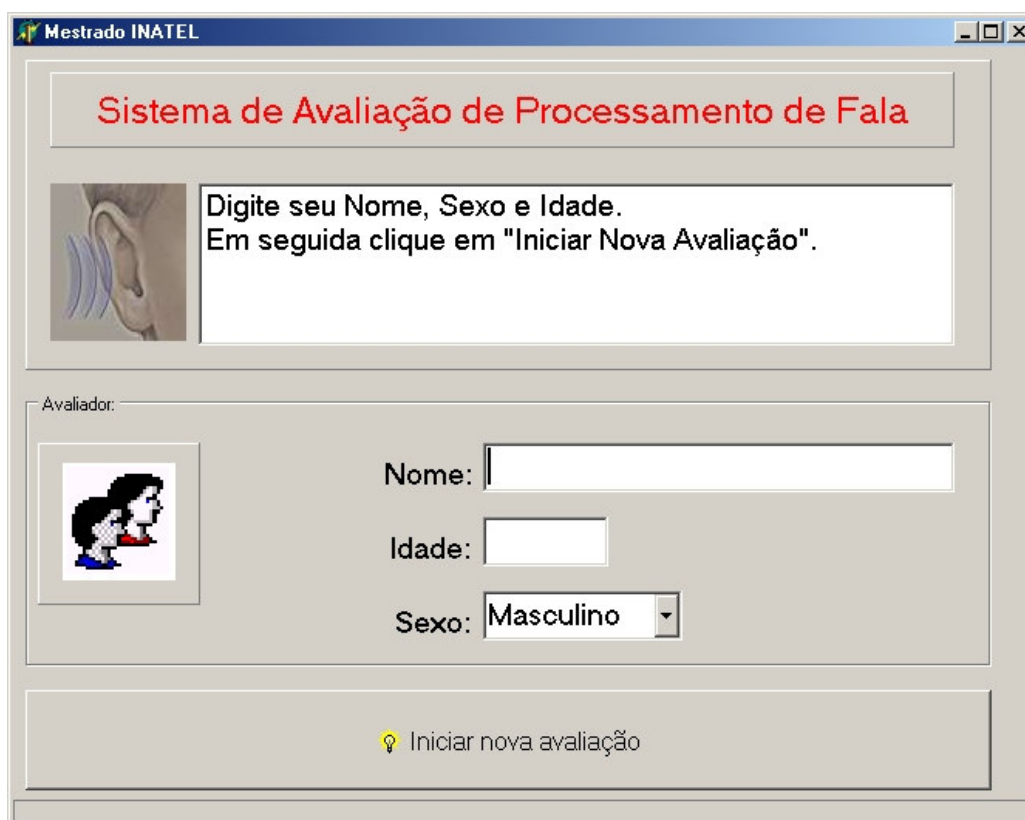
Depois de ter ouvido 3 vezes uma sílaba fonética escolhida aleatoriamente da base de dados final (sem nenhuma informação adicionais sobre estes sons), o ouvinte tem que escolher uma sílaba dentre uma lista de 7 possibilidades. A vogal é a correta

---

<sup>(\*)</sup> A notação usada para representar os sons segue a nomenclatura do Alfabeto Fonético Internacional (IPA – International Phonetic Alphabet)

nestas 7 sílabas, portanto a decisão só será tomada com base nas propriedades acústicas dos sons fricativos processados.

O teste foi implementado na linguagem Delphi, cujo código fonte é o Object Pascal. Na tela de entrada o avaliador se identifica com nome, idade e sexo, cuja finalidade é efetuar o registro dos dados da pessoa, conforme é mostrado na **Figura IX.1**.

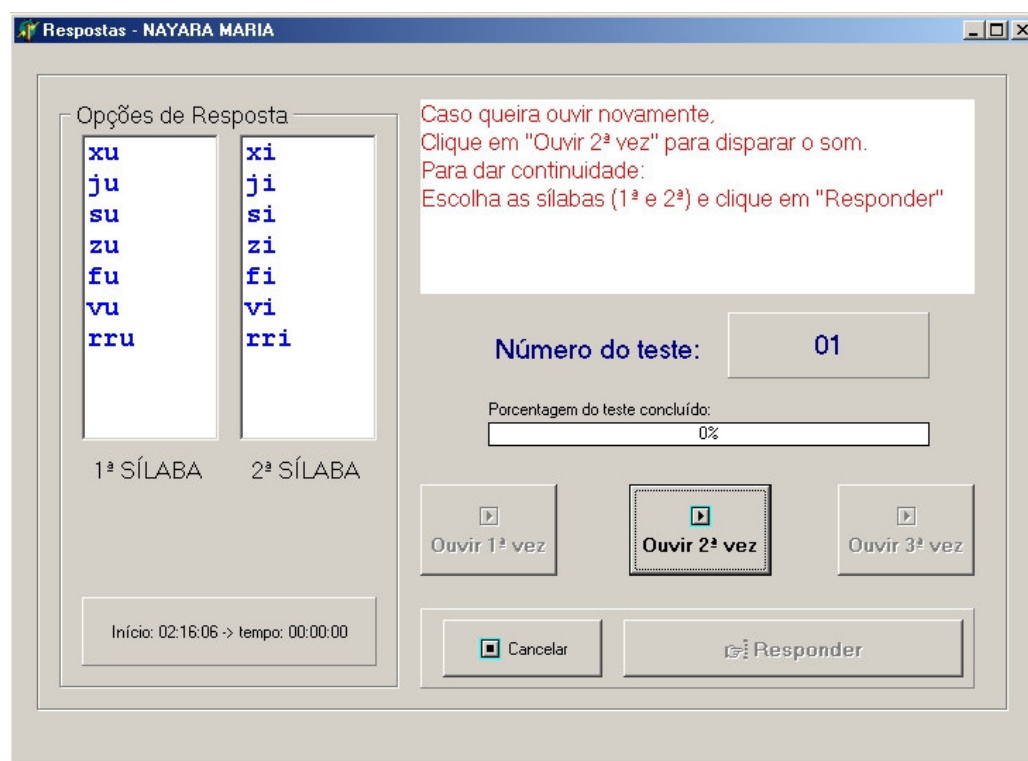


**Figura IX.1** – Tela de Entrada do Programa de Teste de Inteligibilidade

Depois que o avaliador se identifica na tela de entrada, uma segunda tela permite ouvir os arquivos de áudio, sendo obrigatório ouvi-lo 3 vezes, mesmo que as sílabas sejam identificadas na primeira vez. Na **Figura IX.2** é mostrada a tela de respostas no instante em que a lista de possibilidades para as sílabas fica disponível para o avaliador, após o mesmo ter ouvido o arquivo de áudio pela primeira vez. Nesse instante as sílabas de resposta já podem ser escolhidas, mas o botão de

“Responder” só fica disponível após a pessoa ter ouvido o arquivo de áudio pela terceira vez.

Nesta tela existe também uma caixa (branca) contendo instruções sobre como proceder. Logo abaixo desta caixa aparece o número do teste, ou seja, o número da vez em que é apresentada a junção de duas sílabas. Existe ainda uma barra que indica a porcentagem da avaliação que já foi concluída.



**Figura IX.2** – Tela de Respostas do Programa de Teste de Inteligibilidade

O teste implementado foi realizado em duas etapas por dois grupos de pessoas. Para o primeiro grupo, composto de 20 pessoas (5 mulheres e 15 homens), foram apresentadas 18 junções de duas sílabas fonéticas, perfazendo um total de 36 sílabas fonéticas sob avaliação de cada ouvinte. Para o segundo grupo, composto de 22 pessoas (6 mulheres e 16 homens), foram apresentados 27 junções de duas sílabas fonéticas, perfazendo um total de 54 sílabas fonéticas sob avaliação de cada ouvinte.

Os percentuais de acerto do primeiro grupo para cada fonema, tipo de processamento e frequência de corte estão mostrados na *Tabela IX.1*, assim como na *Tabela IX.2* estão os percentuais de acerto do segundo grupo.

Os resultados compilados para ambos os grupos estão na *Tabela IX.3*, onde a coluna *Nenhum* significa que não foi aplicado nenhum processamento ao sinal além da filtragem passa baixas, *Compressão* significa que foi aplicado o processamento de compressão de frequências ao sinal de áudio e *Transladação* significa que foi aplicado o processamento de transladação de frequências ao sinal de áudio. Na primeira coluna temos todas as possibilidades de fricativas para cada um dos filtros nas 3 frequências de corte analisadas. Nas tabelas, os números em **negrito-sublinhado** correspondem à melhor porcentagem de decisões corretas feitas para cada tipo de processamento. Para o primeiro grupo de avaliadores, a escolha das sílabas apresentadas aos ouvintes foi randômica, assim, existiram algumas sílabas que foram menos ouvidas do que outras. Já para o segundo grupo de avaliadores, à medida que os testes foram sendo aplicados, as sílabas foram escolhidas randomicamente dentro do conjunto de sílabas menos escolhidas. Assim, as sílabas foram apresentadas aos ouvintes de forma que todos os fonemas e vogais foram apresentados uma quantidade de vezes aproximadamente igual.



**Tabela IX.1** - Resultados percentuais de reconhecimento do primeiro grupo de ouvintes:

| <i>Fonema<br/>Processado</i> | <i>Freq.<br/>Corte<br/>(Hz)</i> | <i>Processamentos</i> |                             |                     |
|------------------------------|---------------------------------|-----------------------|-----------------------------|---------------------|
|                              |                                 | <i>Nenhum</i>         | <i>Compressão</i>           | <i>Transladação</i> |
| [ f ]<br>( f )               | 1500                            | 61,54                 | <b><u>72,73</u></b>         | 62,50               |
|                              | 2000                            | 50,00                 | 44,44                       | <b><u>71,43</u></b> |
|                              | 2500                            | <b><u>87,50</u></b>   | 50,00                       | 58,33               |
| [ ʒ ]<br>( j )               | 1500                            | 22,22                 | 28,57                       | <b><u>33,33</u></b> |
|                              | 2000                            | 50,00                 | 81,82                       | <b><u>86,67</u></b> |
|                              | 2500                            | 69,23                 | 62,50                       | <b><u>77,78</u></b> |
| [ x ]<br>( r )               | 1500                            | <b><u>57,14</u></b>   | 40,00                       | 12,50               |
|                              | 2000                            | <b><u>46,15</u></b>   | 21,43                       | 38,46               |
|                              | 2500                            | 55,56                 | 38,46                       | <b><u>75,00</u></b> |
| [ s ]<br>( s )               | 1500                            | 53,85                 | <b><u>55,56</u></b>         | 8,33                |
|                              | 2000                            | <b><u>73,33</u></b>   | 36,36                       | 40,00               |
|                              | 2500                            | <b><u>76,92</u></b>   | 41,67                       | 30,77               |
| [ v ]<br>( v )               | 1500                            | 78,57                 | 80,00                       | <b><u>83,33</u></b> |
|                              | 2000                            | <b><u>100,00</u></b>  | 71,43                       | 66,67               |
|                              | 2500                            | 77,78                 | 64,29                       | <b><u>91,67</u></b> |
| [ f ]<br>( x )               | 1500                            | 0,00                  | 20,00                       | <b><u>41,67</u></b> |
|                              | 2000                            | 40,00                 | <b><u>61,11 (11/18)</u></b> | 55,56 (5/9)         |
|                              | 2500                            | 77,78                 | <b><u>100,00</u></b>        | 84,62               |
| [ z ]<br>( z )               | 1500                            | 57,14                 | 50,00                       | <b><u>75,00</u></b> |
|                              | 2000                            | <b><u>70,00</u></b>   | 60,00                       | 40,00               |
|                              | 2500                            | 50,00                 | <b><u>60,00</u></b>         | 33,33               |

**Tabela IX.2** – Resultados percentuais de reconhecimento do segundo grupo ouvintes:

| <i>Fonema<br/>Processado</i> | <i>Freq.<br/>Corte<br/>(Hz)</i> | <i>Processamentos</i> |                             |                             |
|------------------------------|---------------------------------|-----------------------|-----------------------------|-----------------------------|
|                              |                                 | <i>Nenhum</i>         | <i>Compressão</i>           | <i>Transladação</i>         |
| [ f ]<br>( f )               | 1500                            | <b><u>73,68</u></b>   | 50,00                       | 63,16                       |
|                              | 2000                            | 57,89                 | 57,14                       | <b><u>68,42</u></b>         |
|                              | 2500                            | <b><u>63,16</u></b>   | 63,16                       | 55,56                       |
| [ ʒ ]<br>( j )               | 1500                            | 10,53                 | 33,33                       | <b><u>57,89</u></b>         |
|                              | 2000                            | 63,16                 | 70,00                       | <b><u>75,00</u></b>         |
|                              | 2500                            | 73,68                 | <b><u>80,00</u></b>         | 75,00                       |
| [ x ]<br>( r )               | 1500                            | 45,00                 | <b><u>66,67</u></b>         | 42,11                       |
|                              | 2000                            | 57,89                 | 55,56                       | <b><u>68,42</u></b>         |
|                              | 2500                            | 68,42                 | 47,37                       | <b><u>77,78</u></b>         |
| [ s ]<br>( s )               | 1500                            | <b><u>61,11</u></b>   | 47,37                       | 44,44                       |
|                              | 2000                            | <b><u>55,56</u></b>   | 44,44                       | 38,89                       |
|                              | 2500                            | <b><u>50,00</u></b>   | 26,32                       | 30,00                       |
| [ v ]<br>( v )               | 1500                            | 55,56                 | <b><u>65,00</u></b>         | <b><u>65,00</u></b>         |
|                              | 2000                            | <b><u>80,00</u></b>   | 77,78                       | 78,95                       |
|                              | 2500                            | 73,68                 | <b><u>77,78</u></b>         | 75,00                       |
| [ ʃ ]<br>( x )               | 1500                            | 22,22                 | <b><u>41,18</u></b>         | 38,89                       |
|                              | 2000                            | 47,37                 | <b><u>72,22 (13/18)</u></b> | <b><u>72,22 (13/18)</u></b> |
|                              | 2500                            | 78,95                 | <b><u>83,33</u></b>         | <b><u>83,33</u></b>         |
| [ z ]<br>( z )               | 1500                            | 60,00                 | <b><u>63,16</u></b>         | 36,84                       |
|                              | 2000                            | <b><u>77,78</u></b>   | 42,11                       | 26,32                       |
|                              | 2500                            | <b><u>57,89</u></b>   | 26,32                       | 44,44                       |

**Tabela IX.3** – Resultados percentuais de reconhecimento compilados dos dois grupos de ouvintes:

| <i>Fonema Processado</i> | <i>Freq. Corte (Hz)</i> | <i>Processamentos</i> |                             |                             |
|--------------------------|-------------------------|-----------------------|-----------------------------|-----------------------------|
|                          |                         | <i>Nenhum</i>         | <i>Compressão</i>           | <i>Transladação</i>         |
| [ f ]<br>( f )           | 1500                    | <b><u>68,75</u></b>   | 58,62                       | 62,96                       |
|                          | 2000                    | 56,00                 | 53,33                       | <b><u>69,70</u></b>         |
|                          | 2500                    | <b><u>70,37</u></b>   | 57,14                       | 56,67                       |
| [ ʒ ]<br>( j )           | 1500                    | 14,29                 | 31,43                       | <b><u>48,39</u></b>         |
|                          | 2000                    | 57,58                 | 74,19                       | <b><u>80,00</u></b>         |
|                          | 2500                    | 71,88                 | 75,00                       | <b><u>75,86</u></b>         |
| [ x ]<br>( r )           | 1500                    | 48,15                 | <b><u>54,55</u></b>         | 33,33                       |
|                          | 2000                    | 53,13                 | 40,63                       | <b><u>56,25</u></b>         |
|                          | 2500                    | 64,29                 | 43,75                       | <b><u>76,67</u></b>         |
| [ s ]<br>( s )           | 1500                    | <b><u>58,06</u></b>   | 50,00                       | 30,00                       |
|                          | 2000                    | <b><u>63,64</u></b>   | 41,38                       | 39,29                       |
|                          | 2500                    | <b><u>61,29</u></b>   | 32,26                       | 30,30                       |
| [ v ]<br>( v )           | 1500                    | 65,63                 | 70,00                       | <b><u>71,88</u></b>         |
|                          | 2000                    | <b><u>87,50</u></b>   | 76,00                       | 75,00                       |
|                          | 2500                    | 75,68                 | 71,88                       | <b><u>81,25</u></b>         |
| [ ʃ ]<br>( x )           | 1500                    | 14,29                 | 36,36                       | <b><u>40,00</u></b>         |
|                          | 2000                    | 44,83                 | <b><u>66,67 (24/36)</u></b> | <b><u>66,67 (18/27)</u></b> |
|                          | 2500                    | 78,57                 | <b><u>88,00</u></b>         | 83,87                       |
| [ z ]<br>( z )           | 1500                    | <b><u>58,82</u></b>   | 57,14                       | 54,29                       |
|                          | 2000                    | <b><u>75,00</u></b>   | 48,28                       | 31,03                       |
|                          | 2500                    | <b><u>55,17</u></b>   | 37,93                       | 40,74                       |

## IX.2 – Discussão dos resultados

As variações de percentuais de decisões corretas tidas entre o primeiro e o segundo grupo foram mais perceptível nos casos em que o espaço amostral verificado para o primeiro grupo foi pequeno, em alguns casos de até 5 pontos percentuais devido à escolha randômica, como mencionado anteriormente neste capítulo. Assim, uma análise mais confiável deve ser feita sobre os resultados compilados dos dois grupos, pois o espaço amostral é maior (*Tabela IX.3*).

Estes resultados são de difícil análise se considerarmos o conjunto de sílabas como um todo. Mas é interessante analisar cada som fricativo em particular. Por exemplo, podemos concluir dos resultados que para os fonemas [ *ʃ* ] ( *s* ) e [ *ʒ* ] ( *z* ), o melhor é não executar nenhum processamento com eles, mas se considerarmos o caso do fonema [ *ʃ* ] ( *x* ) concluímos justamente o contrário: quando a mais alta frequência audível foi de 2 kHz, a ausência de processamento conduziu a uma porcentagem de inteligibilidade de 44,83%, ao passo que tanto para a Compressão como para a Transladação, a porcentagem de acerto foi de 66,67%. No caso do som fricativo [ *ʒ* ] ( *j* ), a melhor solução foi aplicar o algoritmo de transladação de frequências proposto, para qualquer das 3 frequências de corte consideradas.

Os resultados com o fonema [ *x* ] ( *r* ) em todas as três bandas de frequências analisadas mostraram que, aplicando o processamento, a inteligibilidade foi maior. Para a menor largura de banda (frequência de corte igual a 1.500 Hz), o processamento implicou em 6% a mais de inteligibilidade, podendo chegar a 12% para a maior largura de banda (frequência de corte igual a 2.500 Hz).

O fonema [ *v* ] ( *v* ) apresentou uma porcentagem de acertos 6% maior para a menor largura de banda sobre o sinal não processado. E, em todas as demais frequências de corte e tipos de processamento aplicados, bem como para o caso de não haver processamento algum do sinal, a taxa de acertos foi superior a 70%. Assim, também o fonema [ *f* ] ( *f* ) teve taxa de acertos superior a 50%, porém a melhor solução foi manter o sinal sem processamento.

### IX.3 Matriz de confusão

Uma outra maneira de visualizar os resultados é através de uma matriz que relaciona o fonema apresentado com aquele respondido pelo avaliador. Essa análise é mais detalhada, pois além de conter os acertos, indica também qual o tipo de confusão o fonema causou. Para cada um dos 9 tipos de arquivos de áudio testados, foi gerada uma matriz de confusão. Na *Tabela IX.4*, a seguir, estão dispostas as 9 matrizes geradas para cada combinação das 3 frequências de corte possíveis (1.5, 2 e 2.5 kHz) com cada um dos três processamentos sob teste (nenhum, compressão e transladação):

**Tabela IX.4.a** – Matriz de Confusão

*Proc: nenhum; freq. corte: 1500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 68.8 | 0    | 25.9 | 3.2  | 12.5 | 28.6 | 0    |
| j | 0    | 14.3 | 3.7  | 0    | 3.1  | 3.6  | 5.9  |
| r | 6.3  | 0    | 48.1 | 12.9 | 3.1  | 0    | 0    |
| s | 12.5 | 7.1  | 3.7  | 58.1 | 3.1  | 46.4 | 5.9  |
| v | 3.1  | 28.6 | 11.1 | 0    | 65.6 | 0    | 29.4 |
| x | 3.1  | 0    | 3.7  | 22.6 | 0    | 14.3 | 0    |
| z | 6.3  | 50.0 | 3.7  | 3.2  | 12.5 | 7.1  | 58.8 |
|   | 32   | 28   | 27   | 31   | 32   | 28   | 34   |

**Tabela IX.4.b** – Matriz de Confusão

*Proc: compressão; freq. corte: 1500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 58.6 | 2.9  | 30.3 | 25.0 | 10.0 | 22.7 | 0    |
| j | 0    | 31.4 | 3.0  | 3.6  | 0    | 0    | 14.3 |
| r | 10.3 | 2.9  | 54.5 | 0    | 3.3  | 0    | 8.6  |
| s | 27.6 | 0    | 6.1  | 50.0 | 0    | 31.8 | 0    |
| v | 0    | 11.4 | 3.0  | 0    | 70.0 | 0    | 17.1 |
| x | 0    | 0    | 3.0  | 21.4 | 0    | 36.4 | 2.9  |
| z | 3.4  | 51.4 | 0    | 0    | 16.7 | 9.1  | 57.1 |
|   | 29   | 35   | 33   | 28   | 30   | 22   | 35   |

**Tabela IX.4.c** – Matriz de Confusão

*Proc: transladação; freq. corte: 1500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 63.0 | 0    | 40.7 | 33.3 | 0    | 20.0 | 0    |
| j | 3.7  | 48.4 | 0    | 0    | 6.3  | 6.7  | 25.7 |
| r | 3.7  | 0    | 33.3 | 3.3  | 6.3  | 0    | 5.7  |
| s | 18.5 | 3.2  | 11.1 | 30.0 | 3.1  | 26.7 | 2.9  |
| v | 0    | 6.5  | 0    | 3.3  | 71.9 | 3.3  | 8.6  |
| x | 7.4  | 0    | 11.1 | 26.7 | 0    | 40.0 | 2.9  |
| z | 3.7  | 41.9 | 3.7  | 3.3  | 12.5 | 3.3  | 54.3 |
|   | 27   | 31   | 27   | 30   | 32   | 30   | 35   |

**Tabela IX.4.d** – Matriz de Confusão

*Proc: nenhum; freq. de corte: 2000:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 56.0 | 0    | 21.9 | 3.0  | 0    | 13.8 | 0    |
| j | 0    | 57.6 | 0    | 0    | 3.1  | 3.4  | 14.3 |
| r | 0    | 6.1  | 53.1 | 0    | 0    | 3.4  | 0    |
| s | 32.0 | 0    | 9.4  | 63.6 | 0    | 31.0 | 7.1  |
| v | 4.0  | 15.2 | 0    | 0    | 87.5 | 3.4  | 3.6  |
| x | 0    | 3.0  | 9.4  | 33.3 | 0    | 44.8 | 0    |
| z | 8.0  | 18.2 | 6.3  | 0    | 9.4  | 0    | 75.0 |
|   | 25   | 33   | 32   | 33   | 32   | 29   | 28   |

A diagonal principal de cada matriz contém a porcentagem de acertos de cada fonema. Na última linha está a quantidade de vezes total que cada fonema foi apresentado para avaliação discriminativa.

**Tabela IX.4.e – Matriz de Confusão**

*Proc: compressão; freq. corte: 2000:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 53.3 | 0    | 37.5 | 3.4  | 4.0  | 5.6  | 0    |
| j | 3.3  | 74.2 | 6.3  | 10.3 | 8.0  | 2.8  | 31.0 |
| r | 6.7  | 0    | 40.6 | 0    | 0    | 0    | 3.4  |
| s | 26.7 | 0    | 12.5 | 41.4 | 4.0  | 22.2 | 0    |
| v | 3.3  | 6.5  | 0    | 0    | 76.0 | 0    | 17.2 |
| x | 3.3  | 0    | 0    | 44.8 | 0    | 66.7 | 0    |
| z | 3.3  | 19.4 | 3.1  | 0    | 8.0  | 2.8  | 48.3 |
|   | 30   | 31   | 32   | 29   | 25   | 36   | 29   |

**Tabela IX.4.f – Matriz de Confusão**

*Proc: transladação; freq. corte: 2000:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 69.7 | 0    | 18.8 | 10.7 | 3.6  | 14.8 | 0    |
| j | 0    | 80.0 | 3.1  | 3.6  | 3.6  | 3.7  | 41.4 |
| r | 0    | 0    | 56.3 | 0    | 0    | 0    | 0    |
| s | 15.2 | 0    | 3.1  | 39.3 | 3.6  | 14.8 | 3.4  |
| v | 3.0  | 0    | 6.3  | 0    | 75.0 | 0    | 24.1 |
| x | 12.1 | 0    | 0    | 46.4 | 0    | 66.7 | 0    |
| z | 0    | 20.0 | 12.5 | 0    | 14.3 | 0    | 31.0 |
|   | 33   | 35   | 32   | 28   | 28   | 27   | 29   |

**Tabela IX.4.g – Matriz de Confusão**

*Proc: nenhum; freq. de corte: 2500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 70.4 | 0    | 25.0 | 6.5  | 2.7  | 0    | 0    |
| j | 0    | 71.9 | 0    | 0    | 2.7  | 3.6  | 13.8 |
| r | 0    | 0    | 64.3 | 0    | 0    | 0    | 0    |
| s | 11.1 | 3.1  | 0    | 61.3 | 0    | 17.9 | 3.4  |
| v | 7.4  | 12.5 | 3.6  | 0    | 75.7 | 0    | 20.7 |
| x | 11.1 | 0    | 7.1  | 32.3 | 0    | 78.6 | 6.9  |
| z | 0    | 12.5 | 0    | 0    | 18.9 | 0    | 55.2 |
|   | 27   | 32   | 28   | 31   | 37   | 28   | 29   |

**Tabela IX.4.h – Matriz de Confusão**

*Proc: compressão; freq. corte: 2500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 57.1 | 3.6  | 21.9 | 16.1 | 9.4  | 4.0  | 0    |
| j | 2.9  | 75.0 | 0    | 3.2  | 6.3  | 8.0  | 34.5 |
| r | 0    | 0    | 43.8 | 0    | 0    | 0    | 0    |
| s | 28.6 | 3.6  | 3.1  | 32.3 | 0    | 0    | 0    |
| v | 0    | 17.9 | 12.5 | 0    | 71.9 | 0    | 20.7 |
| x | 11.4 | 0    | 6.3  | 45.2 | 0    | 88.0 | 6.9  |
| z | 0    | 0    | 12.5 | 3.2  | 12.5 | 0    | 37.9 |
|   | 35   | 28   | 32   | 31   | 32   | 25   | 29   |

**Tabela IX.4.i – Matriz de Confusão**

*Proc: transladação; freq. corte: 2500:*

|   | f    | j    | r    | s    | v    | x    | z    |
|---|------|------|------|------|------|------|------|
| f | 56.7 | 0    | 20.0 | 12.1 | 6.3  | 0    | 0    |
| j | 0    | 75.9 | 3.3  | 3.0  | 0    | 0    | 40.7 |
| r | 0    | 0    | 76.7 | 0    | 3.1  | 3.2  | 0    |
| s | 26.7 | 0    | 0    | 30.3 | 0    | 12.9 | 3.7  |
| v | 3.3  | 10.3 | 0    | 0    | 81.3 | 0    | 14.8 |
| x | 6.7  | 6.9  | 0    | 54.5 | 3.1  | 83.9 | 0    |
| z | 6.7  | 6.9  | 0    | 0    | 6.3  | 0    | 40.7 |
|   | 30   | 29   | 30   | 33   | 32   | 31   | 27   |

## **CAPÍTULO X**

### **CONCLUSÕES**

A diferença entre o sinal original (com frequências até 8 kHz) e o sinal filtrado (com frequências até 2 kHz) é grande. Porém, a melhor discriminação de consoantes fricativas observada quando os sinais foram submetidos aos algoritmos de rebaixamento de frequências pode ser uma solução para melhorar a audição de pessoas com deficiência, que nunca tiveram nenhuma percepção de sons com frequências acima de 2 kHz.

Relativamente aos resultados do teste de inteligibilidade, podemos concluir que se incorporássemos um algoritmo de classificação automática de fonemas no sistema, seria possível escolher o melhor algoritmo de rebaixamento de frequências para ser aplicado em cada fonema específico, dada a máxima frequência de audição residual. Isto não é difícil de fazer, considerando os avanços observados no desempenho dos algoritmos automáticos de reconhecimento de fonemas dos últimos anos.

Com todos os processamentos sendo feitos no domínio da frequência, ambos os algoritmos têm demonstrado serem bastante rápidos para possibilitar o uso de aplicações em tempo real.

Alguns aparelhos auditivos fabricados atualmente já fazem uso dos recursos de processamento digital de sinais. Isto possibilita a implementação dos métodos de rebaixamento propostos em aplicações práticas, permitindo o uso dos benefícios proporcionados pelas técnicas apresentadas.



Relativamente à alternativa cirúrgica do implante de cóclea, os aparelhos auditivos digitais possuem a grande vantagem de ser uma técnica não evasiva, evitando assim os riscos de uma intervenção cirúrgica no ouvido interno, que pode até mesmo levar a óbito alguns pacientes.

No atual estágio de implementação em que se encontram as técnicas de processamento propostas neste trabalho, já é possível fazer uso do sistema de rebaixamento de frequências para testes clínicos, permitindo que os deficientes possam testar os sinais processados a fim de melhor ajustar os algoritmos para cada caso.

Finalmente, é importante salientar que, com o passar do tempo, a audição humana tende a perder a capacidade auditiva em altas frequências. Dessa forma, o trabalho desenvolvido constitui um passo na direção de aumentar as possibilidades de proporcionar uma melhor audição para os idosos em geral, e conseqüentemente uma melhoria significativa na sua qualidade de vida.

---

## REFERÊNCIAS BIBLIOGRÁFICAS:

- [HICKS1981] HICKS, B. L., BRAIDA, L. D., DURLACH, N. I.; *Pitch invariant frequency lowering with nonuniform spectral compression*, Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (IEEE New York), pp. 121-124, 1981.
- [REED1983] REED, C. M., HICKS, B. L., BRAIDA L. D., DURLACH, N. I., *Discrimination of speech processed by low-pass filtering and pitch-invariant frequency lowering*, J. Acoust. Soc. Am, vol. 74, pp. 409-419, 1983.
- [REED1985] REED, C. M., SCHULTZ, K. I., BRAIDA, L. D., DURLACH, N. I., *Discrimination and identification of frequency-lowered speech in listeners with high-frequency hearing impairment*, J. Acoust. Soc. Am, vol. 78, pp. 2139-2141, 1985.
- [NELSO1998] NELSON, P., REVOILE, S., *Detection of spectral peaks in noise: Effects of hearing loss and frequency regions*, J. Acoust. Soc. Am, 1998.
- [MUÑOZ1999] MUÑOZ, C. M. Aguilera, PEGGY, B. N., RUTLEDGE, J. C., GAGO, A., *Frequency lowering processing for listeners with significant hearing loss*, IEEE, pp. 741-744, 1999.
- [FROTA2001] FROTA, S., *Fundamentos em Fonoaudiologia*, 1ª ed., vol. 1. Guanabara Koogan, 2001, pp. 40-59.
- [ALSAK1996] ALSAKA, Y. A., MCLEAN, B., *Spectral Shaping for the Hearing Impaired*, IEEE, pp. 103-106, 1996.

- [TEJER1998] TEJERO, J. C., PEGGY, B. N., RUTLEDGE, J. C., *Combination compression and linear gain processing for digital hearing aids*, IEEE, pp. 3140-3143, 1998.
- [NORTH1989] NORTHERN, J. L., DOWNS, M. P., *Audição em crianças*, 3ª ed., Editora Manole Ltda., pp. 7, 1989.
- [PEREI1997] PEREIRA, L. D., SCHOCHAT, E., *Processamento auditivo central – manual de avaliação*, Editora Louise, pp. 99-109, 1997.
- [RUSSO1999] RUSSO, I. C. P., *Acústica e Psicoacústica aplicados à Fonoaudiologia*, 2ª ed., Editora Louise, pp. 203-211, 1999.
- [GUYTO1981] GUYTON, A.C. *Fisiologia Humana*, 5ª ed., Rio de Janeiro, Ed. Interamericana, 1981.