

Universidade Federal do Rio de Janeiro

Escola Politécnica

Departamento de Eletrônica e de Computação

**Ferramentas para Realização de Testes Psicoacústicos -
Aspectos de Percepção do Som**

Autor:

Isabela Ferrão Apolinário

Orientador:

Prof. Luiz Wagner Pereira Biscainho, D. Sc.

Examinador:

Prof. Marcello Luiz Rodrigues de Campos, Ph. D.

Examinador:

Leonardo de Oliveira Nunes, M. Sc.

DEL

Setembro de 2013

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Escola Politécnica - Departamento de Eletrônica e de Computação

Centro de Tecnologia, bloco H, sala H-217, Cidade Universitária

Rio de Janeiro - RJ CEP 21949-900

Este exemplar é de propriedade da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazenar em computador, microfilmear ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es) e do(s) orientador(es).

DEDICATÓRIA

Dedico este trabalho à minha família.

AGRADECIMENTO

Gostaria de agradecer, primeiramente, aos meus pais por todo o carinho e pelo apoio que depositaram em mim.

Agradeço aos meus colegas de turma e muitos outros que passaram pela minha vida nesses cinco anos de faculdade. Cada um contribuiu de alguma forma para que eu chegasse até aqui.

Agradeço aos meus fiéis amigos Carlos Rossi e Allan Freitas, que permaneceram ao meu lado em todos os instantes, não importando a circunstância. Agradeço a Rossi, especialmente, por muitas conversas, às vezes mais francas, às vezes mais descontraídas, mas todas com sua devida importância. Agradeço a Allan, especialmente, por ter me ajudado diversas vezes durante as disciplinas cursadas e por, do seu jeito, se mostrar um verdadeiro amigo.

Agradeço a Carlos Vinícius Campos por ter confiado na minha capacidade desde o princípio e, mais que isso, por ter me feito acreditar em mim nas vezes que pensei que não fosse conseguir.

Agradeço a todos aqueles que aceitaram participar do meu trabalho por meio dos testes subjetivos. Sua ajuda foi fundamental.

Agradeço, por fim, ao meu orientador Luiz Wagner, que esteve sempre disponível para me ajudar. Sua dedicação e amizade me deram forças durante todo o desenvolvimento deste trabalho.

RESUMO

Este trabalho de engenharia elétrica se enquadra na área de processamento de sinais, em particular aplicado ao áudio. Está associado ao estudo de Psicoacústica, que por sua vez inclui elementos de áreas como a Medicina e a Psicologia.

Existem diversas aplicações que utilizam princípios da Psicoacústica modelados por meio de ferramentas de processamento digital de sinais. São exemplos a codificação de áudio, avaliação automática de qualidade, recuperação de informação musical e mesmo composição musical.

Este trabalho trata dos principais conceitos da área de Psicoacústica como audibilidade, *pitch*, JND (menor diferença perceptível ou, do inglês, *Just Noticeable Difference*) e mascaramento. Cada tema é abordado com detalhes e suas aplicações são expostas aqui.

Um pacote de programas foi criado para ilustrar de forma prática a teoria abordada. É possível, com isso, analisar diferentes casos para cada conceito, o que permite ao usuário ter uma ideia mais sólida do seu significado. Um exemplo é a nossa percepção de *pitch*, que pode variar se a frequência fundamental estiver ou não presente no sinal ou, ainda, de acordo com os harmônicos que o compõem. Em um dos programas implementados, o usuário pode selecionar quais harmônicos, incluindo a fundamental, farão parte do sinal a ser escutado.

Os experimentos psicoacústicos elaborados neste trabalho possuem uma finalidade didática, ou seja, procuram complementar a teoria com ilustrações práticas. Além disso, alguns deles também permitem a realização de testes subjetivos. Assim, com o intuito de testar esta funcionalidade, quatro programas foram selecionados. São eles: mascaramento temporal e frequencial, JND de frequência e curvas de Fletcher-Munson. Por fim, os resultados são apresentados e é feita uma análise detalhada dos mesmos.

Palavras-Chave: Psicoacústica, audibilidade, *pitch*, JND, mascaramento, teste subjetivo.

ABSTRACT

This electrical engineering work pertains to the signal (in particular, audio) processing area. It is related to the study of Psychoacoustics, which in turn contains elements from areas such as Medicine and Psychology.

There are many applications that uses the principles of Psychoacoustics modeled by digital signal processing techniques. Some examples are audio codings, automatic quality evaluation, music information retrieval, and even musical composition.

This work is about the main concepts of Psychoacoustics such as loudness, pitch, JND (Just Noticeable Difference), and masking. Each topic is presented in details along with applications.

A package of programs was created to illustrate in a practical way the theory explained herein. They allow to analyze different cases for each concept, allowing the user to have a clearer understanding of its meaning. One example is the way we perceive pitch, which can vary if the fundamental frequency is present in the audio signal or not, or even according to which harmonics the signal is composed of. In one of the programs implemented, the user can choose which harmonics, including the fundamental, will be part of the test signal.

The proposed Psychoacoustics-oriented programs have a didactic purpose, in other words, they complement the related theory with practical examples. Moreover, some of them also allow the execution of subjective tests. They are: temporal and frequency masking, JND of frequency, and Fletcher-Munson curves. At last, the results are presented along with their detailed analysis.

Key-words: Psychoacoustics, loudness, pitch, JND, masking, subjective test.

SIGLAS

UFRJ - Universidade Federal do Rio de Janeiro

JND - *Just noticeable difference*

SPL - *Sound Pressure Level*

ADSR - *Attack, decay, sustain, release*

Sumário

1	Introdução	1
1.1	Tema	1
1.2	Motivação	1
1.3	Objetivos	2
1.4	Alguns Aspectos do Sistema Implementado	3
1.5	Organização deste Trabalho	3
2	Fundamentos de Psicoacústica	5
2.1	Introdução	5
2.2	Fisiologia do Ouvido Humano	6
2.3	Principais Conceitos de Psicoacústica	8
2.3.1	Audibilidade	9
2.3.2	<i>Pitch</i>	11
2.3.3	JND	12
2.3.4	Mascaramento Temporal e Freqüencial	12
2.3.5	Mascaramento Binaural	15
2.3.6	Efeito de Precedência	16
2.3.7	Timbre	17
2.4	Percepção de <i>Pitch</i>	18
2.4.1	Definição de <i>Pitch</i>	19
2.4.2	Sensação de <i>Pitch</i>	19
2.4.3	Teoria do Lugar	20
2.4.4	Teoria do Tempo	20
2.4.5	<i>Pitch</i> de Tons Complexos	21
2.4.6	Tons de Shepard	22

2.4.7	Definições de <i>Pitch</i>	22
2.5	Aplicações de Psicoacústica	23
2.6	Considerações Finais	25
3	Testes Psicoacústicos	26
3.1	Introdução	26
3.2	Experimentos Implementados	27
3.2.1	Invólucro e Calibração	28
3.2.2	Curvas de Fletcher-Munson	29
3.2.3	Atrasa e Soma	31
3.2.4	<i>Pitch</i> sem a Fundamental	34
3.2.5	<i>Pitch</i> e Intensidade	35
3.2.6	<i>Pitch</i> vs. Clique	36
3.2.7	Teoria Dúplex do <i>Pitch</i>	38
3.2.8	JND de Frequência	39
3.2.9	Mascaramento Temporal	41
3.2.10	Mascaramento Frequencial	43
3.2.11	Mascaramento Binaural	45
3.2.12	Efeito de Precedência	47
3.3	Detalhes da Implementação	48
3.4	Considerações Finais	49
4	Testes Subjetivos e Resultados	50
4.1	Introdução	50
4.2	Experimentos Selecionados	50
4.2.1	Calibração	51
4.2.2	Mascaramento Temporal	52
4.2.3	Mascaramento Frequencial	53
4.2.4	JND de Frequência	53
4.2.5	Curvas de Fletcher-Munson	53
4.3	Procedimento	54
4.4	Resultados	56
4.4.1	Mascaramento Temporal	56

4.4.2	Mascaramento Freqüencial	58
4.4.3	JND de Freqüência	61
4.4.4	Curvas de Fletcher-Munson	64
4.5	Considerações Finais	69
5	Conclusões	70
	Bibliografia	72

Lista de Figuras

2.1	Fisiologia do ouvido humano.	7
2.2	Envoltórias de máxima deflexão.	8
2.3	Curvas de audibilidade de Fletcher-Munson.	10
2.4	Mascaramento temporal.	13
2.5	Decaimento no mascaramento temporal.	14
2.6	Decaimento no mascaramento frequencial.	14
2.7	Mascaramento Binaural.	16
2.8	Exemplos de curvas ADSR: flauta e piano.	18
2.9	Regra de Stevens.	21
2.10	Tons de Shepard.	23
3.1	Interface gráfica do invólucro.	28
3.2	Interface gráfica da calibração.	29
3.3	Interface gráfica principal do experimento “Curvas de Fletcher-Munson”.	30
3.4	Diálogo pertencente ao experimento “Curvas de Fletcher-Munson”.	30
3.5	Exemplo de gráfico gerado pelo experimento “Curvas de Fletcher-Munson”.	31
3.6	Diagrama de blocos de um sistema “atrasa e soma”.	32
3.7	Interface gráfica do experimento “Atrasa e Soma”.	32
3.8	Resposta em frequência da saída $y(t)$ do experimento “Atrasa e Soma”.	33
3.9	Interface gráfica do experimento “ <i>Pitch</i> sem a Fundamental”.	34
3.10	Interface gráfica do experimento “ <i>Pitch</i> e Intensidade”.	35
3.11	Exemplo sintético de gráfico gerado pelo experimento “ <i>Pitch</i> e Intensidade”.	36
3.12	Interface gráfica do experimento “ <i>Pitch</i> vs. Clique”.	37
3.13	Interface gráfica do experimento “Teoria Dúplex do <i>Pitch</i> ”.	38
3.14	Interface gráfica principal do experimento “JND de Frequência”.	39
3.15	Diálogo pertencente ao experimento “JND de Frequência”.	40

3.16	Exemplo de gráfico gerado pelo experimento “JND de Frequência”.	41
3.17	Interface gráfica do experimento “Mascaramento Temporal”.	42
3.18	Exemplo de gráfico gerado pelo experimento “Mascaramento Temporal”.	43
3.19	Interface gráfica principal do experimento “Mascaramento Freqüencial”.	44
3.20	Diálogo pertencente ao experimento “Mascaramento Freqüencial”.	44
3.21	Exemplo de gráfico gerado pelo experimento “Mascaramento Freqüencial”.	45
3.22	Interface gráfica principal do experimento “Mascaramento Binaural”.	46
3.23	Diálogo pertencente ao experimento “Mascaramento Binaural”.	47
3.24	Interface gráfica do experimento “Efeito de Precedência”.	48
4.1	Interface gráfica inicial do teste subjetivo.	51
4.2	Interface gráfica principal pertencente ao teste subjetivo.	52
4.3	Ilustração da interface gráfica do experimento “Mascaramento Temporal”.	55
4.4	Exemplos de curvas de decaimento obtidas no experimento “Mascaramento Temporal”.	56
4.5	Efeito da calibração nos resultados do experimento “Mascaramento Temporal”.	57
4.6	Média e desvio padrão dos resultados do experimento “Mascaramento Temporal”.	58
4.7	Exemplos de curvas de decaimento obtidas no experimento “Mascaramento Freqüencial”.	59
4.8	Efeitos do limiar de audibilidade, batimento e calibração nos resultados do experimento “Mascaramento Freqüencial”.	60
4.9	Média e desvio padrão dos resultados do experimento “Mascaramento Freqüencial”.	61
4.10	Exemplos de resultados obtidos no experimento “JND de Frequência”.	63
4.11	Exemplos de resultados particulares obtidos no experimento “JND de Frequência”.	64
4.12	Curvas obtidas com os resultados do experimento “JND de Frequência”.	65
4.13	Exemplos de resultados obtidos no experimento “Curvas de Fletcher-Munson”.	66
4.14	Exemplos de resultados particulares obtidos no experimento “Curvas de Fletcher-Munson”.	67

4.15 Média e desvio padrão dos resultados do experimento “Curvas de Fletcher-Munson”	68
--	----

Capítulo 1

Introdução

1.1 Tema

O tema principal deste trabalho é a Psicoacústica, que é o estudo da percepção do som e seus impactos fisiológicos. Há uma série de fenômenos característicos desse tema, tais como audibilidade, *pitch*, mascaramento e JND (menor diferença perceptível ou, do inglês, *Just Noticeable Difference*).

A audibilidade pode ser definida como a intensidade percebida de um som. Da mesma forma, o *pitch* pode ser definido como a altura (ou frequência) percebida. O mascaramento está relacionado ao fato de não percebermos um som de baixa intensidade na presença de um outro de maior intensidade. A JND, por fim, é a menor diferença que conseguimos detectar em um sinal quando variamos um parâmetro. No caso de áudio, os parâmetros mais comuns para os quais se define uma JND são frequência e intensidade.

1.2 Motivação

A Psicoacústica possui diversas aplicações práticas, dentre elas a codificação de áudio e a avaliação automática de qualidade. As duas técnicas procuram avaliar um dado sinal de áudio num domínio perceptivo, no qual as informações relevantes para a percepção são destacadas. Assim, podemos retirar do sinal tudo aquilo que não é percebido por nós, como é o caso da informação mascarada. Da mesma

forma, podemos avaliar a qualidade de um sinal de áudio comprimido comparando-o, perceptivamente, ao sinal de áudio original (sem compressão).

1.3 Objetivos

O objetivo geral deste trabalho é criar um conjunto de programas com experimentos práticos que ilustrem conceitos importantes dentro da Psicoacústica, chamados experimentos psicoacústicos. Os programas possuem duas finalidades. A primeira delas é permitir a realização sistemática e controlada de testes, chamados testes subjetivos¹. A segunda finalidade é ilustrar conceitos de Psicoacústica com exemplos práticos.

Os objetivos específicos são, dentre outros, aumentar a usabilidade dos programas por meio de interfaces gráficas amigáveis ao usuário; realizar, com o uso de alguns dos programas criados, testes subjetivos para comprovar sua funcionalidade e comparar os resultados obtidos com os teóricos.

Com o pacote de programas, é possível escutar os diversos fenômenos relacionados à Psicoacústica. Assim sendo, este trabalho é do interesse tanto de alunos de engenharia, que procuram entender na prática os conceitos estudados para, por exemplo, aplicá-los em codificação de áudio, como de alunos de música, que possuem interesse em entender mais a fundo os aspectos da percepção auditiva para relacioná-los à sua área.

Há, atualmente, uma grande variedade de material didático disponível sobre o tema Psicoacústica. Para o aprendizado, porém, é importante escutar e perceber por conta própria os efeitos explicados em livros. Nesse sentido, este projeto complementa o ensino no campo da Psicoacústica com um conjunto de programas experimentais. Como já foi dito, estes permitem ao usuário a percepção auditiva dos conceitos e a realização de testes subjetivos.

¹Os testes subjetivos têm como objetivo medir a resposta humana (média) a diferentes estímulos. Dessa forma, o participante deve dizer qual foi a mudança percebida no sinal de saída quando são variados os seus parâmetros de entrada.

1.4 Alguns Aspectos do Sistema Implementado

Os programas implementados utilizam ferramentas de processamento digital de sinais. Os sinais de entrada são constituídos, basicamente, por tons puros (senoides), tons complexos (compostos por uma senoide de frequência fundamental f_0 e outras de frequências múltiplas desta: $2f_0$, $3f_0$, etc.), pulsos triangulares e ruído branco. Os programas oferecem ao usuário a possível alteração de parâmetros desses sinais, tais como a sua frequência, intensidade, duração e número de ciclos, entre outros. Na maioria dos casos, há um botão que toca o sinal alterado pelos parâmetros configurados e um outro que toca um sinal de referência, que dependerá do programa em questão. A maioria dos experimentos exige o uso de fones de ouvido.

Os testes subjetivos foram realizados em uma sala bem isolada de ruídos externos e com fones de ouvido de boa qualidade. Uma pequena introdução esclarecendo a motivação e o funcionamento do teste foi realizada no começo de cada um. Os resultados foram armazenados no término. Todos os resultados parciais obtidos foram utilizados para estimar um resultado final.

Há algumas limitações presentes no projeto, tais como o *software* e o sistema operacional utilizado para rodar os programas implementados, o local onde estes devem ser utilizados e os equipamentos necessários. São essenciais, para rodar os programas aqui apresentados, computadores com o sistema operacional Windows e o *software* Matlab[®] instalados e fones de ouvido de boa qualidade. Para a realização de testes subjetivos, os programas devem ser utilizados, preferencialmente, em estúdios (ambientes com isolamento acústico). Se utilizados somente com finalidade didática, os testes podem ser rodados em ambientes silenciosos.

1.5 Organização deste Trabalho

No Capítulo 2 será feita uma introdução teórica. Os principais temas abordados serão a fisiologia do ouvido humano, os principais conceitos de Psicoacústica e, por fim, as aplicações dessa área.

O Capítulo 3 apresenta os testes psicoacústicos implementados e sua funcionalidade. Cada teste será explicado com detalhes e serão feitas relações com a teoria explicada no Capítulo 2.

Os resultados dos testes subjetivos realizados são apresentados no Capítulo 4. Nele serão mostrados os gráficos gerados e justificativas teóricas para cada resultado obtido.

Por fim, o Capítulo 5 apresenta as conclusões do trabalho.

Capítulo 2

Fundamentos de Psicoacústica

2.1 Introdução

A Psicoacústica estuda como os seres humanos percebem o som externo. Esse estudo engloba todo o caminho realizado pelo som, desde a sua emissão pela fonte sonora, transmissão no espaço físico entre a fonte e o ouvinte, captação e processamento por meio dos receptores sensoriais do ouvinte, transmissão do sinal resultante ao cérebro e, por fim, interpretação.

A percepção do som pode ser, então, dividida em processos ou etapas. O estudo destas etapas é importante para a Psicoacústica, já que, em cada uma delas, o som sofre um tipo de processamento, ou seja, ele é alterado ao longo do caminho. A sua versão final, o sinal resultante que chega ao cérebro, é bem diferente da versão emitida originalmente pela fonte sonora.

O modo como o ambiente—no qual estão presentes o ouvinte e a fonte sonora—altera o sinal emitido está diretamente relacionado ao nosso senso de espacialidade. Somos capazes, por exemplo, de identificar a proximidade de uma fonte em uma sala fechada devido à influência desta no sinal sonoro.

Às vezes, porém, não estamos interessados no senso de espacialidade, ou seja, em como o espaço altera o som. Se o ouvinte estiver utilizando fones de ouvido, o processamento realizado ao longo dos ouvidos externo, médio e interno até o cérebro é o fator crucial para entendermos como é feita a percepção do som. Ao longo deste

caminho, uma análise em frequência do sinal captado é realizada no ouvido interno antes de ser transmitido ao cérebro. Essa análise é um aspecto fundamental para a percepção de *pitch*.

Além da percepção em frequência—altura—, também é importante analisar outros aspectos da nossa percepção, como o temporal e o de intensidade. Por exemplo, para um dado tom, a menor intensidade percebida pode variar de acordo com sua altura; se dois pulsos são tocados em sucessão e um destes estiver suficientemente próximo do outro e com uma intensidade suficientemente pequena, nós deixamos de escutá-lo; se duas notas de instrumentos diferentes são tocadas com a mesma altura—*pitch*—e mesma intensidade, somos, mesmo assim, capazes de identificar a diferença. Esses e outros aspectos serão explicados de forma mais aprofundada neste capítulo.

Além da definição dos conceitos de Psicoacústica, serão explicadas algumas das principais aplicações desse tema, tais como codificação de áudio e síntese de música, e a relação entre elas e os fenômenos envolvidos, tais como mascaramento e percepção de *pitch*.

É importante ressaltar que, apesar de serem apresentadas definições formais para cada conceito, estas não são estritas, já que representam medidas subjetivas. Os valores de *pitch* e audibilidade, por exemplo, podem variar com o gênero e a idade das pessoas.

2.2 Fisiologia do Ouvido Humano

O ouvido humano é dividido em ouvido externo, contendo o pavilhão auditivo e o conduto auditivo; ouvido médio, contendo o tímpano e os ossículos; e ouvido interno, contendo os canais semicirculares, a cóclea e o nervo auditivo. Cada elemento é responsável por uma determinada função. O conjunto destes é que faz o som chegar ao cérebro com o processamento necessário para, então, ser interpretado. Uma ilustração da fisiologia do ouvido humano é dada na Fig. 2.1.

Na primeira etapa, o som é captado pelo pavilhão auditivo. Esse funciona, também, como um filtro direcional [1], ou seja, o sinal captado é alterado de acordo

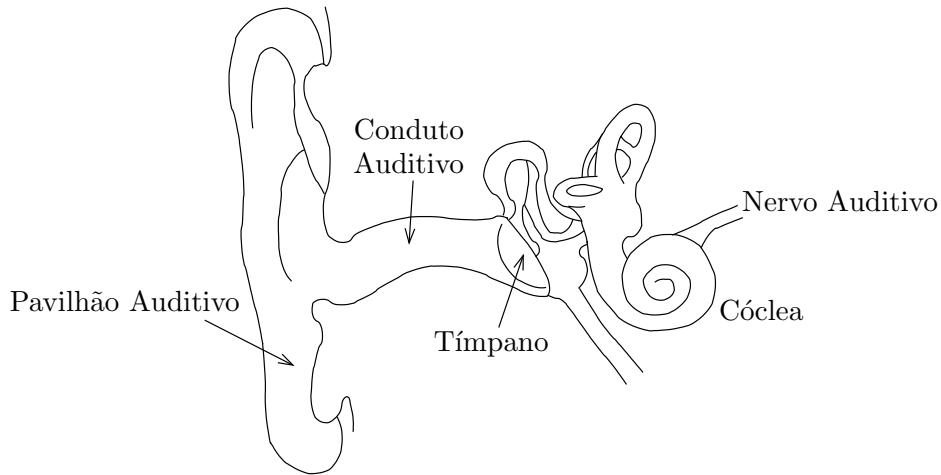


Figura 2.1: Fisiologia do ouvido humano.

com a sua direção de emissão e não somente com sua forma de onda. Além disso, o som ainda sofre efeitos direcionais ao interagir com os ombros, a cabeça e o próprio pavilhão auditivo. Esse efeito é importante para o nosso senso de direita, esquerda, cima e baixo.

A seguir, as ondas sonoras são guiadas para o interior do ouvido por meio do conduto auditivo. O tímpano, presente no começo do ouvido médio, vibra em resposta a estas ondas e os ossículos transferem essas vibrações até a cóclea, presente no ouvido interno.

A cóclea é uma cavidade em formato de caracol, repleta de fluido e composta por três partes: as escalas timpânica, vestibular e média. Possui, também, uma membrana—membrana basilar—que a divide em duas passagens e cujo comportamento pode ser interpretado como uma análise em frequência realizada sobre o sinal captado pelo ouvido. Isso porque, para um determinado estímulo senoidal, uma deflexão ocorre ao longo dessa membrana e a posição de pico da deflexão está diretamente relacionada à frequência de entrada [1]. A Fig. 2.2 ilustra as deflexões ocorridas na membrana basilar em função da distância ao longo desta e da frequência. Na figura, a cóclea e a membrana basilar, que possuem um formato espiral, estão “desenroladas”, passando a apresentar o formato de um tubo. O estribo é um pe-

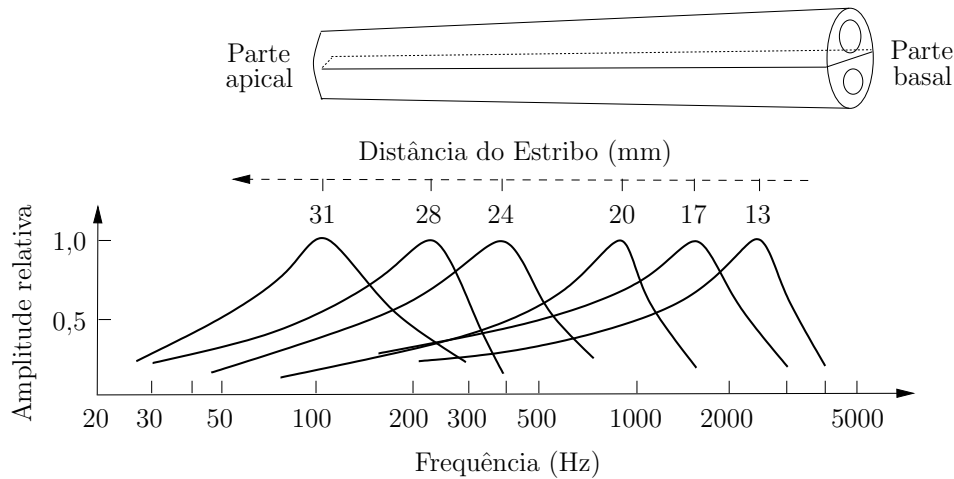


Figura 2.2: Envoltórias de máxima deflexão em função da distância ao longo da membrana basilar.

queno osso localizado no ouvido médio, próximo à parte basal da cóclea.

Por fim, o nervo auditivo carrega impulsos nervosos aferentes da cóclea ao cérebro e eferentes do cérebro à cóclea. Esses impulsos foram originados a partir das excitações (ocasionadas por ondas sonoras) nas células ciliadas presentes na membrana basilar.

2.3 Principais Conceitos de Psicoacústica

A Psicoacústica é composta por uma série de conceitos associados à percepção sonora. Estes dizem respeito às influências que grandezas como intensidade e frequência exercem na forma que escutamos o som, como, por exemplo, a audibilidade, a JND, o *pitch* e o mascaramento frequencial. Também podem estar relacionados à nossa percepção temporal e espacial, como o mascaramento temporal, o mascaramento binaural e o efeito de precedência. Por fim, podem se referir às envoltórias temporal e frequencial para explicar algum outro tipo de interpretação do som, como o timbre. Todos esses conceitos serão descritos a seguir.

2.3.1 Audibilidade

Quando escutamos um sinal de áudio, percebemos que sua intensidade pode variar, dependendo da potência com que este chega ao nosso ouvido. A essa intensidade percebida atribuímos o nome de audibilidade.

A intensidade é uma medida relativa, ou seja, seu valor é sempre dado em relação a uma referência. Esta é, normalmente, a potência do som de um milionésimo de milionésimo de watt por metro quadrado [1]. Assim sendo, dado um tom de amplitude A_1 e supondo que o tom de referência possua uma amplitude A_2 , a intensidade do primeiro é dada por:

$$I_1 = 20 \log_{10}(A_1/A_2).$$

Esta é chamada de nível de pressão sonora (do inglês, *Sound Pressure Level* ou SPL) e é expressa em decibéis (dB). Para não haver dúvidas, a unidade “dB” vem acompanhada da sigla “SPL”, como “10 dB SPL”, por exemplo.

É importante ressaltar que a unidade em decibéis utilizada é conveniente, já que o alcance entre as intensidades mínima e máxima que conseguimos escutar é muito alta (da ordem de 10^{12}). Além disso, nossa percepção de intensidade não é linear com o aumento desta.

A intensidade percebida—audibilidade—é, portanto, diferente do SPL de um sinal sonoro, já que, para diferentes frequências com o mesmo SPL, percebemos diferentes intensidades. Assim, podemos plotar curvas que variem com a frequência e que representem igual audibilidade, ou seja, tais que tons puros de diferentes alturas cujas intensidades estejam sobre a curva soem igualmente audíveis. Essas curvas são denominadas curvas de Fletcher-Munson [1] e são dadas na Fig. 2.3.

A cada curva atribui-se um nível de audibilidade especificado em *phons*¹ (1 *phon* equivale a 1 dB SPL em 1 kHz) [1]. A curva tracejada representa uma curva de audibilidade 0 *phon*. As curvas acima desta representam audibilidades de 10, 20,

¹A palavra *phon* vem do grego, que significa som ou voz.

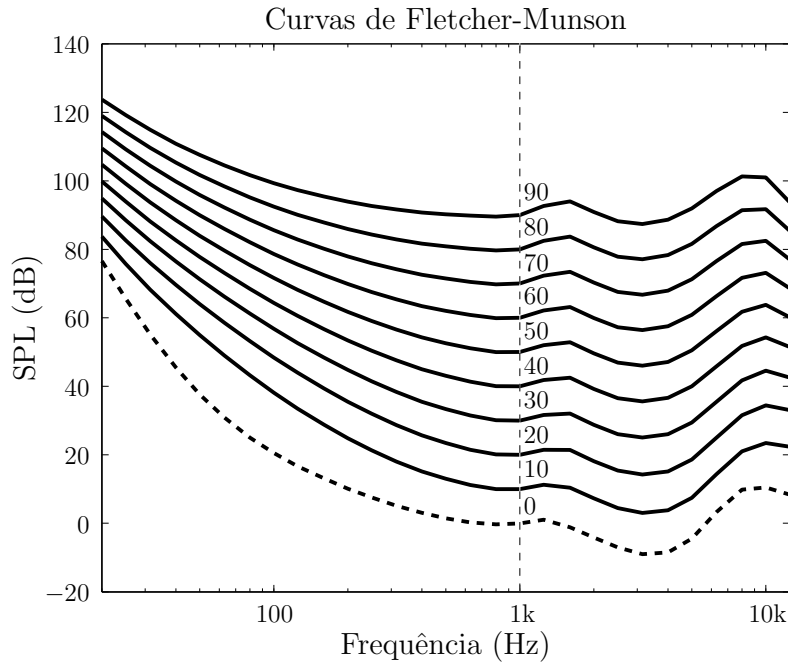


Figura 2.3: Curvas de audibilidade de Fletcher-Munson.

30 e assim sucessivamente até 90 *phons*. Estas são construídas levando-se em consideração um tom de referência (1 kHz) com uma intensidade fixa (no caso da Fig. 2.3, 0, 10, 20, ..., 90 dB SPL). Para os tons com as demais frequências, a sua intensidade é ajustada de forma que estes possuam a mesma audibilidade do tom de referência [2].

Analisando, agora, o formato das curvas obtidas, podemos perceber que o nosso ouvido possui uma maior sensibilidade em algumas regiões de frequência. Para as frequências em torno de 3–4 kHz, o nosso sistema auditivo possui sua sensibilidade máxima. Esta diminui bruscamente para frequências abaixo de 100 Hz.

Além disso, quanto maior o nível de audibilidade, menor é a nossa sensibilidade, já que a variação das curvas nessa região é bem menor. Isso significa que não somos capazes de identificar variações em intensidades muito altas com facilidade; ou ainda que, neste mesmo caso, baixas frequências irão contribuir de forma mais significativa.

A curva tracejada representa o limiar de audibilidade. Este é o nível de pressão sonora (SPL) mínimo de um tom que um ser humano é capaz de perceber. A curva que representa o limiar da dor, que alcança níveis de intensidade capazes de machucar nossos ouvidos, seria equivalente a uma curva de audibilidade de aproximadamente 120 *phons* [1].

Até agora, no entanto, só foi feita uma análise para tons senoidais. Os sons que escutamos podem ser compostos, por exemplo, pela frequência do tom original, denominada fundamental, mais harmônicos (ou parciais) desta; ou, ainda, incluir ruído com larga faixa de frequência. Podemos ajustar o nível de uma senoide pura para que esta possua a mesma audibilidade do som em questão [1]. Em seguida, para obter a audibilidade do som, vemos a quantos *phons* essa intensidade calculada representa para a frequência da senoide escolhida.

2.3.2 *Pitch*

Um sinal periódico possui, na maioria das vezes, um *pitch* que está associado ao inverso do seu período. Se ele for um tom puro, o seu *pitch* será equivalente à sua frequência. Quando escutamos, no entanto, um tom complexo (composto por uma fundamental e suas parciais), o nosso sistema auditivo não distingue separadamente as frequências presentes. Pelo contrário, percebe o som como sendo uma entidade única caracterizada por uma altura bem definida, o seu *pitch* [2].

Podemos dizer, assim, que o *pitch* é uma medida de percepção de altura. Ele pode ser definido de forma análoga à determinação da audibilidade. Ao ajustar a frequência de uma senoide pura para que esta possua a mesma frequência percebida do som em questão, a frequência obtida será o *pitch* [3].

A percepção de *pitch* está diretamente relacionada à música. Um instrumento musical, na maioria das vezes, emite sons compostos por parciais harmônicas (ou quase harmônicas) que, dessa forma, apresentam formas de onda (quase) periódicas. O *pitch* percebido está associado a essa periodicidade [1].

A Psicoacústica estuda, então, as diversas formas de percepção de *pitch* e propõe modelos, que se baseiam no estudo do nosso sistema auditivo, para a percepção automática. Esse estudo, além de em música, também pode ser aplicado a sinais de fala.

A discussão em torno do conceito de *pitch* é bastante ampla. Portanto, foi reservada uma seção para tratar dos principais tópicos relacionados a esse tema. Aqui foram citadas apenas algumas ideias iniciais de o que é a percepção de *pitch* e do que ela representa para a Psicoacústica.

2.3.3 JND

A menor diferença perceptível (do inglês, *Just Noticeable Difference* ou JND) de um parâmetro é definida como sendo a menor variação em seu valor que consegue ser detectada de forma confiável por uma pessoa. No caso de áudio [3], a JND se refere, tipicamente, à percepção da variação de intensidade ou de frequência.

A JND de frequência é, normalmente, maior para valores menores de intensidade (como, por exemplo, tons entre 5 e 10 dB SPL) e menor para valores maiores (como tons acima de 20 dB SPL). Ela varia de acordo com a frequência do sinal que estamos escutando. No caso de tons senoidais, a JND mínima é menor de 1% e ocorre para tons com frequências entre 1 e 2 kHz [3].

A JND de intensidade está diretamente relacionada com as curvas de audibilidade. Alguns resultados obtidos em [4] mostram que tons de igual audibilidade produzem, aproximadamente, JND's de intensidade iguais.

2.3.4 Mascaramento Temporal e Freqüencial

O mascaramento está relacionado à nossa sensibilidade para escutarmos um som na presença de um outro. Quando escutamos, por exemplo, um som isolado, este possui um determinado limiar de audibilidade. Ao escutarmos o mesmo som em conjunto com um mais intenso, tal limiar deverá aumentar. O mascaramento é,

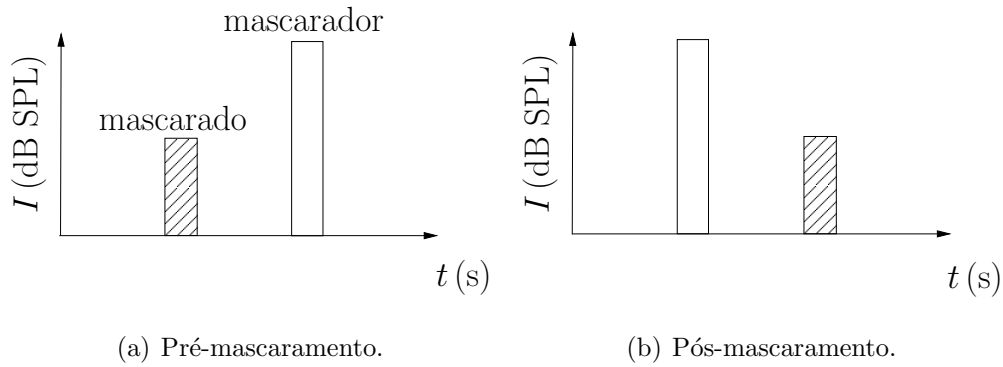


Figura 2.4: Mascaramento temporal.

então, definido como o aumento no limiar de audibilidade de um som na presença de outro [5].

Ao som que sofreu o aumento no limiar—ou diminuição de audibilidade—, denominamos mascarado, e ao som que provocou esse aumento, mascarador. O mascarador, como já foi dito, possui uma intensidade maior e é capaz de inibir—mascarar—a percepção de outros de menor intensidade. Normalmente, o mascaramento ocorre simultaneamente no tempo—mascaramento temporal—e na frequência—mascaramento frequencial. Os tipos de mascaramento (temporal e frequencial) serão inicialmente explicados em separado.

No caso do mascaramento temporal, os sinais sob consideração não ocorrem simultaneamente. O som mascarado pode ser emitido antes ou após o mascarador, contanto que essa diferença temporal seja suficientemente pequena. Esses dois casos de mascaramento temporal são denominados pré- e pós-mascaramento, respectivamente. A Fig. 2.4 ilustra esse conceito.

O mascaramento temporal depende do intervalo de tempo entre os sinais mascarador e mascarado. Quanto maior for esse intervalo, menor será a intensidade máxima do sinal de teste—mascarado—capaz de torná-lo inaudível. Além disso, esse intervalo é diferente para o pré- e pós-mascaramento. No primeiro caso, o mascaramento decai exponencialmente até se tornar praticamente inexistente em aproximadamente 20 ms. No segundo caso, o decaimento também é exponencial, porém, com uma duração que pode chegar a 200 ms [5], como ilustra a Fig. 2.5.

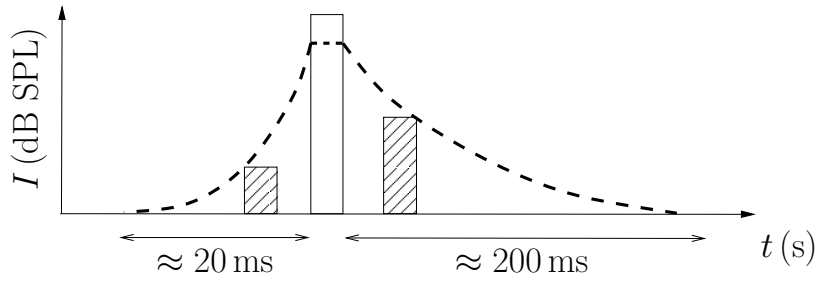


Figura 2.5: Decaimento no mascaramento temporal.

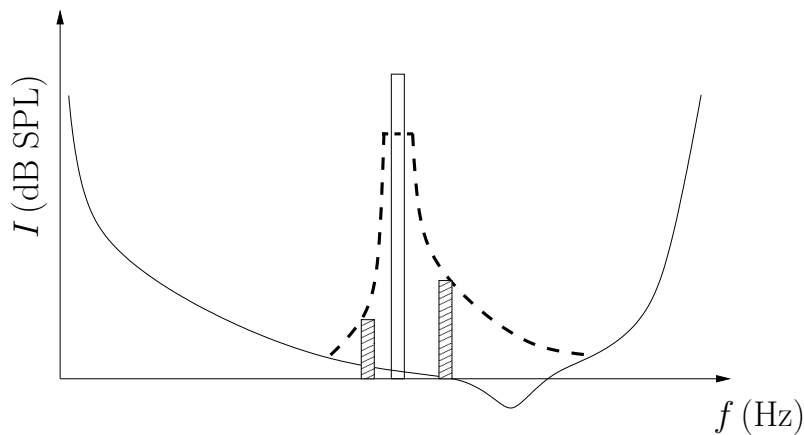


Figura 2.6: Decaimento no mascaramento frequencial.

No mascaramento frequencial, os sinais sob consideração são simultâneos e ocorrem em frequências próximas. No geral, sinais de baixa frequência mascaram de forma efetiva sinais de alta frequência, mas não o contrário [5]. Assim sendo, podemos construir curvas de decaimento para o mascaramento frequencial semelhantes às do mascaramento temporal. Nesse caso, teríamos gráficos de intensidade I (dB SPL) por frequência f (Hz), como na curva tracejada da Fig. 2.6.

No gráfico da Fig. 2.6, há um sinal de maior intensidade que mascara outros dois sinais em frequências próximas. Há também a curva—linha cheia—que representa o limiar de audibilidade em silêncio, ou seja, a intensidade mínima necessária para escutarmos um tom sem a interferência de outros sinais de áudio. A presença de um tom mascarador de frequência f eleva o limiar de audibilidade de tons com

frequências próximas a ela. Assim sendo, podemos dizer que, na presença de um tom externo, as curvas de decaimento—linha tracejada—juntamente com a curva do limiar de audibilidade em silêncio formam uma curva de limiar “global” de mascaramento [6].

O mascaramento frequencial pode ser explicado se levarmos em conta as deflexões que ocorrem ao longo da membrana basilar da cóclea na presença de estímulos senoidais [6]. As deflexões, ilustradas na Fig. 2.2, inibem a detecção de demais tons de excitação de frequências próximas e menor intensidade, o que caracteriza o mascaramento frequencial.

Até agora, os tipos de mascaramento foram explicados separadamente e somente para sinais que são pulsos no tempo, no caso do mascaramento temporal, ou tons senoidais, no caso do mascaramento frequencial. No caso de sinais complexos, o procedimento para a análise do mascaramento é diferente. Para esses sinais, é necessário, primeiro, identificar possíveis mascaradores no domínio da frequência; em seguida, estimar, para os mascaradores encontrados, curvas de mascaramento temporal e frequencial e; por fim, combinar as curvas encontradas de decaimento frequencial com o limiar de audibilidade em silêncio e com as curvas de decaimento temporal para criar um limiar global de audibilidade [6]. Pode-se dizer, então, que este limiar representa curvas de mascaramento frequencial variantes com o tempo.

2.3.5 Mascaramento Binaural

O mascaramento binaural é a redução no nível de mascaramento que ocorre quando escutamos os sinais mascarador e mascarado em circunstâncias diferentes nos dois ouvidos. Por exemplo, ao escutarmos dois tons simultâneos em fase, sendo um capaz de mascarar a presença do outro, só identificamos o tom mascarador. Ao alterarmos a fase do tom mascarado de 180° em somente um dos ouvidos, passamos a ser capazes de percebê-lo também. Nesse caso, a redução de mascaramento foi possível pois o nosso cérebro é capaz de identificar essa diferença de fase e, de alguma forma, somar construtivamente os tons mascarados [1]. O mesmo não ocorre para os tons mascaradores.

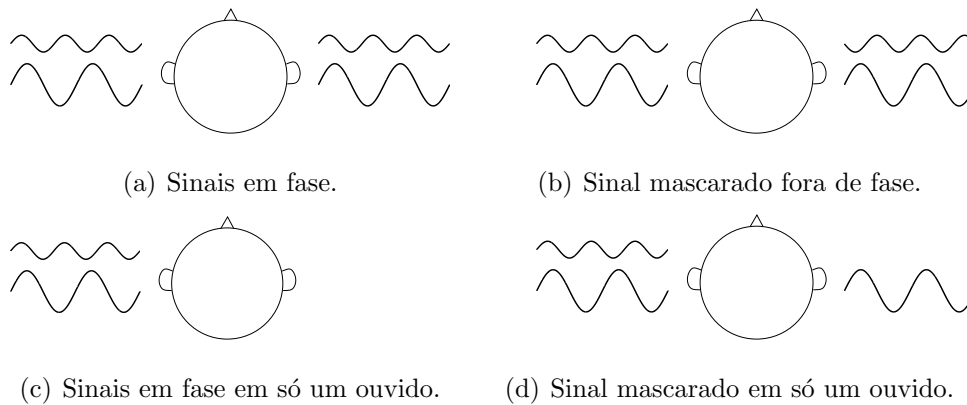


Figura 2.7: Mascaramento Binaural. Nos itens (b) e (d), o ouvinte é capaz de identificar a presença do sinal mascarado.

Um outro caso é quando escutamos, primeiramente, os dois sinais sob consideração em fase em apenas um dos ouvidos. Em seguida, o mascarador passa a ser tocado, também em fase, nos dois ouvidos. Com isso, passamos a perceber a presença do sinal mascarado, já que, com a adição do mascarador no outro ouvido, fomos capazes de “filtrá-lo” [1]. A Fig. 2.7 ilustra os quatro casos citados aqui.

2.3.6 Efeito de Precedência

Quando uma fonte sonora é emitida em uma sala, escutamos o som proveniente diretamente da fonte—som direto—e suas reflexões. Apesar disso, somos capazes, na maioria dos casos, de identificar corretamente a sua direção. Isso ocorre porque nós “escolhemos” essa direção como sendo a da primeira forma de onda que chega aos nossos ouvidos, ou seja, o som direto. Assim, a percepção direcional das reflexões que chegam poucos milissegundos depois são, nesse caso, suprimidas. A esse efeito, damos o nome de efeito de precedência [7].

Se escutarmos de uma posição centralizada um sinal $x(t)$ e uma versão defasada deste de τ segundos, $x(t - \tau)$, vindos de alto-falantes 1 e 2, respectivamente, perceberemos que o efeito de precedência ocorre para um dado intervalo de valores de τ . Para valores de até 1 ms, iremos escutar a fonte como vindo de uma posição intermediária entre os dois alto-falantes. Para valores entre 1 e aproximadamente 30 ms, percebemos o som como vindo somente do primeiro alto-falante, que é quando dizemos que ocorre o efeito de precedência. Para valores acima de 30 ms, escutamos

dois sons: o som “direto”, vindo do alto-falante 1, e um eco, vindo do alto-falante 2 [7].

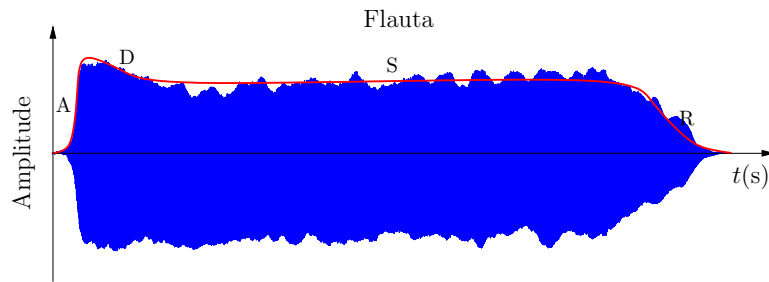
2.3.7 Timbre

O timbre de um som pode ser definido como a qualidade que o diferencia de um outro som de mesmo *pitch* e audibilidade. É o timbre que nos possibilita perceber a diferença entre sons emitidos por instrumentos diferentes. Há dois principais fatores que caracterizam o timbre de um sinal: a envoltória frequencial e a envoltória temporal [1].

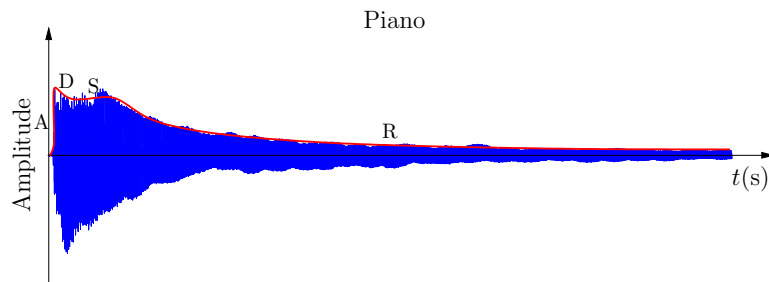
O primeiro caso se refere, principalmente, ao fator de decaimento espectral dos seus harmônicos, ou seja, a taxa com a qual a energia destes decai. Sinais que possuem um decaimento espectral baixo ou nulo, são sons “ásperos” e se parecem com zumbidos. Já os sinais que possuem um decaimento alto, são sons “brandos” e menos brilhantes.

Quando falamos de decaimento espectral, consideramos que a envoltória do espectro do sinal analisado é monotônica. Devemos, também, considerar o caso contrário. Alguns sinais, como o do violino e o da fala, possuem elevações na envoltória dos seus espectros. Estas elevações são chamadas de formantes e sua origem está relacionada às frequências de ressonância do corpo do instrumento ou do trato vocal, no caso da voz. Os formantes podem ser fixos, como é o caso do violino, ou podem variar com o tempo, como é o caso da fala. Eles também podem ser invariantes ou variantes com o *pitch*.

O segundo caso está relacionado ao ataque, ao decaimento, à sustentação e à relaxação do som. Essas quatro etapas, que compõem o tradicional modelo ADSR, representam a sua envoltória temporal. O ataque se refere ao início do som, é o tempo que este leva para ir do silêncio até sua intensidade máxima. O decaimento se refere aos instantes seguintes ao ataque, nos quais a intensidade inicial diminui. A sustentação se refere ao tempo de manutenção da nota. A relaxação se refere à fase final. Cada instrumento possui características próprias de ADSR e estas são importantes para a determinação do timbre de um sinal de áudio.



(a) Curva ADSR para a flauta.



(b) Curva ADSR para o piano.

Figura 2.8: Exemplos de curvas ADSR: flauta e piano. A curva em azul representa o sinal de áudio e a curva em vermelho, uma aproximação da curva ADSR.

Dois exemplos de curvas ADSR podem ser vistos nas Figs. 2.8(a) e 2.8(b) para a flauta e o piano, respectivamente. Ambos os instrumentos estão tocando uma nota dó por aproximadamente 2 s. Vemos que a diferença entre as envoltórias dos sinais aparece, principalmente, na duração do relaxamento. O relaxamento da flauta é mais brusco devido ao corte de fluxo de ar quando o instrumentista para de tocar. Já o do piano é lento e pode se estender por vários segundos se o seu pedal de sustentação² estiver acionado. Também é possível perceber que a nota gerada pela flauta se sustenta por mais tempo que a nota gerada pelo piano. Esse tempo, no entanto, somente pode ser controlado no caso da flauta.

2.4 Percepção de *Pitch*

Como já foi dito, o *pitch* pode ser definido como a frequência percebida de um som, que está diretamente relacionada à fundamental. Existem alguns modelos psicoacústicos que procuram justificar essa nossa percepção de *pitch*. Estes serão

²Um piano possui, tipicamente, dois ou três pedais. O pedal de sustentação prolonga a duração da nota, já que permite que as cordas vibrem livremente.

explicados aqui.

Nesta seção, também analisaremos o *pitch* para diferentes tons complexos e o conceito de Tons de Shepard. Por fim, abordaremos algumas outras definições de *pitch* existentes na literatura.

2.4.1 Definição de *Pitch*

De acordo com [3], o *pitch* de um sinal é a frequência selecionada ao ajustarmos a frequência de uma senoide de forma que esta corresponda corretamente à frequência percebida do sinal. Assim, a unidade de medida do *pitch* é o Hertz (Hz). Por convenção, a intensidade da senoide é de 40 dB SPL.

Por essa definição, não é possível ter um valor determinístico para o *pitch* de um sinal, já que diferentes pessoas percebem diferentes frequências. Assim, haverá uma média e um desvio padrão da distribuição do *pitch* estimado. Essa distribuição pode ser, no entanto, bimodal. Dessa forma, dizemos que o sinal terá um determinado *pitch* somente se a grande maioria dos ouvintes selecionar frequências próximas a este.

2.4.2 Sensação de *Pitch*

O conceito de brilho é diferente do conceito de *pitch*. O brilho de um som está associado a quantidade de componentes de alta frequência que este possui em seu espectro—se tiver muitas, será um som brilhante—, enquanto que o *pitch* está associado à sua periodicidade [1].

Um sinal de poucos ciclos irá soar como um clique, que pode ser ou não brilhante dependendo da sua frequência. Este sinal, no entanto, não nos dá uma sensação de *pitch*, que só pode ser obtida com mais ciclos do mesmo sinal. Os valores que determinam essa transição variam entre quatro e dez ciclos. Com quatro ciclos, percebemos somente o clique; com dez, a sensação já é inteiramente de *pitch*; com valores entre quatro e dez, a sensação é uma mistura de clique e *pitch*. Esses valores aumentam ligeiramente com o aumento da frequência do sinal.

Para sinais de frequências extremamente baixas ou altas—abaixo de 20 Hz ou acima de 2 kHz—, no entanto, não é possível termos uma sensação de *pitch*. Dessa forma, dizemos que estes sinais não possuem um *pitch* definido.

2.4.3 Teoria do Lugar

A Teoria do Lugar do *pitch* está relacionada ao funcionamento da cóclea, explicado no início deste capítulo. Nesta, a membrana basilar vibra em diferentes regiões de acordo com a frequência do estímulo de entrada. Assim, dizemos que é organizada de forma tonotópica [3].

Ao vibrar, a membrana excita diferentes neurônios presentes no nervo auditivo, formando uma mensagem codificada—composta dos neurônios que estão ativos e inativos—que será enviada ao cérebro. Esta mensagem mantém sua estrutura tonotópica até sua chegada ao cérebro, onde é decodificada. O cérebro é, assim, capaz de identificar o *pitch* do sinal.

Alguns efeitos são capazes de comprovar a veracidade da Teoria do Lugar. Um deles, conhecido como regra de Stevens, diz que, quando a intensidade aumenta, o *pitch* de um tom grave diminui e o de um tom agudo aumenta. Esse efeito é ilustrado na Fig. 2.9. Da mesma forma, estudos mostram que, para estímulos mais intensos, o pico do padrão de excitação formado pela membrana basilar se altera [3].

2.4.4 Teoria do Tempo

A Teoria do Tempo do *pitch* está relacionada à taxa de disparos de neurônios no nervo auditivo ao escutarmos um determinado som. Essa taxa codifica, assim como na Teoria do Lugar, informação de frequência presente no estímulo, permitindo, assim, a interpretação do *pitch* por parte do nosso cérebro.

Os impulsos nervosos disparam a uma taxa aleatória antes da presença do som. À medida que a intensidade deste aumenta, os impulsos se organizam no tempo, de forma que sua taxa esteja sincronizada com o período—inverso da frequência fundamental—do sinal. Se aumentarmos ainda mais sua intensidade, o número de

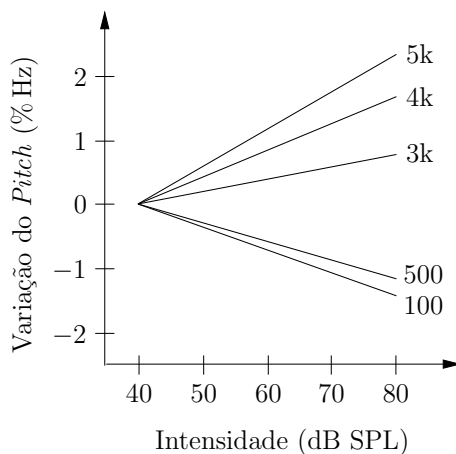


Figura 2.9: Regra de Stevens: variação do *pitch* (em % Hz) em função da intensidade (em dB SPL).

impulsos e a sincronia aumentam. Esta sincronia, no entanto, ocorre somente para sinais mais graves. Em tons que possuem uma frequência fundamental a partir de 2 kHz, a taxa de disparos deixa de seguir o estímulo.

Assim sendo, podemos dizer que há uma espécie de “contagem” realizada por nós que está diretamente relacionada ao *pitch* do som escutado. Se a sua frequência fundamental aumentar, a nossa “resolução temporal” já não será suficiente para distinguir o seu período e a sincronização da taxa de disparos se perde.

2.4.5 *Pitch* de Tons Complexos

Instrumentos como o violão, a flauta e até a voz possuem uma estrutura parecida, em geral, com a de tons complexos: frequência fundamental e harmônicos. É, portanto, mais interessante analisar a percepção do *pitch* de tons complexos do que de senoides puras.

O *pitch* percebido de um tom complexo é, normalmente, igual à sua frequência fundamental. Há, porém, um fenômeno conhecido em Psicoacústica que é o fato de atribuímos ao *pitch* de alguns sinais a sua frequência fundamental quando esta não está presente. A Teoria do Lugar falha em explicar esse efeito. Por outro lado,

segundo a Teoria do Tempo, percebemos o *pitch* como sendo o período do sinal. Se este não possui uma fundamental— f_0 —, mas somente harmônicos— $2f_0, 3f_0, 4f_0$, etc.—, a frequência de sua envoltória ainda será igual a f_0 . Esse efeito só ocorre, porém, para baixas frequências.

2.4.6 Tons de Shepard

Os tons de Shepard são uma ilusão da percepção de *pitch*. Eles são compostos por um determinado número de parciais separadas entre si por uma oitava. Há uma envoltória fixa em frequência que vai para zero em frequências baixas e altas, como se pode ver na Fig. 2.10 [1].

A ideia é sempre aumentarmos (ou diminuirmos) a frequência das parciais. Dessa forma, parciais de baixa frequência que eram antes inaudíveis (se encontravam abaixo do limiar de audição) se tornam audíveis e parciais de alta frequência que eram audíveis deixam de ser escutadas (já que estas passam a ficar abaixo do limiar). Ao aumentarmos, por exemplo, a frequência em um semitom, perceberemos o *pitch* aumentar. Se assim for, sucessivamente, até a frequência aumentar em uma oitava (ou seja, dobrar a frequência de todos os harmônicos), teremos o mesmo espectro dos tons de Shepard iniciais. O *pitch* percebido cresce indefinidamente, apesar de as parciais estarem presentes sempre na mesma oitava.

Se aumentarmos ou diminuirmos, agora, a frequência em seis semitons³ (harmônicos pontilhados na Fig. 2.10), a percepção será diferente para cada ouvinte. Alguns percebem o *pitch* aumentar e outros, diminuir.

2.4.7 Definições de *Pitch*

As duas principais definições de *pitch*—*pitch* virtual e *pitch* espectral—foram elaboradas por Ernst Terhardt em seu trabalho de percepção de *pitch* [1]. Na primeira definição, o *pitch* é caracterizado pela presença de harmônicos e, no segundo, por componentes senoidais puras.

³As curvas obtida ao diminuir ou aumentar a frequência em seis semitons é exatamente a mesma.

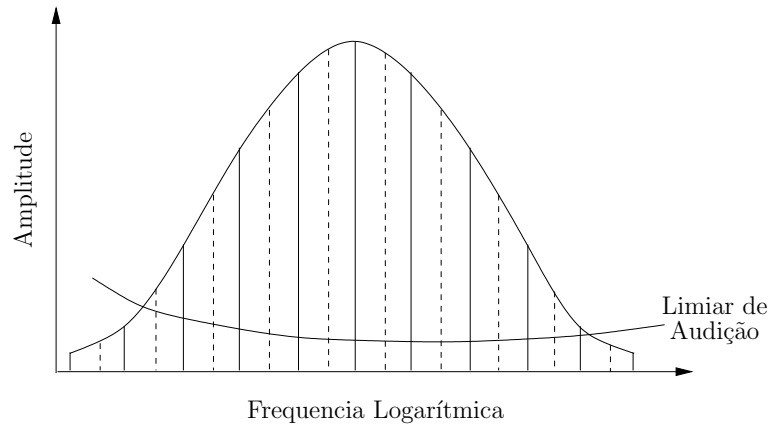


Figura 2.10: Tons de Shepard.

O *pitch* virtual é o que escutamos quando não há a presença da fundamental. Para isso, o nosso sistema auditivo utiliza a informação de frequência presente nos demais harmônicos do sinal. Um exemplo é quando escutamos uma voz grave por rádio ou telefone. Nesse caso, somos capazes de atribuir corretamente o *pitch* do sinal à sua fundamental, apesar de esta ter sido cortada na transmissão.

O *pitch* espectral, por outro lado, é o que atribuímos à própria frequência da senoide que estamos escutando, no caso de senoídes puras, ou à fundamental, quando presente, no caso de sons complexos. Também podemos associá-lo ao *pitch* percebido de um sinal que possui alguma singularidade espectral. Exemplos são o *pitch* de um ruído branco após passar por um filtro passa-faixas de largura de faixa estreita ou o mesmo ruído após passar por um filtro pente. Este último é chamado de *pitch* de repetição.

2.5 Aplicações de Psicoacústica

A aplicação mais direta de Psicoacústica é em compressão de áudio. A codificação, nesse caso, é com perdas e o objetivo é eliminar do sinal de áudio aquilo que não escutamos, como sons mascarados e sons abaixo do limiar de audição.

Utilizamos as informações presentes no limiar de audição e no mascaramento frequencial para a quantização em um processo de compressão de um sinal de áudio. Sabemos que, para cada componente frequencial, o limiar de audição é diferente. Sabemos, também, que na presença de um mascarador em frequência, o limiar de audição aumenta nas faixas próximas a este. Isso nos permite alocar os bits de acordo com a nossa sensibilidade em cada faixa, ou seja, diminuimos o número de bits alocados em faixas menos sensíveis (como é o caso de frequências abaixo de 100 Hz ou acima de 10 kHz ou próximas a um mascarador) e aumentamos o número de bits alocados em faixas mais sensíveis (como a faixa entre 3 e 4 kHz). Esse processo nos permite diminuir o número total de bits utilizados enquanto mantemos o erro de quantização inaudível.

Os principais codificadores de áudio com perdas utilizam, como uma de suas etapas de codificação, modelos psicoacústicos. Um exemplo é o MPEG (do inglês, *Movie Picture Expert Group*) Audio Layer III, também conhecido como MP3, que é amplamente utilizado hoje em dia para codificar sinais de música para distribuição na internet. Ele é capaz de reduzir sinais de áudio em até dez vezes com transparência, ou seja, sem diferenças perceptivas em relação ao sinal de áudio original [6].

Outras aplicações de Psicoacústica são relacionadas à música. Entender como funciona a nossa percepção de *pitch* é uma ferramenta para a análise e síntese musical. Alguns modelos que procuram explicar os fenômenos associados a esse tema foram elaborados a partir da Teoria do Lugar e Teoria do Tempo. Estes, no entanto, não são completos e ainda possuem falhas na identificação do *pitch* de alguns sinais [3]. Há ainda outros conceitos, como o timbre, que também possuem aplicações em síntese de música. Nesse caso, o conhecimento do modelo ADSR e o decaimento espectral de diversos instrumentos nos permite sintetizar sons com características próprias de cada um deles.

2.6 Considerações Finais

Este capítulo teve como objetivo fornecer uma base teórica para um melhor entendimento das próximas etapas desse projeto. Abordamos os principais temas da área de Psicoacústica procurando sempre esclarecer as principais ideias relacionadas a cada um.

Os principais conceitos explicados aqui foram a audibilidade, relacionada à percepção de intensidade; o *pitch*, relacionado à percepção de altura; a JND, relacionada, normalmente, à percepção de variação de altura ou intensidade; o mascaramento, relacionado a percepção de sons na presença de outros de maior intensidade; e o timbre, relacionado às características próprias de cada instrumento. Cada um desses conceitos possui sua importância e aplicação em áudio, seja ela em codificação ou em análise e síntese música. É, portanto, importante entender as ideias presentes e, mais que isso, saber identificá-las em um sinal de áudio.

Para isso, foi implementado um conjunto de testes psicoacústicos com base nos conceitos descritos. Eles tem como finalidade a ilustração prática de cada fenômeno e, além disso, alguns permitem a realização de testes subjetivos para a comprovação dos resultados teóricos. O próximo capítulo explica com detalhes cada um dos programas implementados e os relaciona com a teoria aqui explicada.

Capítulo 3

Testes Psicoacústicos

3.1 Introdução

Neste capítulo, vamos explicar com mais detalhes os testes psicoacústicos, mencionados no Cap. 1, implementados. As interfaces feitas serão expostas aqui e, quando existentes, os resultados serão exibidos em gráficos. Além disso, todos os parâmetros de entrada serão explicados, deixando claro quais os seus possíveis valores e que impactos possuem no sinal de saída.

Um trabalho prévio a este [8] apresenta, com a mesma finalidade, programas de interface gráfica que também ilustram conceitos de Psicoacústica. Os testes propostos aqui foram, de forma geral, selecionados para complementar (em termo de conteúdo) os já implementados.

Foram implementados, no total, 11 testes. Estes são listados a seguir, juntamente com um breve resumo da sua finalidade.

1. Curvas de Fletcher-Munson - permite ao usuário estimar as curvas de audibilidade;
2. Atrasa e Soma - ilustra o conceito *pitch* de repetição através de um sistema “atrás e soma”;
3. *Pitch* sem a Fundamental - permite ao usuário comparar os *pitches* percebidos de um tom complexo com e sem a sua fundamental;

4. *Pitch* e Intensidade - permite ao usuário estimar a variação do *pitch* de um tom quando há um aumento na sua intensidade;
5. *Pitch* vs. Clique - permite ao usuário estimar o número de ciclos necessários para perceber corretamente o *pitch* de um tom;
6. Teoria Dúplex do *Pitch* - ilustra os conceitos de *pitch* virtual e *pitch* espectral;
7. JND de Frequência - permite ao usuário estimar a JND de frequência para tons com intensidade e frequência especificadas;
8. Mascaramento Temporal - permite ao usuário estimar a curva de decaimento do mascaramento temporal;
9. Mascaramento Freqüencial - permite ao usuário estimar a curva de decaimento do mascaramento freqüencial para uma frequência de mascarador específica;
10. Mascaramento Binaural - ilustra o mascaramento binaural com os quatro casos apresentados no capítulo anterior; e
11. Efeito de Precedência - permite ao usuário variar o atraso de um som que será emitido em um alto-falante (dado que este mesmo som será emitido sem atraso em um outro alto-falante) para ilustrar o efeito de precedência.

Os testes psicoacústicos listados podem ser divididos tanto por tipo quanto por funcionalidade. No primeiro caso, eles são divididos em testes que pertencem ao tema percepção de *pitch* (*Atrasa e Soma*; *Pitch* sem a Fundamental; *Pitch* e Intensidade; *Pitch* vs. Clique; e Teoria Dúplex do *Pitch*) e testes relacionados a outros conceitos de Psicoacústica. No segundo caso, eles são divididos em testes que servem somente para ilustrar algum conceito (*Atrasa e Soma*; *Pitch* sem a Fundamental; Teoria Dúplex do *Pitch*; Mascaramento Binaural; e Efeito de Precedência) ou testes que servem tanto para ilustração, como para a aplicação de testes subjetivos.

3.2 Experimentos Implementados

O objetivo desta seção é apresentar os programas implementados e, além disso, mostrar sua relação com a teoria explicada no Cap. 2. Os valores dos parâmetros



Figura 3.1: Interface gráfica do invólucro.

de entrada de cada programa serão expostos e alguns resultados serão apresentados. Uma análise mais aprofundada destes, no entanto, será dada no próximo capítulo, que trata dos testes subjetivos realizados.

3.2.1 Invólucro e Calibração

As interfaces gráficas do invólucro e da calibração podem ser vistas nas Figs. 3.1 e 3.2. O invólucro é responsável por unir os programas implementados em uma só interface. A partir dele, o usuário é capaz de acessar cada um deles.

Antes de iniciar algum experimento, no entanto, é necessário realizar a calibração. Nesta, há dois tipos de sinais: um tom de frequência arbitrada 1 kHz e um pulso, que representam os dois tipos de sinais utilizados no experimento. Para cada um destes, o usuário deve selecionar a menor amplitude na qual ainda seja possível escutar o sinal. A amplitude pode variar de 0 até 10^{-5} . O valor máximo de 10^{-5} foi selecionado

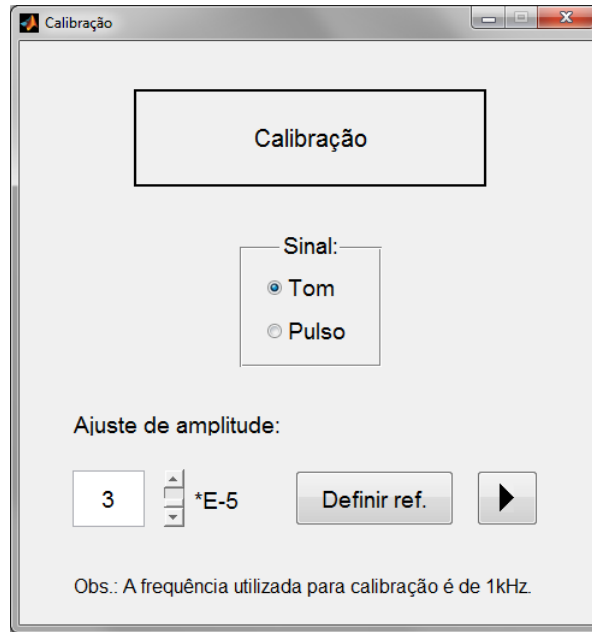


Figura 3.2: Interface gráfica da calibração.

já que, nessa amplitude, os dois sinais testados apresentam uma percepção clara. Os valores ajustados serão usados como referência para o cálculo de intensidades em dB SPL nos demais experimentos.

3.2.2 Curvas de Fletcher-Munson

O experimento ilustra o conceito de curvas de Fletcher-Munson, visto no Cap. 2. A interface gráfica principal pode ser vista na Fig. 3.3. A Fig. 3.4 mostra um diálogo pertencente ao programa.

No teste, o usuário escuta um tom de referência com frequência de 880 Hz e intensidade constante ao longo da experiência. Esta é ajustada pelo próprio usuário logo no início, no diálogo da Fig. 3.4. Existe, também, um conjunto de sinais de análise composto por tons de frequências 55, 110, 220, 440, 1760 e 3250 Hz.

O usuário irá, então, para cada sinal de análise, regular a sua intensidade (que pode ir de 0 até 90 dB SPL) de forma que esta e a intensidade do sinal de referência sejam igualmente percebidas, ou seja, que os dois sinais tenham a mesma audibi-



Figura 3.3: Interface gráfica principal do experimento “Curvas de Fletcher-Munson”.

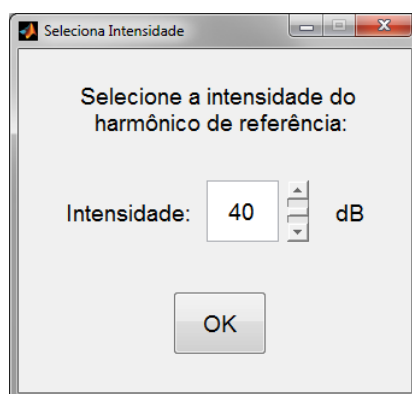


Figura 3.4: Diálogo pertencente ao experimento “Curvas de Fletcher-Munson”.

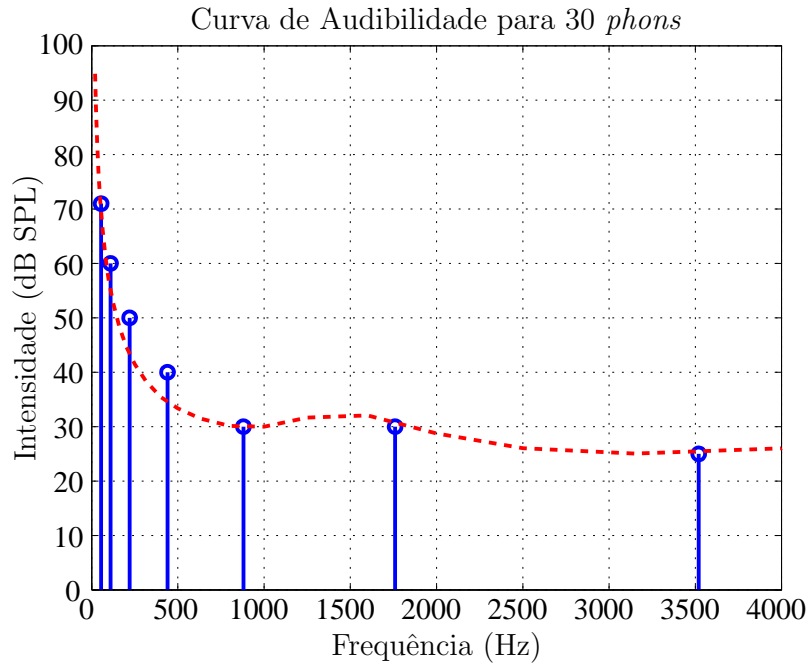


Figura 3.5: Exemplo de gráfico gerado pelo experimento “Curvas de Fletcher-Munson”.

lidade. Isso resultará em um vetor de sete valores correspondentes às intensidades ajustadas para cada tom. Por fim, esse resultado é salvo em um arquivo `.mat`, um gráfico de Intensidade (dB SPL) por Frequência (Hz) é gerado e o programa se fecha. Um exemplo de gráfico é o da Fig. 3.5, onde os pontos azuis representam o vetor de resultados e a curva em vermelho, os valores originais da curva de Fletcher-Munson retirados da norma ISO 226:2003. A intensidade de referência utilizada foi de 30 dB SPL.

A análise do resultado será feita no próximo capítulo.

3.2.3 Atrasa e Soma

O experimento procura ilustrar o conceito de *pitch* de repetição. Ele é gerado através de um sistema “atrasa e soma” visto na Fig. 3.6. Nesta, as variáveis T e g são, respectivamente, um atraso temporal e um ganho. A interface gráfica do experimento pode ser visto na Fig. 3.7. O sinal de entrada $x(t)$ é um ruído branco.

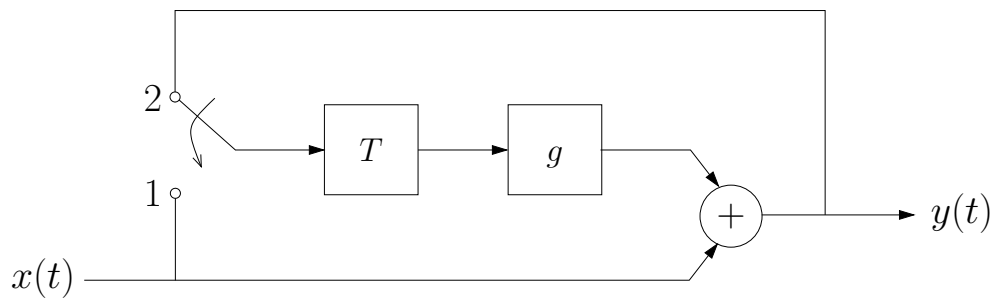


Figura 3.6: Diagrama de blocos de um sistema “atrasa e soma”. É usado para simular o *pitch* de repetição.

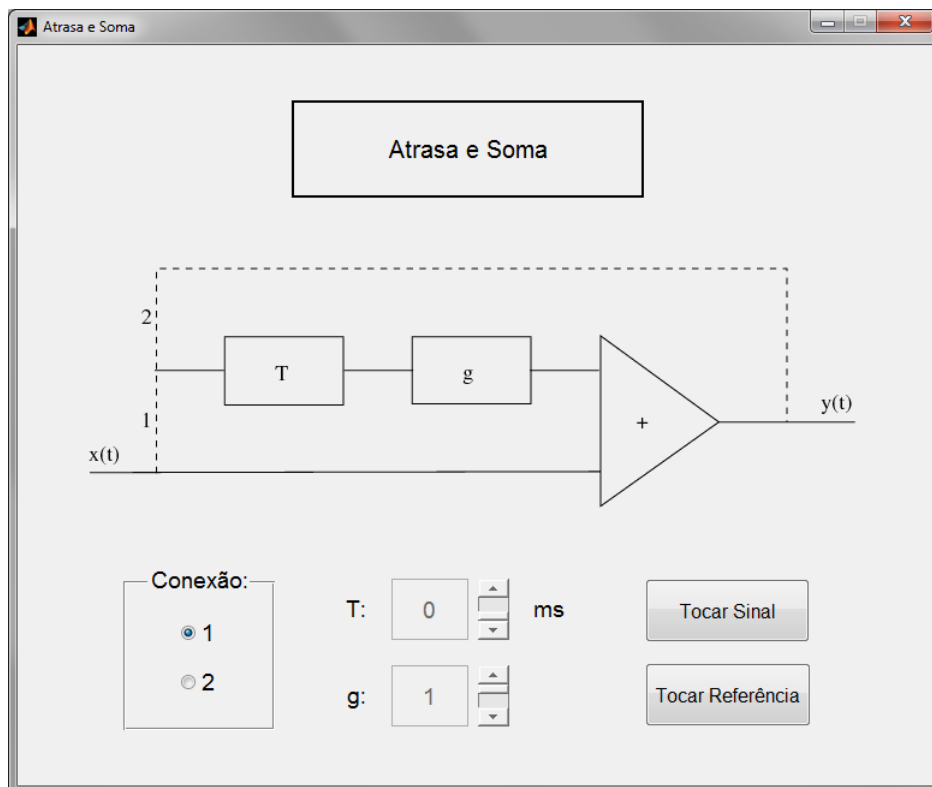


Figura 3.7: Interface gráfica do experimento “Atrasa e Soma”.

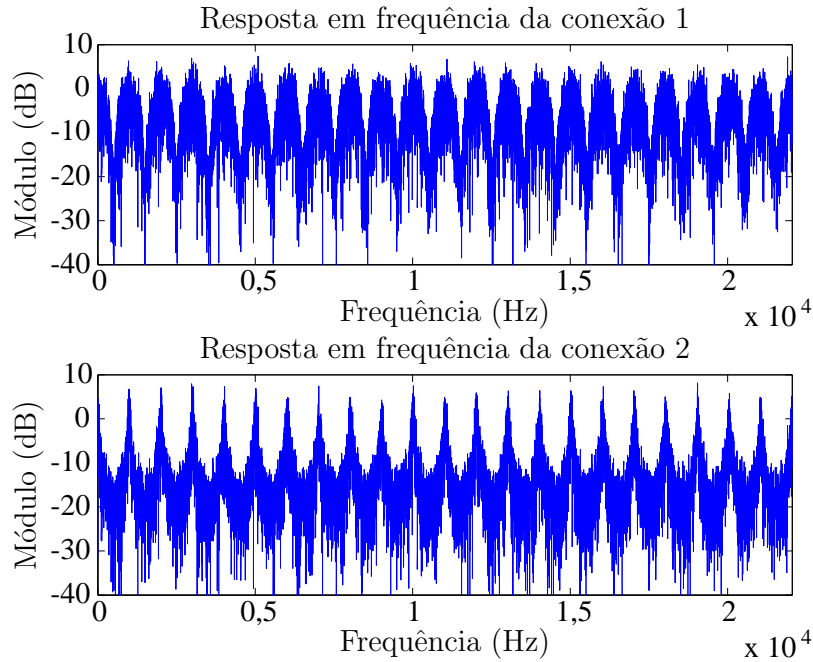


Figura 3.8: Resposta em frequência da saída $y(t)$ para as conexões 1 e 2 (com $T = 1$ ms e $g = 0,8$) do experimento “Atrasa e Soma”.

No experimento, o usuário escuta, primeiramente, a saída do sistema com a conexão 1. Pode-se perceber que o *pitch* se ajusta ao inverso do atraso escolhido (que pode variar de 0 até 20 ms) e é mais evidente quanto maior for o ganho (que pode variar de 0 até 1). A seguir, a conexão 1 é desfeita e o usuário passa a escutar a saída do sistema com a conexão 2. O *pitch* percebido é o mesmo, só que mais evidente. Esse efeito pode ser justificado pelas respostas em frequência em cada conexão, vistas na Fig. 3.8.

O conceito de *pitch* de repetição pode ser associado ao *pitch* que percebemos quando estamos em um ambiente fechado com uma fonte sonora. O som que chega aos nossos ouvidos é o emitido diretamente da fonte e suas versões atenuadas e atrasadas, que são as reflexões. A iteração “sinal direto-reflexões” nos faz ter a percepção de um *pitch*. Essa percepção, no entanto, é somente válida para pequenos valores de atraso, já que, caso contrário, as reflexões se transformam em ecos, que são percebidos de forma separada do som direto [9].

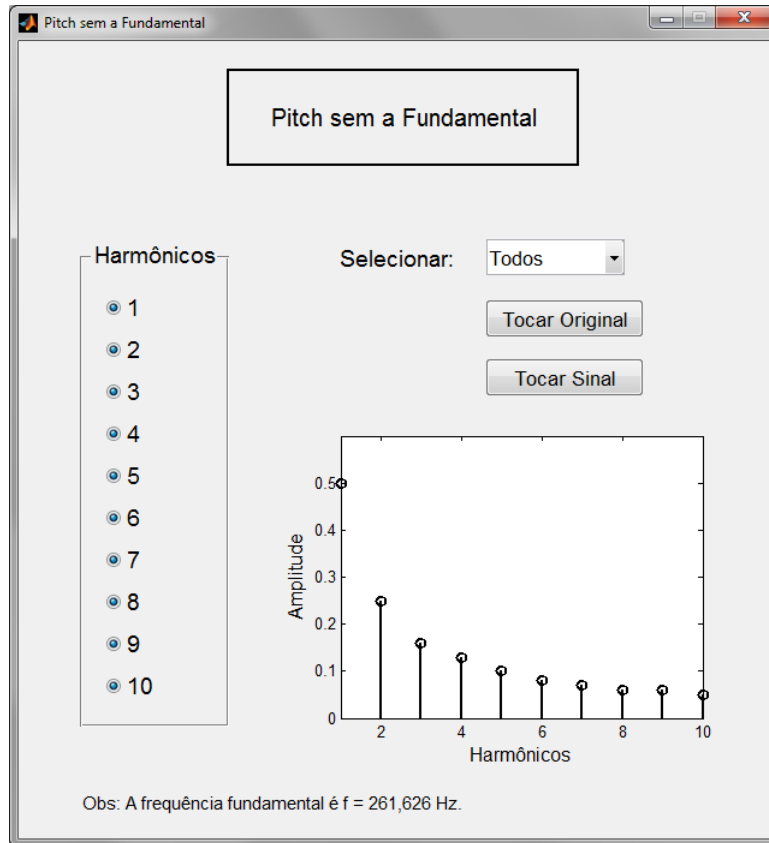


Figura 3.9: Interface gráfica do experimento “*Pitch sem a Fundamental*”.

3.2.4 *Pitch sem a Fundamental*

O experimento procura ilustrar a relação entre o *pitch* de um som e sua frequência fundamental. A interface gráfica pode ser vista na Fig. 3.9.

O experimento possui um sinal de referência, que é um tom complexo com uma fundamental de aproximadamente 261 Hz (equivalente a uma nota dó) e mais nove harmônicos. A referência pode ser visualizada no gráfico da Fig. 3.9. A partir dela, o usuário cria um sinal de análise: ele escolhe se a fundamental fará ou não parte do sinal resultante e, ainda, quais harmônicos estarão presentes. Por fim, é possível escutar o sinal original e o sinal gerado para comparação.

Este experimento nos permite analisar o *pitch* de um tom complexo com e sem a sua fundamental e compararmos os resultados. Além disso, também é possível analisar o *pitch* de alguns outros sinais relevantes, como sinais com somente harmônicos

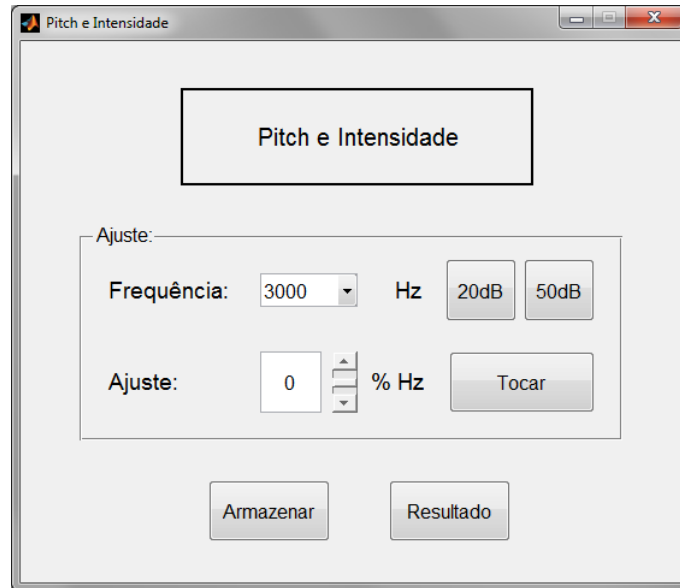


Figura 3.10: Interface gráfica do experimento “*Pitch* e Intensidade”.

ímpares, sinais com somente os três últimos harmônicos e sinais sem um dos harmônicos centrais, entre outros.

3.2.5 *Pitch* e Intensidade

Esse experimento procura ilustrar a regra de Stevens, explicada no Cap. 2. Ele também permite a realização de um teste para determinar a variação do *pitch* percebido em função da intensidade e da frequência. A interface gráfica do experimento pode ser vista na Fig. 3.10.

No teste, o usuário escuta, para uma determinada frequência, um tom de 20 dB SPL e um de 50 dB SPL. A seguir, ele ajusta a frequência do tom de 20 dB SPL para que o seu *pitch* percebido seja o mesmo que o percebido para 50 dB SPL. O ajuste pode variar de -4 até 4% da frequência central. Esse procedimento é realizado para as frequências 3000, 4000, 5000, 500, 200 e 100 Hz. Estes valores foram retirados de [3]. O vetor de valores selecionados pelo usuário é salvo em um arquivo `.mat`, um gráfico de resultados é gerado e o programa se fecha. O gráfico é de variação de *pitch* por intensidade e um exemplo é o da Fig. 3.11.

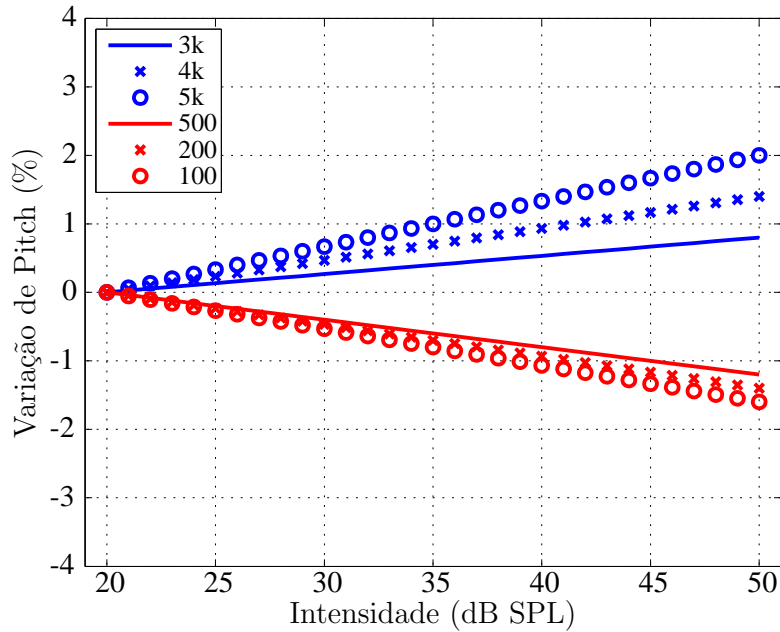


Figura 3.11: Exemplo sintético de gráfico gerado pelo experimento “*Pitch e Intensidade*”.

3.2.6 *Pitch vs. Clique*

Esse experimento procura ilustrar o efeito que o número de ciclos de um sinal tem na sensação de *pitch* por ele gerada. Ele também permite a realização de um teste para determinar o número de ciclos necessários para percebermos corretamente o *pitch*. A interface gráfica do experimento pode ser vista na Fig. 3.12.

No teste, o usuário deve ajustar o número de ciclos, que pode variar de 0 até 50, de um sinal senoidal de forma que, ao escutá-lo, seja capaz de perceber corretamente o seu *pitch*. Para isso, pode escutar um sinal de referência, que é um sinal senoidal com a frequência analisada e 1 s de duração. Esse procedimento é realizado para as frequências 55, 110, 220, 440, 880, 1760 e 3520 Hz, que representam desde tons graves (55 Hz) até agudos (3520 Hz). Por fim, o vetor de valores ajustados pelo usuário é salvo em um arquivo `.mat` e o programa se fecha.

O usuário é capaz de perceber, à medida que aumenta o número de ciclos, que a sensação de clique existente a princípio diminui aos poucos até desaparecer. Enquanto isso, a sensação de *pitch* já aparece com poucos ciclos, ainda quando podemos

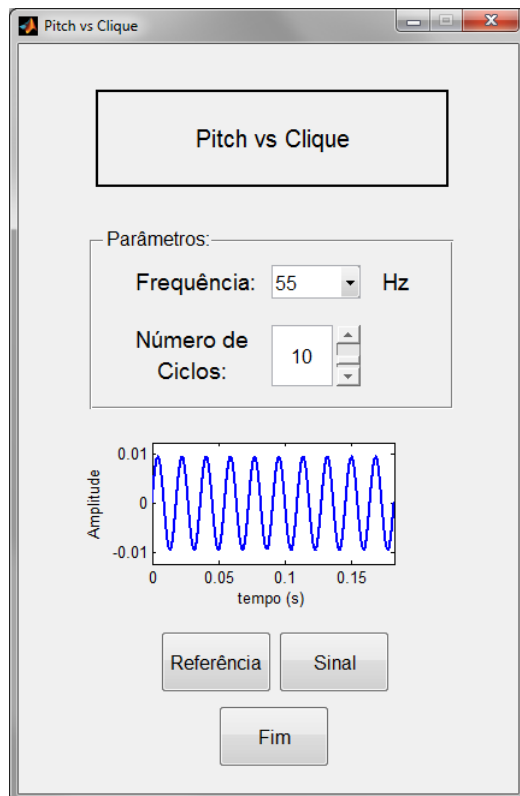


Figura 3.12: Interface gráfica do experimento “*Pitch vs. Clique*”.

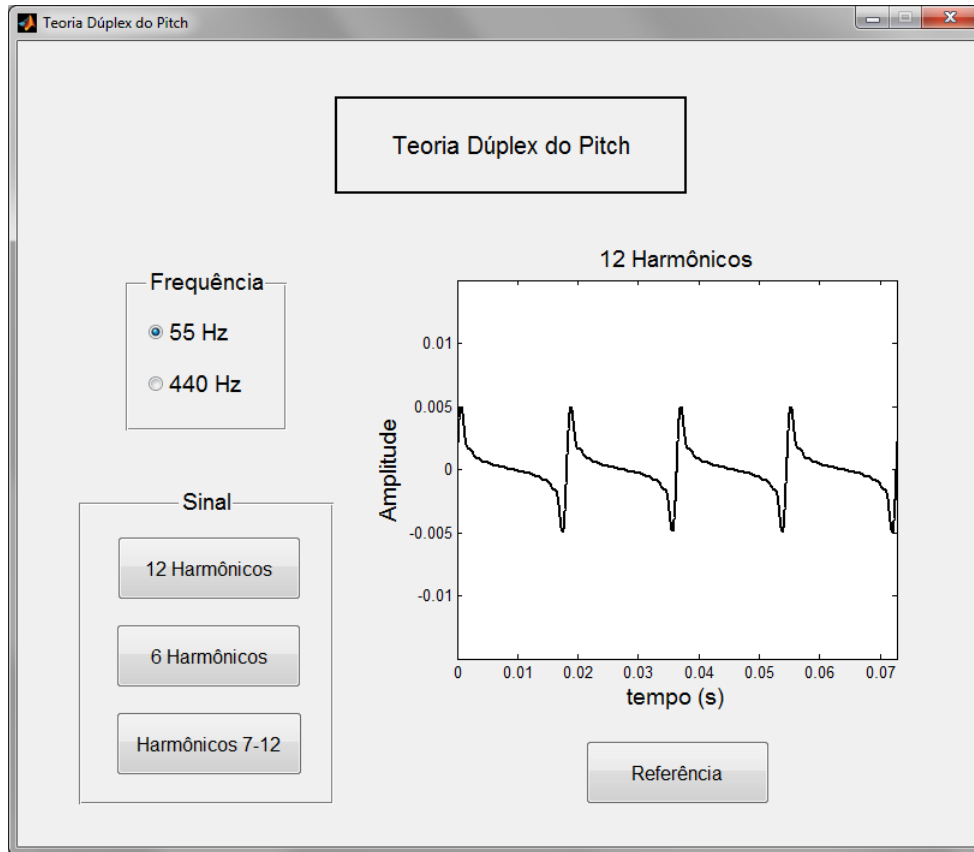


Figura 3.13: Interface gráfica do experimento “Teoria Dúplex do *Pitch*”.

perceber o clique, e se torna mais evidente com o aumento destes.

3.2.7 Teoria Dúplex do *Pitch*

Esse experimento procura ilustrar as diferenças na percepção de *pitch* para sinais de alta e baixa frequência. Ele é chamado de Teoria Dúplex porque relaciona, com exemplos, os dois modelos de percepção de *pitch* conhecidos—Teoria do Lugar e Teoria do Tempo. A interface gráfica do experimento pode ser vista na Fig. 3.13.

No experimento, o usuário escuta três tons complexos para duas frequências fundamentais diferentes: 55 e 440 Hz. Estes valores foram retirados de [1]. O primeiro sinal possui os 12 primeiros harmônicos, incluindo a fundamental; o segundo, os seis primeiros, também incluindo a fundamental; e o terceiro, os seis últimos.

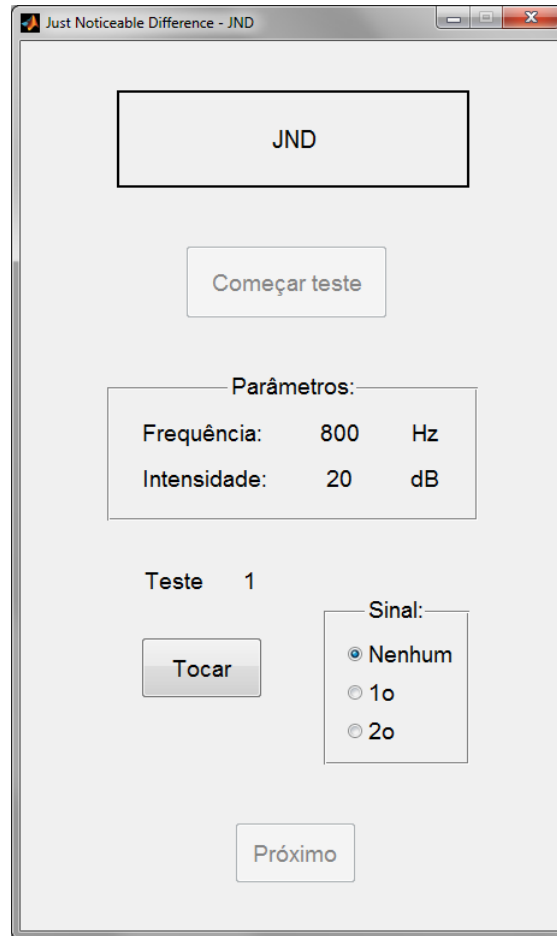


Figura 3.14: Interface gráfica principal do experimento “JND de Frequência”.

O *pitch* percebido para os três tons de fundamental 55 Hz é o mesmo, enquanto que para os tons de fundamental 440 Hz, o *pitch* do terceiro tom (que possui somente os seis últimos harmônicos) é ligeiramente maior que o dos dois primeiros. Isso ocorre já que, segundo a Teoria do Tempo, a nossa percepção de *pitch* para baixas frequências está relacionada à periodicidade do sinal, ou seja, a frequência da sua envoltória. Já para frequências maiores, a Teoria do Tempo deixa de ser válida.

3.2.8 JND de Frequência

O objetivo desse experimento é ilustrar o efeito de JND de frequência e, além disso, possibilitar a realização de um teste para estimá-la. As interfaces gráficas do experimento podem ser vistas nas Figs. 3.14 e 3.15.

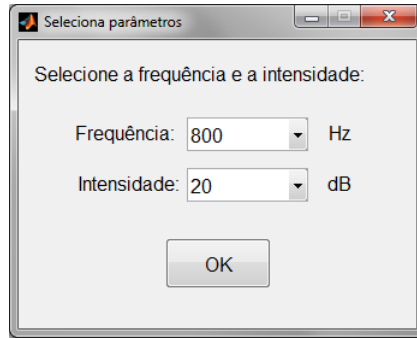


Figura 3.15: Diálogo pertencente ao experimento “JND de Frequência”.

No programa, o usuário seleciona em um diálogo, visto na Fig. 3.15, a frequência e intensidade com que deseja realizar o teste. Os valores possíveis para frequência são 200, 400, 600, 800, 1000, 2000, 4000 e 8000 Hz e para intensidade são 5, 10, 20, 40 e 80 dB SPL. Estes valores foram retirados de [3] e, provavelmente, foram selecionados para facilitar a realização do teste, já que valores fora dessas faixas se tornariam inaudíveis ou prejudiciais ao ouvido. Um tom de referência é criado com os parâmetros escolhidos. A seguir, o usuário escuta dois tons sucessivos: a referência (de frequência f) e um tom ligeiramente mais agudo (de frequência $f + \Delta f$), chamado de sinal de teste. O seu objetivo é determinar qual dos sinais escutados é o que possui maior frequência. Esse procedimento é repetido 20 vezes para diferentes valores de Δf .

Os valores de Δf aumentam geometricamente até um valor pré-estabelecido. A ordem com que aparecem para o usuário é aleatória e a ordem {sinal de referência, sinal de teste} também. Por fim, é criado um vetor de acertos¹, que é salvo em um arquivo `.mat`, o resultado é mostrado ao usuário em um gráfico de acertos por variação de frequência e o programa se fecha. Um exemplo de gráfico gerado é o da Fig. 3.16. Nesse caso, foram usadas frequência de 200 Hz e intensidade de 40 dB SPL.

A análise do resultado será feita no próximo capítulo.

¹Um acerto equivale a selecionar corretamente o sinal de maior frequência. A ele, é atribuído o valor um; a um erro, o valor zero.

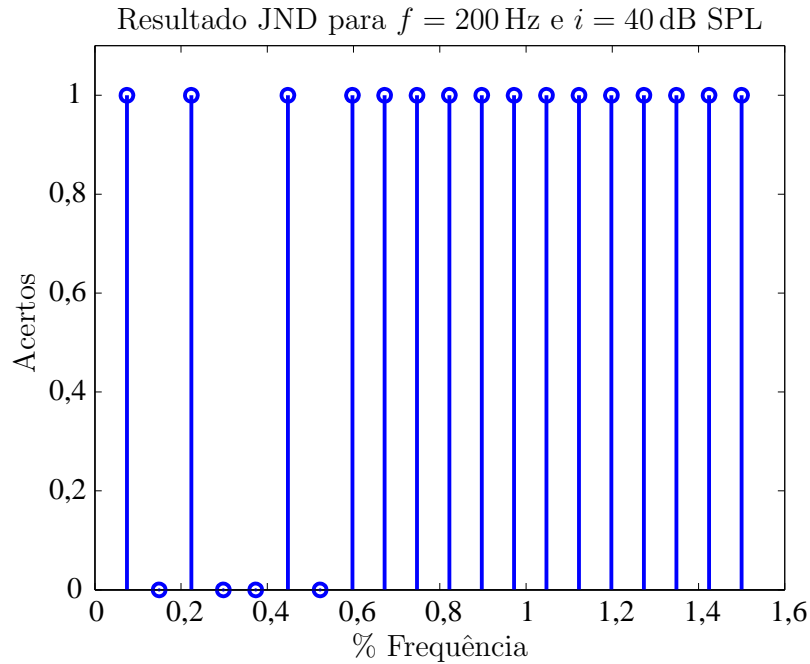


Figura 3.16: Exemplo de gráfico gerado pelo experimento “JND de Frequência”.

3.2.9 Mascaramento Temporal

Esse experimento ilustra o conceito de mascaramento temporal e permite ao usuário estimar uma curva de decaimento. Como explicado no Cap. 2, esta curva mostra, para cada intervalo de tempo entre os sinais tocados, a máxima intensidade que um sinal de teste de menor intensidade pode ter de forma que ele ainda seja mascarado. Os sinais utilizados são pulsos temporais. A interface gráfica do experimento pode ser vista na Fig. 3.17.

No teste, o usuário deve, para diferentes distâncias temporais entre os sinais mascarador e mascarado, selecionar a intensidade máxima do pulso fraco—mascarado—de forma que este não seja percebido na presença do pulso forte—mascarador. As distâncias temporais no pré-mascaramento vão de 3 até 30 ms e, no pós-mascaramento, de 5 até 190 ms. Estes valores foram selecionados de acordo com a teoria apresentada no Cap. 2. Os possíveis valores de intensidade do pulso mascarado variam de 0 até 50 dB SPL. Este possui uma intensidade pré-determinada. O mascaramento pode ocorrer antes (pré-mascaramento) ou após (pós-mascaramento) o pulso mas-

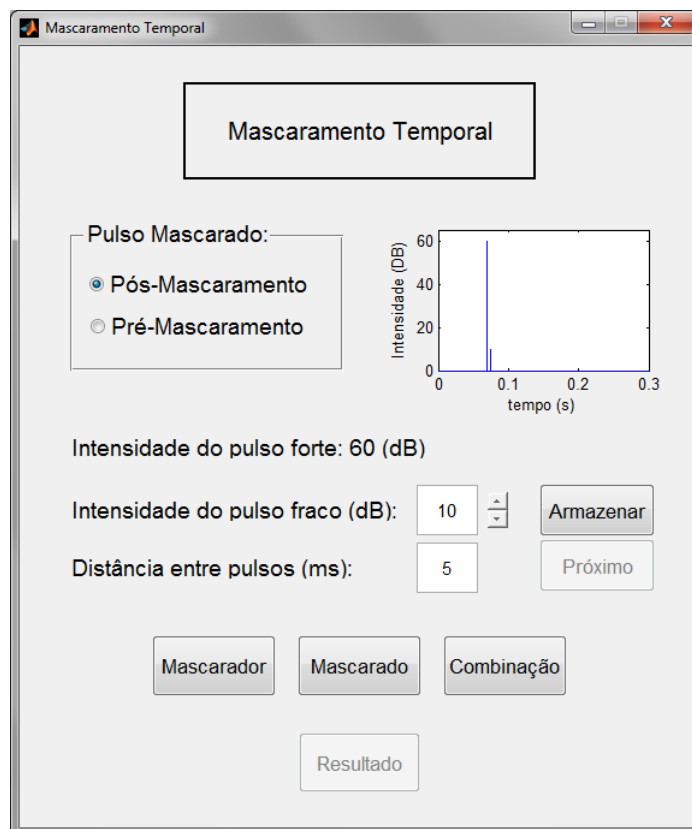


Figura 3.17: Interface gráfica do experimento “Mascaramento Temporal”.

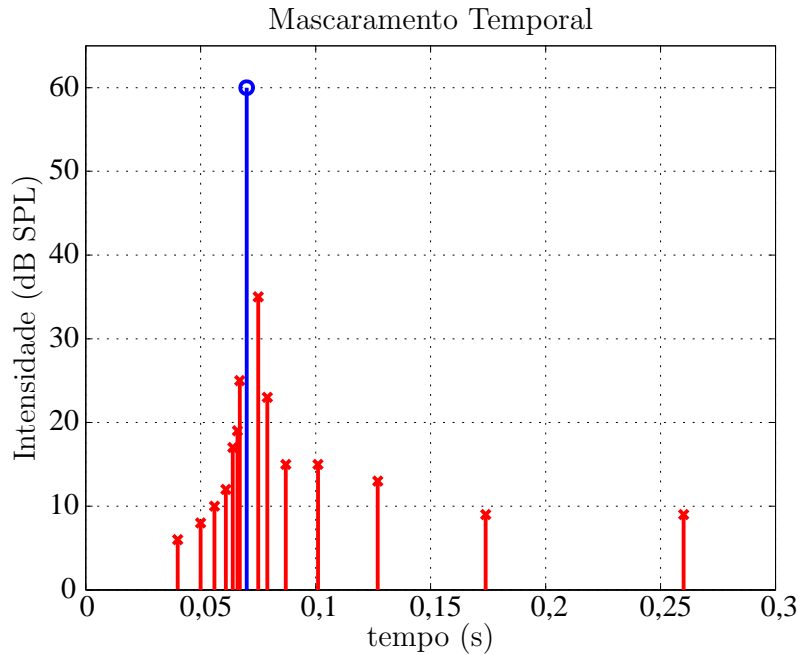


Figura 3.18: Exemplo de gráfico gerado pelo experimento “Mascaramento Temporal”.

carador. Por fim, os valores selecionados pelo usuários são salvos em um arquivo `.mat`, a curva de decaimento, mostrada em um gráfico de intensidade por tempo, é gerada e o programa se fecha. Um exemplo de curva é dado na Fig. 3.18.

A análise do resultado será feita no próximo capítulo.

3.2.10 Mascaramento Freqüencial

Esse experimento é análogo ao mascaramento temporal, porém, ilustrando o efeito do mascaramento no domínio da frequência. Os sinais utilizados são tons de 2 segundos. As interfaces gráficas do experimento podem ser vistas nas Figs. 3.19 e 3.20.

O usuário seleciona, primeiramente, a frequência do tom mascarador. A seleção é feita por um diálogo, visto na Fig. 3.20, e os valores possíveis de frequência são 500, 1000 e 2000 Hz. Em seguida, o usuário deve, para diferentes distâncias em frequência entre os tons, ajustar a intensidade máxima do tom fraco de forma que

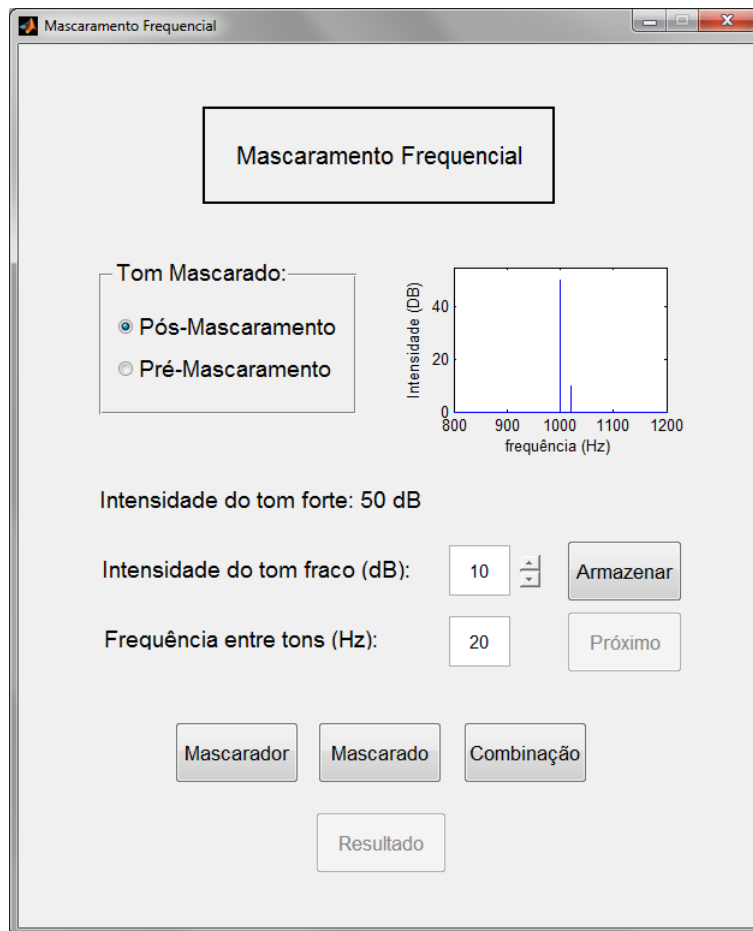


Figura 3.19: Interface gráfica principal do experimento “Mascaramento Freqüencial”.

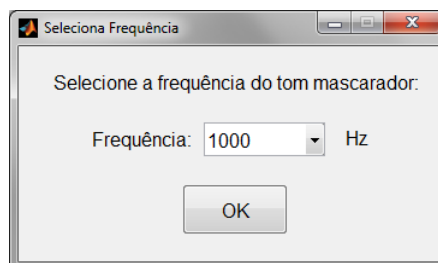


Figura 3.20: Diálogo pertencente ao experimento “Mascaramento Freqüencial”.

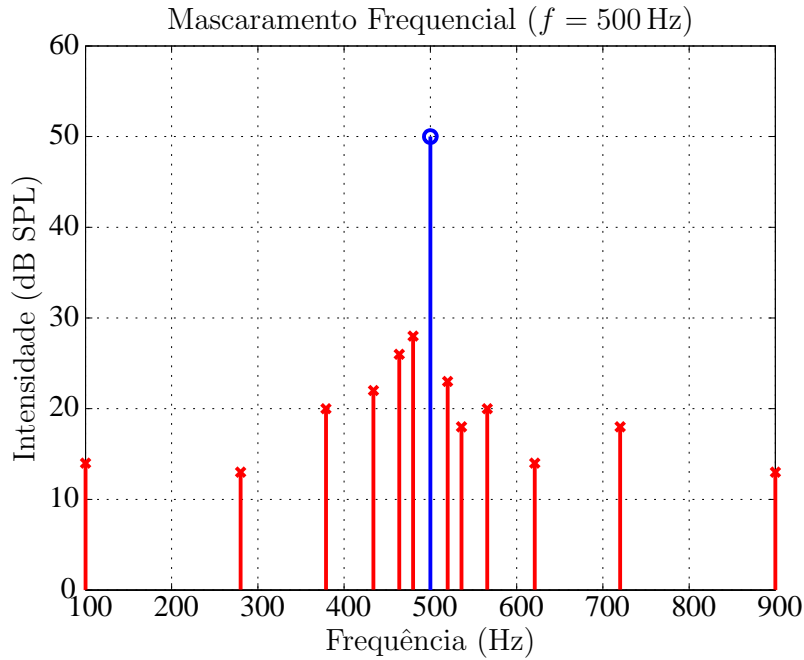


Figura 3.21: Exemplo de gráfico gerado pelo experimento “Mascaramento Freqüencial”.

este não seja percebido na presença do tom mais forte, que possui uma intensidade pré-determinada. Novamente, o mascaramento pode ocorrer antes ou depois do tom mascarador. As distâncias em módulo variam de 20 até 400 Hz. Os valores de intensidade podem variar de 0 até 50 dB SPL. Por fim, os valores selecionados pelo usuários são salvos em um arquivo `.mat`, a curva de decaimento, mostrada em um gráfico de intensidade por frequência, é gerada e o programa se fecha. Um exemplo de curva é o dado na Fig. 3.21. Nesse caso, a frequência do mascarador é 500 Hz.

A análise do resultado será feita no próximo capítulo.

3.2.11 Mascaramento Binaural

Esse experimento consiste em ilustrar o conceito de mascaramento binaural. São representados aqui os quatro casos ilustrados no Cap. 2 e que podem ser vistos na Fig. 2.7. No programa, no entanto, foi usado ruído branco no lugar do tom mascarador. As interfaces gráficas do experimento podem ser vistas nas Figs. 3.22 e 3.23.

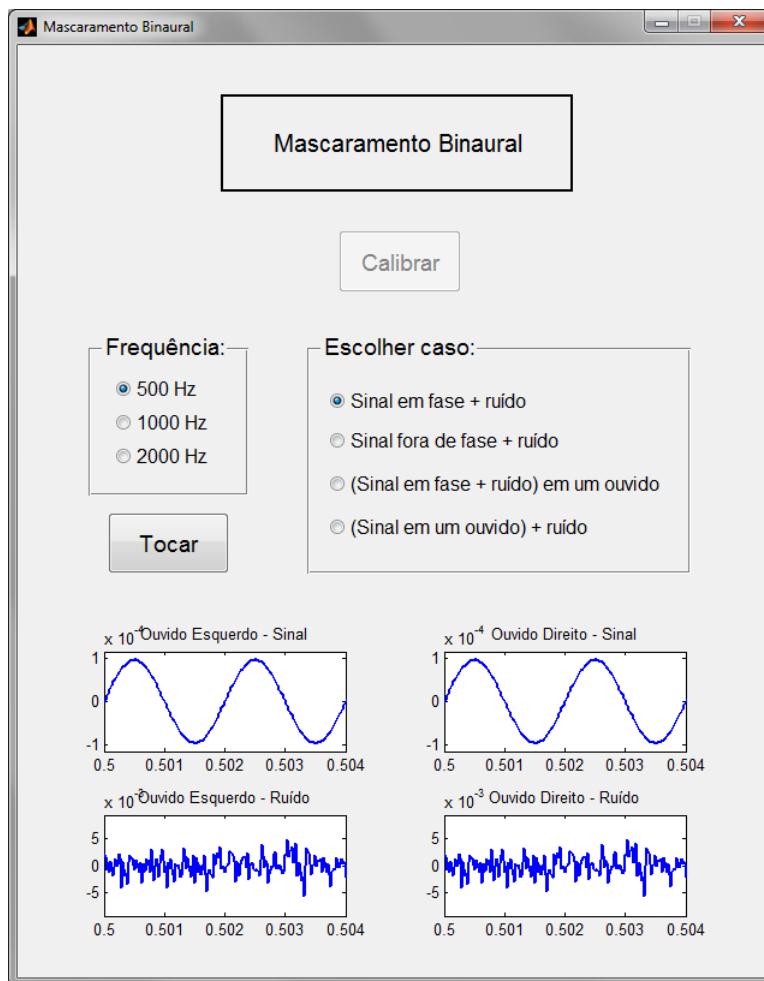


Figura 3.22: Interface gráfica principal do experimento “Mascaramento Binaural”.

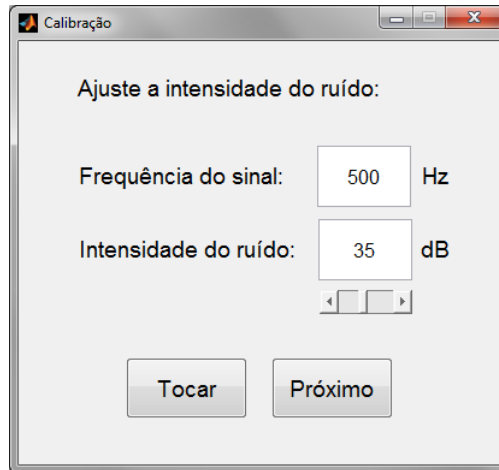


Figura 3.23: Diálogo pertencente ao experimento “Mascaramento Binaural”.

No programa, o usuário deve, inicialmente, fazer uma calibração. Ela é realizada por um diálogo (sua interface é dada na Fig. 3.23) e consiste em, para três valores diferentes de frequência do tom mascarado—500, 1000 e 2000 Hz—, regular a intensidade do ruído de forma que esta seja apenas suficiente para mascarar o tom. Os valores de intensidade podem variar de 0 até 80 dB SPL. A seguir, o usuário pode visualizar cada um dos quatro casos para as três frequências de pulso fraco calibradas.

3.2.12 Efeito de Precedência

Esse experimento consiste em ilustrar o conceito de efeito de precedência. A interface gráfica do experimento pode ser vista na Fig. 3.24.

O usuário pode verificar esse efeito com três sinais diferentes: uma voz falando, uma voz cantando e uma nota de violino. Também é possível ajustar o atraso (em ms) e escolher em qual dos ouvidos este será aplicado. O atraso varia de 0 ms até 100 ms. Percebe-se, à medida que o atraso aumenta, a transição entre escutar a fonte como vindo de uma posição entre os alto-falantes, escutá-la como vindo somente de um deles e escutar a presença do eco (fonte vindo de um dos alto-falantes e, um tempo depois, vindo do segundo alto-falante).

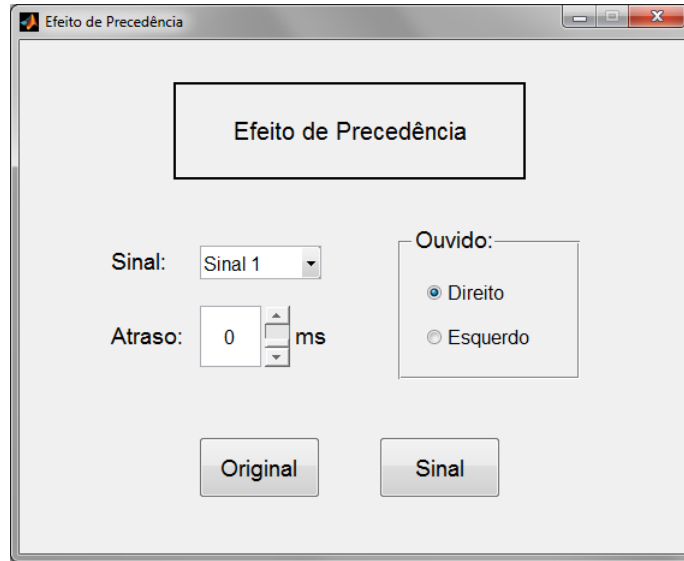


Figura 3.24: Interface gráfica do experimento “Efeito de Precedência”.

3.3 Detalhes da Implementação

Nesta seção, serão discutidos dois detalhes mais específicos na implementação dos programas referentes aos mascaramentos temporal e frequencial. Serão discutidas, também, algumas das principais funções da interface gráfica, como botões, *sliders* e alguns outros elementos.

No mascaramento temporal, os pulsos criados são, na verdade, pulsos triangulares. Estes foram escolhidos pois, no caso de pulsos retangulares, o que se escutava eram dois estalos: um no início e outro no final do pulso.

No mascaramento frequencial, os sinais mascarados são gerados ao passar um ruído branco por um filtro sintonizado na frequência desejada. Esse procedimento é feito para reduzir o efeito do batimento², que ocorre quando os sinais mascarado e mascarador estão muito próximos em frequência.

²O batimento é um fenômeno que ocorre quando dois sons com frequências muito próximas são escutados simultaneamente. O que percebemos são variações periódicas—com período igual ao inverso da diferença entre as frequências dos dois sinais—na intensidade do sinal resultante.

Para inserir e selecionar parâmetros, o usuário utiliza alguns elementos presentes na interface gráfica do experimento. Os *slider's* e caixas de texto são utilizados, na maior parte das vezes, para alterar os parâmetros de entrada como frequência e intensidade. Os botões possuem mais funcionalidades, dentre elas confirmar ou armazenar um valor escolhido pelo usuário, tocar algum sinal de áudio e ir para o próximo ponto a ser analisado no programa. Há, ainda, outras funções como menus e *radio button's*, que também são responsáveis pela seleção de parâmetros.

3.4 Considerações Finais

Este capítulo apresentou os testes psicoacústicos implementados e mostrou sua relação com a teoria explicada anteriormente. Além disso, alguns resultados foram apresentados. Estes serão interpretados com maior detalhes no próximo capítulo.

Como já foi dito, a importância dos testes é principalmente didática. A interface gráfica permite a um estudante de Psicoacústica comprovar de forma fácil e rápida os seus estudos teóricos. Assim, torna-se possível uma maior integração teoria-prática, que pode ser difícil com o uso de somente livros didáticos.

A maioria dos testes implementados cumpre corretamente seu objetivo. Eles ilustram os conceitos abordados e apresentam resultados compatíveis com a teoria. No entanto, como será visto, alguns testes podem produzir resultados que parecem inesperados. Todos esses pontos serão explorados mais a fundo no Cap. 4.

Capítulo 4

Testes Subjetivos e Resultados

4.1 Introdução

Após a implementação dos experimentos psicoacústicos, alguns deles foram selecionados para a realização de testes subjetivos. O objetivo é comprovar a funcionalidade dos experimentos. Para isso, um número significativo de participantes realizou testes, e alguns resultados foram gerados e comparados a outros já conhecidos.

Neste capítulo serão apresentados os experimentos selecionados e algumas pequenas modificações feitas em cada um. Será descrito, também, o procedimento adotado para os testes subjetivos. Por fim, serão apresentados resultados e uma análise de cada um deles.

4.2 Experimentos Selecionados

Foram selecionados quatro experimentos para a realização dos testes subjetivos. São eles: Mascaramento Temporal e Freqüencial, JND de Frequência e Curvas de Fletcher-Munson. Alguns destes testes sofreram pequenas mudanças para que o tempo total do teste subjetivo se ajustasse ao disponível. O tempo estimado é de 40 minutos, sem pausas.

Foi criada uma nova interface, dada na Fig. 4.1, na qual o nome do usuário é inserido juntamente com outros dois parâmetros. O primeiro é a frequência do mascarador a ser utilizado no experimento “Mascaramento Temporal”, que pode

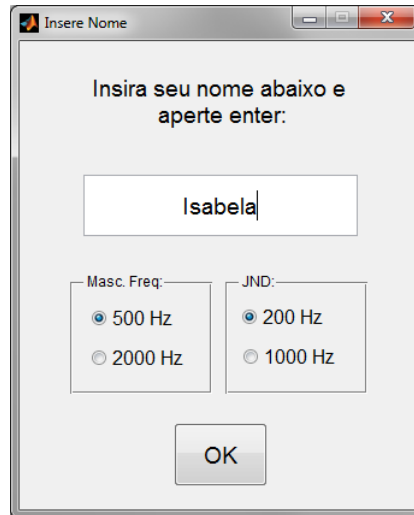


Figura 4.1: Interface gráfica inicial do teste subjetivo.

ser de 500 ou 2000 Hz; o segundo é a frequência do sinal de referência a ser utilizado no experimento “JND de Frequência”, que pode ser 200 ou 1000 Hz. Nesse ponto, o participante é inserido em um de dois grupos: o que possui os parâmetros {500 Hz, 200 Hz} ou o que possui os parâmetros {2000 Hz, 1000 Hz}.

Há, ainda, uma interface principal, que é responsável por interligar os testes. Esta é dada na Fig. 4.2. Inicialmente, o botão Calibração deve ser pressionado. A seguir, o botão Começar se ativa e uma contagem dos testes é realizada ao lado. Ao final de cada teste, o participante deve retornar a essa janela e pressionar novamente Começar para iniciar o próximo. Por fim, o texto deste botão se converte em “Fim” e, ao pressioná-lo, o teste subjetivo se encerra.

4.2.1 Calibração

A calibração possui a mesma interface apresentada na Seção 3.2.1. Sua função é a mesma, e os valores selecionados se mantêm por todo o teste subjetivo de um participante. A cada novo participante, esta etapa se repete. A calibração é sempre realizada pela mesma pessoa, nesse caso a pessoa que aplicou os testes subjetivos.

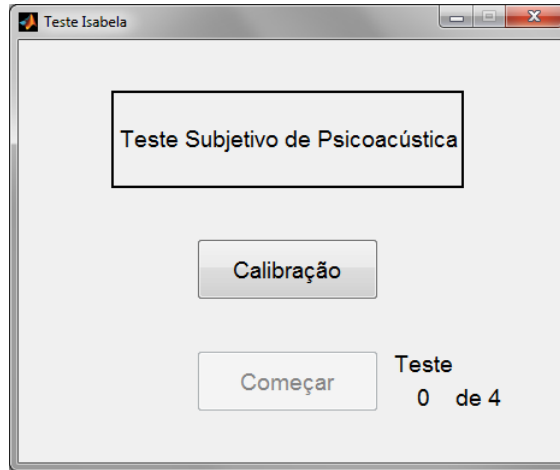


Figura 4.2: Interface gráfica principal pertencente ao teste subjetivo.

O procedimento descrito acima foi adotado pela impossibilidade prática de se realizar a calibração através de medidas da efetiva intensidade do sinal de áudio apresentado ao ouvinte. Entretanto, uma outra possibilidade seria o próprio participante realizar sua calibração. Dessa forma, os valores de referência para o cálculo das intensidades em dB seriam diferentes para cada ouvinte. Isso tornaria os testes subjetivos coerentes para cada um deles, porém incompatíveis com o cálculo de uma média global.

A estratégia adotada aqui foi, a princípio, tornar a amplitude de referência única para todos os participantes. O ambiente onde foram realizados os testes, porém, não apresentava condições ideais. Por isso, tornamos o procedimento um pouco mais flexível permitindo que a calibração se ajustasse, com o passar do tempo, às condições da sala.

4.2.2 Mascaramento Temporal

O primeiro experimento a ser realizado é o do “Mascaramento Temporal”. Algumas pequenas alterações foram realizadas neste programa. A primeira foi a redução do número total de pontos para 14, sendo 7 para cada caso—pulso fraco antes e depois do pulso forte. Os alcances de 3 até 30 ms e 5 até 190 ms para o pré- e pós-mascaramento, respectivamente, foram mantidos.

A segunda alteração é que, agora, o usuário é forçado a selecionar uma intensidade para passar para o próximo ponto. O botão Próximo, visto na interface da Fig. 3.17, permanece desativado até o usuário apertar o botão Selecionar, indicando que ele escolheu um valor de intensidade. O processo se repete para os 14 pontos (começando com o pós-mascaramento e seguindo, automaticamente, para o pré-mascaramento). Por fim, o usuário deve pressionar o botão Resultado, que estava desativado até então, para finalizar o teste.

4.2.3 Mascaramento Freqüencial

O segundo experimento a ser realizado é o do “Mascaramento Freqüencial”. A freqüência do mascarador será aquela selecionada no começo. O número de pontos também foi reduzido, passando a ser 12, sendo 6 para o pré-mascaramento e 6 para o pós-mascaramento. O alcance permaneceu de 20 até 400 Hz. O procedimento aqui é o mesmo que o do experimento anterior, ou seja, as alterações realizadas no programa original foram basicamente as mesmas.

4.2.4 JND de Freqüência

O terceiro experimento a ser realizado é o do “JND de Freqüência”. A única alteração feita aqui é que o usuário não seleciona a freqüência de referência, esta já foi especificada no começo. O número de pontos permanece em 20. O teste é realizado duas vezes para duas intensidades diferentes: 40 e 10 dB SPL, nessa ordem.

O usuário deve selecionar, como já foi explicado, o sinal de maior freqüência. A seguir, deve pressionar o botão Próximo, que se habilita após a escolha. Por fim, o texto presente no botão Próximo se converte em “Fim” e, ao pressioná-lo, o programa se fecha. Ao término da primeira vez, o programa se abre novamente, agora com uma intensidade de referência de 10 dB SPL. Ao término da segunda, o teste está finalizado.

4.2.5 Curvas de Fletcher-Munson

O quarto e último experimento a ser realizado é o do “Curvas de Fletcher-Munson”. A única alteração feita aqui é que o usuário não seleciona a intensidade

do harmônico de referência. Esta possui valores especificados de 50, 30 e 10 dB SPL, ou seja, o usuário deve realizar este teste para cada uma dessas intensidades, nessa ordem. Ao final das duas primeiras vezes, o programa abre novamente com um novo valor de intensidade. Ao final da terceira, o teste está finalizado.

4.3 Procedimento

O procedimento para a realização de cada experimento do teste subjetivo foi descrito na seção anterior. Aqui, será explicado o procedimento adotado para a sua aplicação.

Um roteiro foi elaborado e entregue para cada participante. Neste, há a ordem na qual as etapas do teste subjetivo devem acontecer:

- Inserir nomes e parâmetros;
- Realizar a Calibração;
- Experimento 1: Mascaramento Temporal;
- Experimento 2: Mascaramento Freqüencial;
- Experimento 3: JND de Frequência; e
- Experimento 4: Curvas de Fletcher-Munson.

Além disso, há a ordem na qual devem ser realizados os passos de cada teste e o número de vezes que cada um será executado. Para o experimento “Mascaramento Temporal”, por exemplo, há as seguintes instruções:

- Botão **M**ascarador;
- Botão **m**ascarado;
- Botão **C**ombinação;
- Ajustar **I**ntensidade;
- Botão **A**rmazenar;

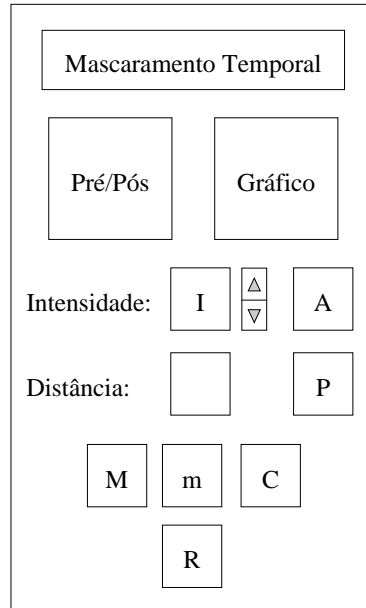
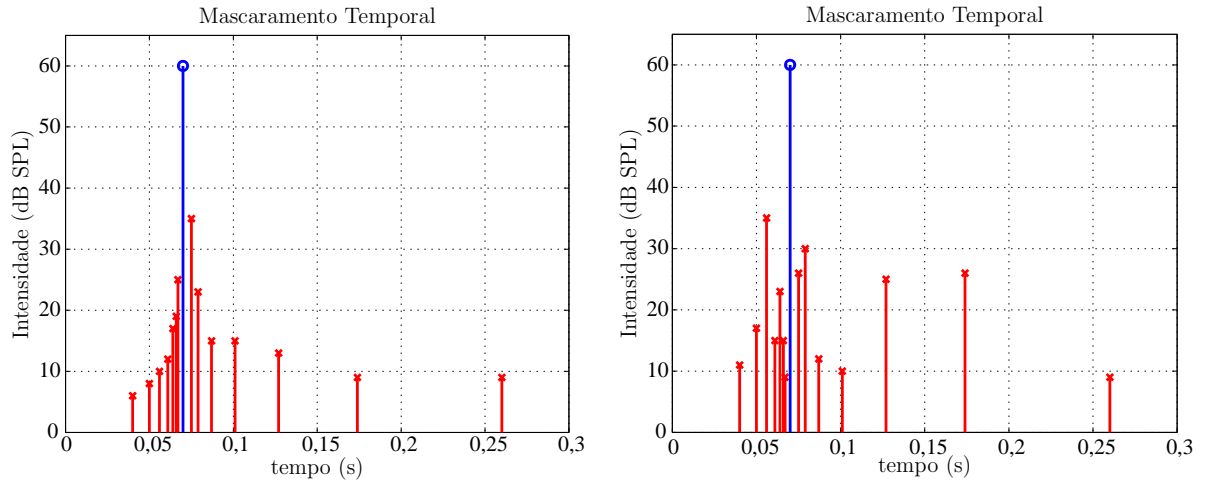


Figura 4.3: Ilustração da interface gráfica do experimento “Mascaramento Temporal”.

- Botão **P**róximo; e
- Botão **R**esultado.

Em conjunto com as instruções, foram dadas explicações sobre a funcionalidade de cada elemento da interface gráfica. Para escutar o sinal mascarador, por exemplo, o participante deve pressionar o botão Mascarador, para escutar o sinal mascarado, ele deve pressionar o botão mascarado, e assim por diante. Uma ilustração, como a dada na Fig. 4.3 para o Mascaramento Temporal, foi feita para cada teste e utilizada na explicação.

Após a explicação, o teste subjetivo iniciava. O participante fica, a partir desse momento, sozinho na sala para a realização dos experimentos. Ele é capaz, com as instruções passadas, de realizar todos os quatro testes. No caso de dúvida, pode consultar o roteiro elaborado ou consultar diretamente quem está aplicando o teste subjetivo, disponível durante sua execução. Ao final, os resultados são armazenados em vetores para serem posteriormente analisados.



(a) Curva com comportamento exponencial. (b) Curva com comportamento aparentemente aleatório.

Figura 4.4: Exemplos de curvas de decaimento obtidas no experimento “Mascaramento Temporal”. A primeira figura apresenta, visualmente, um resultado mais próximo do ideal, enquanto que a segunda, um resultado fora do esperado.

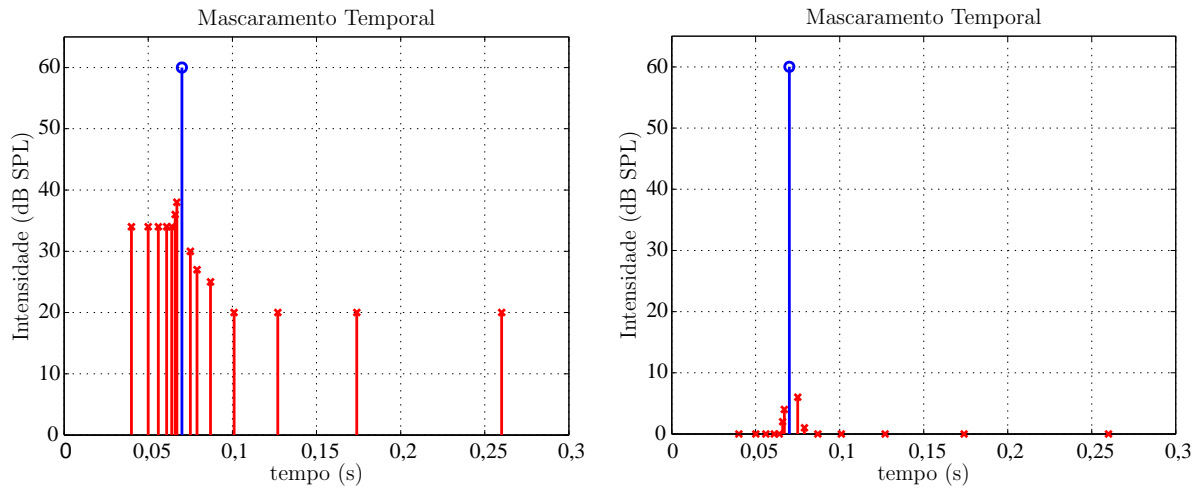
4.4 Resultados

Aqui serão apresentados os resultados obtidos com os testes subjetivos. O número de participantes foi de 24 e sua idade média é de 20 anos. Para cada experimento, serão mostrados gráficos com a média de todos os pontos. Serão apresentados também os resultados que tiveram um comportamento mais próximo ao esperado e alguns outros casos interessantes. Cada gráfico será, por fim, analisado.

4.4.1 Mascaramento Temporal

O resultado esperado para o experimento “Mascaramento Temporal” é composto, no pré- e no pós-mascaramento, por curvas exponenciais que decrescem com o aumento da distância temporal do pulso fraco ao pulso forte. No pré-mascaramento, essa queda é, normalmente, mais acentuada. O gráfico dado pela Fig. 4.4(a) é um bom exemplo desse resultado. Já o gráfico da Fig. 4.4(b) apresenta um comportamento aparentemente aleatório, fora da média.

Um ponto importante na análise dos resultados é o efeito claro da calibração presente em alguns deles. Exemplos são dados na Fig. 4.5. Como a calibração é sempre realizada pela mesma pessoa, isso significa que, para diferentes ouvintes, o limiar de



(a) Calibração com valores baixos para o ouvinte. (b) Calibração com valores altos para o ouvinte.

Figura 4.5: Efeitos da calibração nos resultados do experimento “Mascaramento Temporal”.

audição não estará em 0 dB SPL. Isso ocorre porque cada participante possui seu próprio limiar. Vemos, por exemplo, que na Fig. 4.5(a) após uma certa distância temporal, os valores de intensidade permaneceram os mesmos. Isso provavelmente ocorre pois essa intensidade corresponde ao limiar de audição desse ouvinte. Já na Fig. 4.5(b), os valores de intensidade caem rapidamente a 0 dB SPL e, como este é seu limite inferior, permanecem constantes. Novamente, isso é possivelmente explicado pelo fato de que o limiar de audição desse ouvinte é inferior ao limiar da pessoa que realizou a calibração.

A média e o desvio padrão dos resultados podem ser vistos no gráfico da Fig. 4.6. A média, em preto, se comporta visualmente como uma exponencial decrescente para o pré- e para o pós-mascaramento. Vemos que, para o pré-mascaramento, a curva de decaimento parece poder se estender até uma distância maior que 30 ms. O mesmo acontece no caso do pós-mascaramento, no qual a curva parece poder se estender até uma distância maior que 190 ms¹. A teoria vista no Cap. 2, no entanto, afirma que as curvas decaem até, no máximo, 20 e 200 ms para os casos pré- e pós-mascaramento, respectivamente. Como vimos, essa diferença no alcance se deve

¹A distância de 30 ms para o pré-mascaramento corresponde ao “x” preto mais à esquerda do gráfico. A distância de 190 ms para o pós-mascaramento corresponde ao “x” preto mais à direita do gráfico.

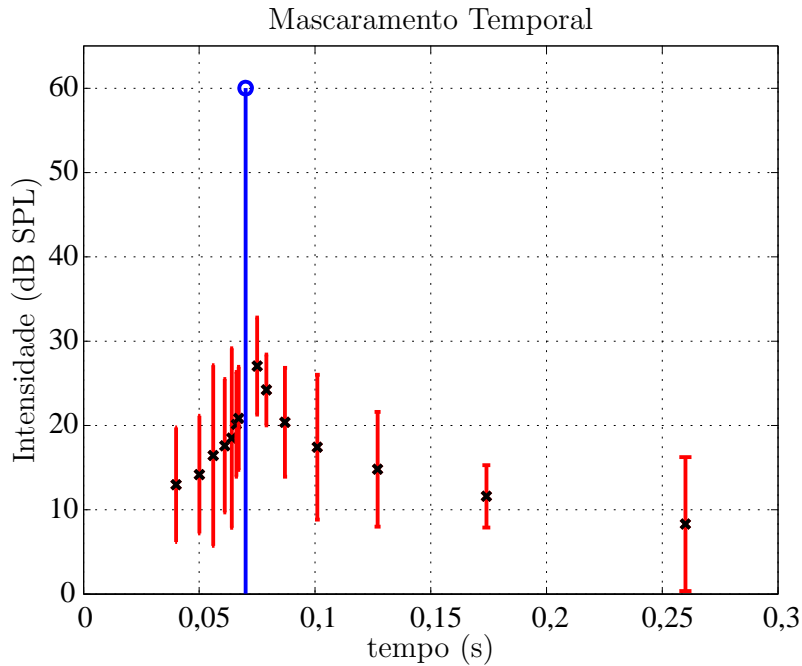


Figura 4.6: Média e desvio padrão dos resultados do experimento “Mascaramento Temporal”.

possivelmente à calibração.

Percebemos, ainda, que o desvio padrão dos pontos é elevado. Esse fato pode ser explicado pela própria variabilidade da percepção auditiva entre diversos indivíduos.

4.4.2 Mascaramento Freqüencial

O resultado esperado para o experimento “Mascaramento Freqüencial” é análogo ao do experimento anterior, mas com exponenciais decrescentes com o aumento da distância freqüencial. Vemos dois exemplos na Fig. 4.7 que correspondem a esse resultado. Percebemos no caso dos gráficos mostrados que, para a freqüência de mascarador de 500 Hz—Fig. 4.7(a)—, o espalhamento da curva é menor que para a freqüência de mascarador de 2000 Hz—Fig. 4.7(b). Esse resultado será comprovado mais adiante, quando analisarmos o gráfico com os valores médios.

Há alguns pontos importantes relacionados aos resultados obtidos a serem apresentados aqui. Um deles é a influência do limiar de audibilidade em boa parte dos

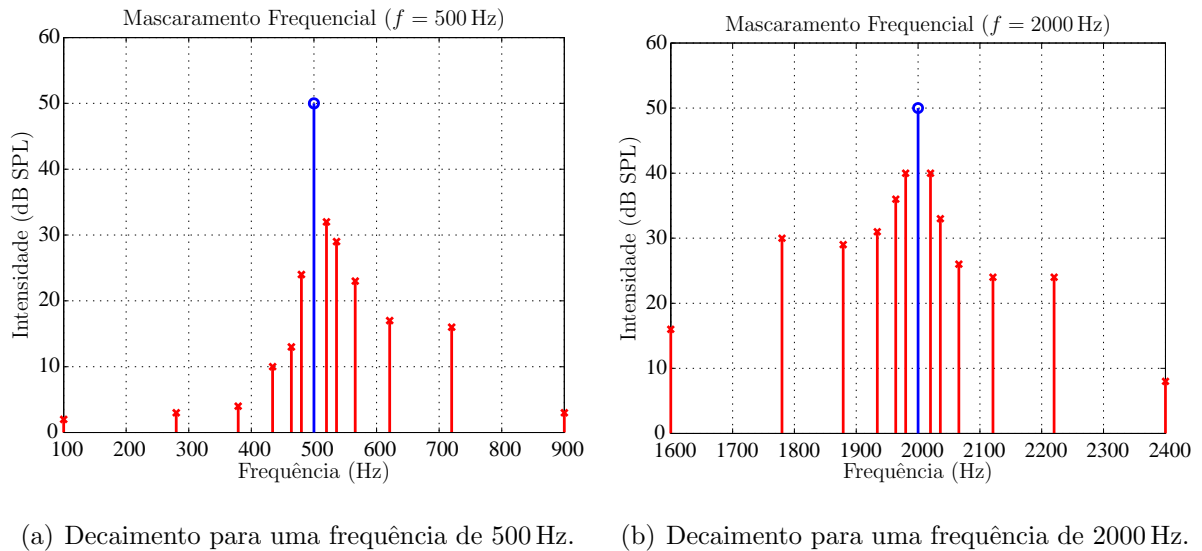
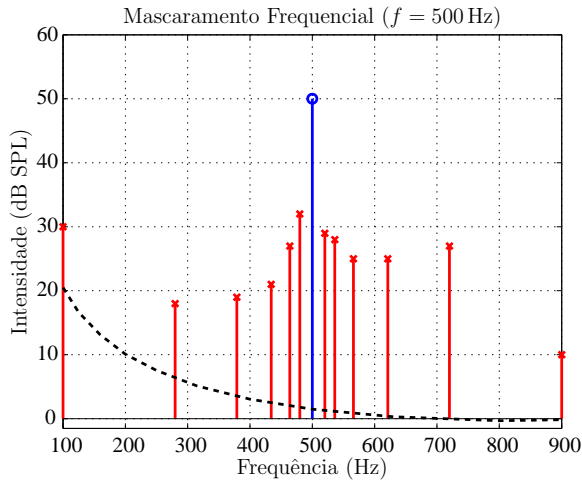


Figura 4.7: Exemplos de curvas de decaimento obtidas no experimento “Mascaramento Frecuencial”. Os gráfico foram obtidos realizando o experimento com frequências de mascarador iguais a 500 Hz e 2000 Hz para a primeira e segunda figura, respectivamente.

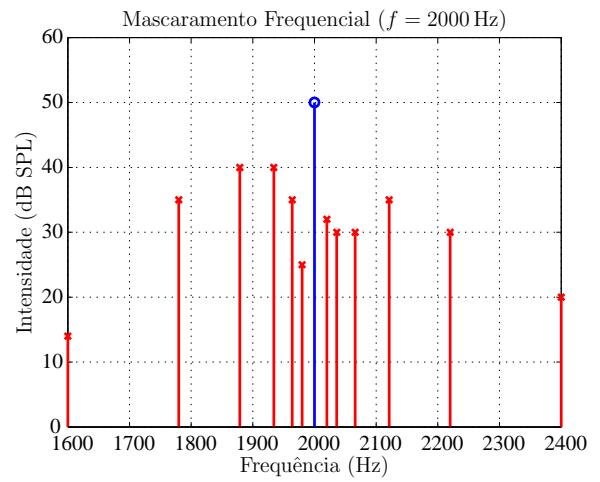
resultados do teste de frequência de mascarador de 500 Hz. Vemos na Fig. 4.8(a), por exemplo, que os valores de intensidade no pré-mascaramento só decrescem, com exceção do último ponto de 100 Hz. Ele possui uma intensidade maior, já que, nessa frequência, nosso limiar de audição, representado na curva tracejada em preto, também é maior.

Outro ponto a se observar é a presença do batimento, principalmente para os resultados do teste de frequência de mascarador de 2000 Hz. Neles, para pontos mais próximos do sinal mascarador, a intensidade é, muitas vezes, menor que para pontos mais distantes. Isso ocorre porque o batimento ocorre para duas senoides de frequências próximas. Um exemplo pode ser visto no gráfico da Fig. 4.8(b). É importante ressaltar, no entanto, que esse fenômeno ocorre em menor escala do que poderia ocorrer caso fosse utilizado como sinal mascarado um tom puro em vez do ruído de banda estreita.

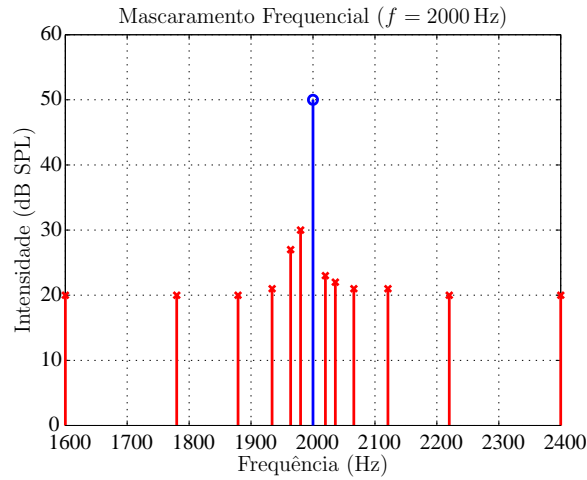
Por fim, vemos que o efeito causado pela calibração também está presente. Um exemplo é dado na Fig. 4.8(c), na qual a curva de decaimento, após uma determinada distância frequencial, fica constante.



(a) Efeito do limiar de audibilidade.

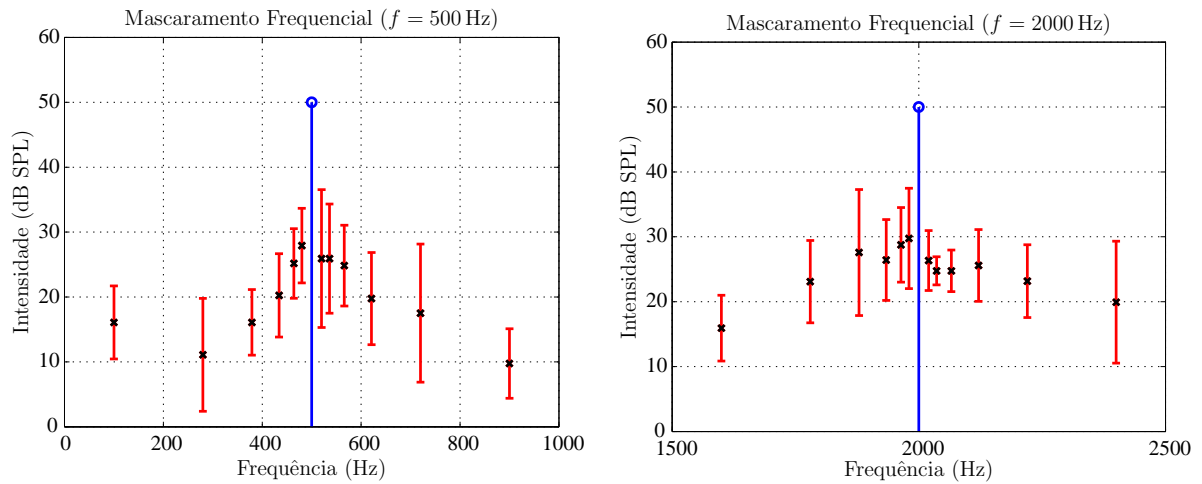


(b) Efeito do batimento.



(c) Efeito da calibração.

Figura 4.8: Efeitos do limiar de audibilidade, batimento e calibração nos resultados do experimento “Mascaramento Frecuencial”.



(a) Decaimento para uma frequência de 500 Hz. (b) Decaimento para uma frequência de 2000 Hz.

Figura 4.9: Média e desvio padrão dos resultados do experimento “Mascaramento Freqüencial”.

A média e o desvio padrão dos resultados podem ser vistos nos gráficos da Fig. 4.9. A média, em preto, se comporta visualmente como uma exponencial decrescente para os resultados do teste de frequência de mascaramento de 500 Hz, com exceção do ponto em 100 Hz. Como foi dito, isso pode ser explicado pelo aumento no nosso limiar de audição nessa região. Para o teste de frequência de mascaramento de 2000 Hz, a média dos pontos é decrescente, apesar de não ter um aspecto exponencial. Esse formato está relacionado à interferência do batimento nas frequências mais próximas ao mascarador. Vemos ainda que o decaimento é menor neste caso, ou seja, o espalhamento da curva de decaimento é maior para frequências de mascarador maiores. Como explicado no Cap. 2, essas curvas estão relacionadas às envoltórias das deflexões que ocorrem na membrana basilar na presença de um estímulo sonoro senoidal.

Como no experimento do “Mascaramento Temporal”, o desvio padrão dos pontos é elevado. A explicação é a mesma.

4.4.3 JND de Frequência

O resultado esperado para o experimento “JND de Frequência” para somente um ouvinte é que a taxa de acertos inicial—para as menores variações de frequência—seja de aproximadamente 50%. Nesse caso, o ouvinte ainda não foi capaz de iden-

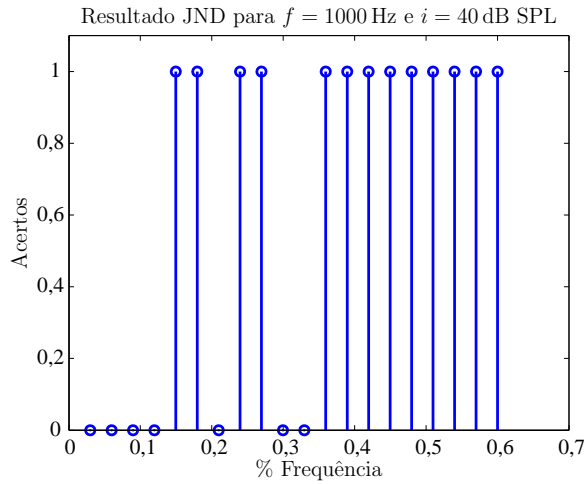
tificar a diferença entre os dois tons tocados e ele selecionou, arbitrariamente, um dos sinais como sendo o de maior frequência. A partir da variação de frequência que corresponde à sua JND, o ouvinte é capaz de identificar corretamente o tom mais agudo e, conseqüentemente, a sua taxa de acertos esperada é de 100%.

Alguns exemplos de resultados individuais são dados na Fig. 4.10. Nas Figs. 4.10(a) e 4.10(b), vemos como, para os primeiros pontos, os ouvintes erraram em torno de metade das vezes. No entanto, a partir de uma determinada variação de frequência, eles só obtiveram acertos. Nas Figs. 4.10(c) e 4.10(d), os ouvintes obtiveram muitos erros, mesmo para variações de frequência mais elevadas. Isso pode ter ocorrido por alguns motivos. O primeiro, mais provável, é o fato de a JND destes participantes ser maior que o alcance do teste. Há, no entanto, a possibilidade de erro na detecção do sinal mais agudo, apesar de a percepção da diferença entre os sinais escutados ser clara. Esse tópico será ilustrado com um exemplo mais à frente.

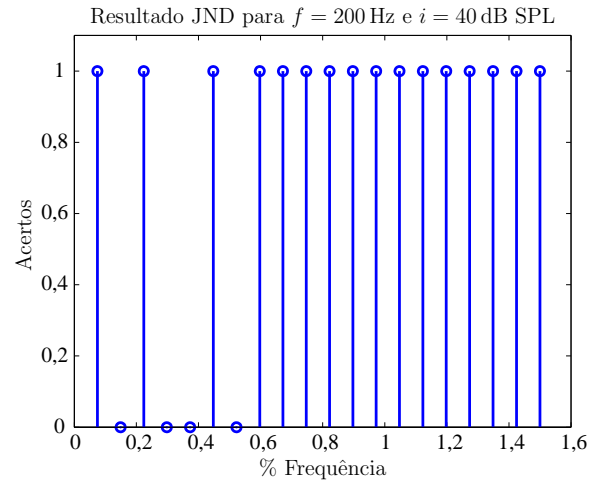
É importante ressaltar que, mesmo para os resultados das Figs. 4.10(a) e 4.10(b), há uma pequena probabilidade de que os acertos seguidos tenham sido ao acaso. Para diminuir essa probabilidade e obter uma estimação de sua JND, é necessário que ele repita o teste várias vezes e, em seguida, realize uma média de acertos para cada ponto. Por definição, a sua JND será aquela que corresponder à uma taxa de acertos média de 75% [3].

Alguns resultados, como os da Fig. 4.11, ilustram situações particulares. Na Fig. 4.11(a), o ouvinte parece possuir uma JND pequena e o alcance de frequência utilizado no teste foi, nesse caso, muito alto. Já para os testes das Figs. 4.10(b) e 4.10(d), vemos que o mesmo alcance foi ideal e ruim, respectivamente. A outra situação particular é dada na Fig. 4.11(b). Nela, vemos uma sequência de muitos acertos com apenas um erro entre eles. Como já foi dito, o que houve foi, provavelmente, um erro de detecção do sinal de maior frequência.

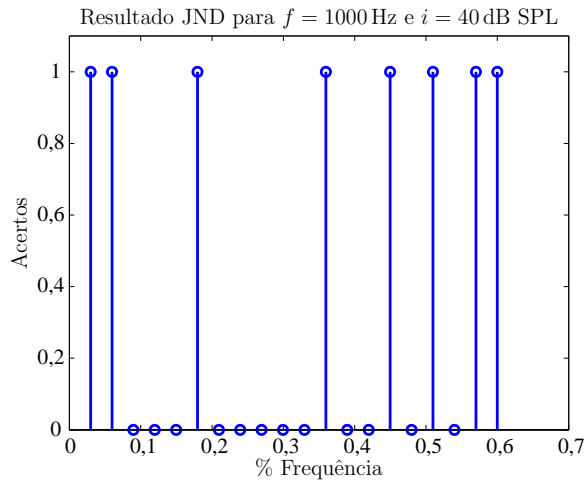
Por fim, com os resultados individuais obtidos, foram geradas as curvas da Fig. 4.12. Elas representam as médias entre todos os ouvintes para cada ponto. É, portanto, esperado que variem de 50% até 100%, onde o valor 50% corresponde a uma seleção



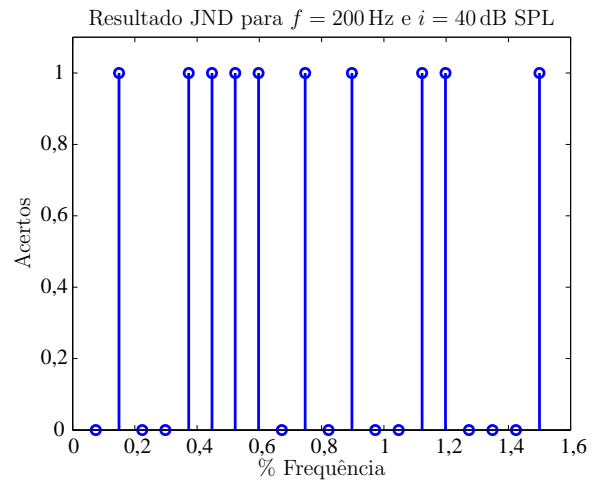
(a) Resultado ideal para frequência de 1000 Hz.



(b) Resultado ideal para frequência de 200 Hz.

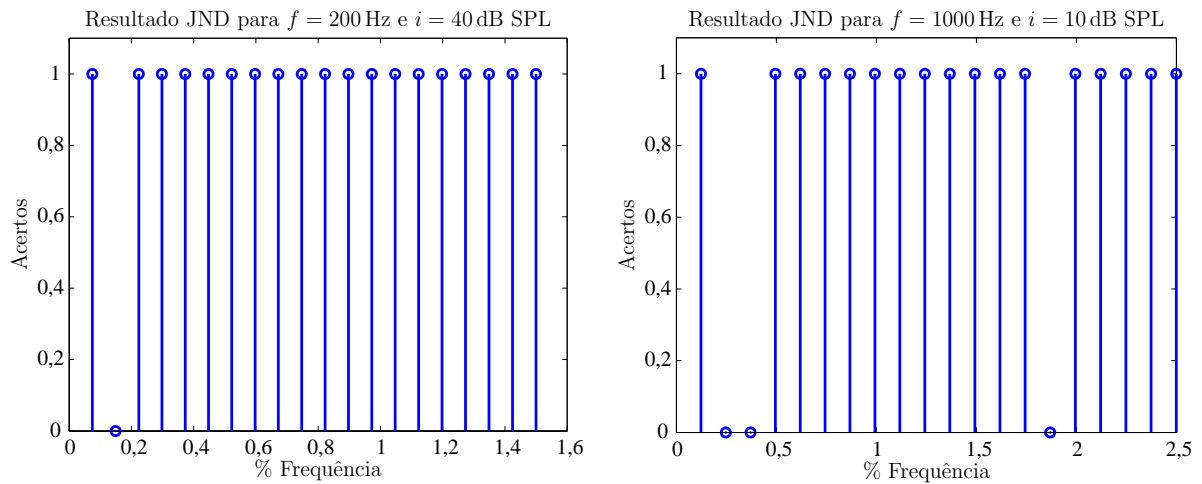


(c) Resultado aparentemente aleatório para frequência de 1000 Hz.



(d) Resultado aparentemente aleatório para frequência de 200 Hz.

Figura 4.10: Exemplos de resultados obtidos no experimento “JND de Frequência”. As figuras de cima ilustram resultados próximos dos ideais e, as de baixo, resultados com um comportamento aleatório. O valor um corresponde a um acerto e o valor zero, a um erro.



(a) Resultado para um ouvinte com JND pequena.

(b) Resultado com erro de detecção.

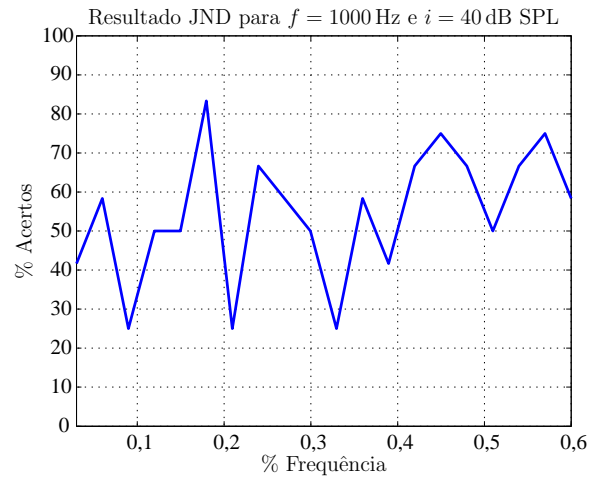
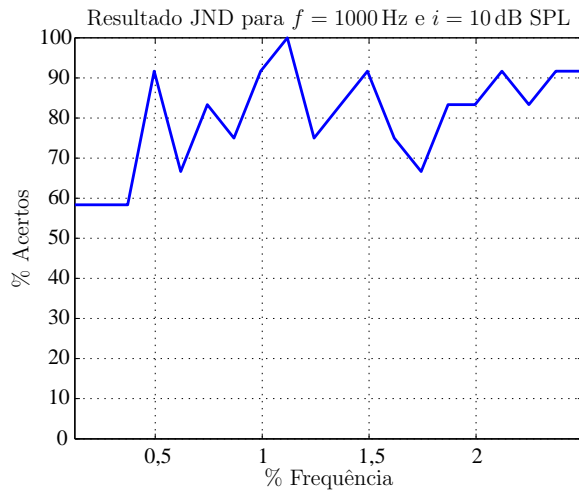
Figura 4.11: Exemplos de resultados particulares obtidos no experimento “JND de Frequência”. O valor um corresponde a um acerto e o valor zero, a um erro.

arbitrária e o valor 100%, a uma seleção certa. A variação de frequência correspondente ao valor de 75%, como mencionado acima, será a JND.

As curvas mostradas possuem um aspecto ruidoso e algumas dão pouca informação, como as das Figs. 4.12(b) e 4.12(d). Vemos que as Figs. 4.12(a) e 4.12(c), apesar de ruidosas, apresentam uma inclinação positiva, e seus valores finais são consideravelmente maiores que os iniciais. Dentre outros motivos, como os citados acima, uma causa provável do comportamento das curvas é o número pequeno de participantes. Foram, no total, 24 e estes foram divididos em dois grupos. Assim, para cada frequência de referência—200 e 1000 Hz—, somente 12 participantes realizaram os testes. Diante dos resultados inconclusivos, achamos desnecessário tabulá-los.

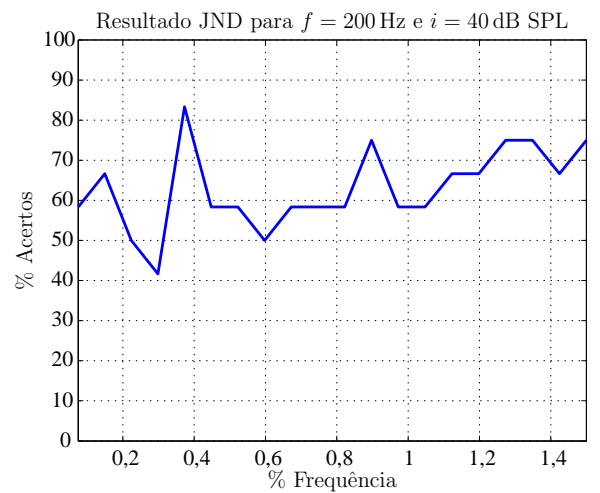
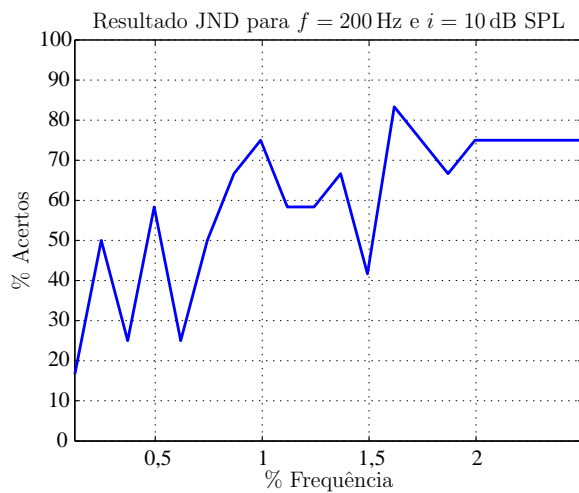
4.4.4 Curvas de Fletcher-Munson

O resultado esperado para o experimento “Curvas de Fletcher-Munson” são pontos pertencentes às curvas de audibilidade apresentadas no Cap. 2. Estas podem ser vistas em vermelho na Fig. 4.13. Os pontos em azul são as intensidades selecionadas pelos ouvintes. O ponto em preto é a intensidade, fixa durante todo o experimento, do harmônico de referência (tom puro de frequência 880 Hz). Deve-se notar que este ponto traduz a calibração feita pelo aplicador dos testes, que não necessariamente corresponde ao limiar de percepção da pessoa sob teste. Assim sendo, serviu apenas



(a) Resultado para frequência de 1000 Hz e intensidade de 10 dB SPL.

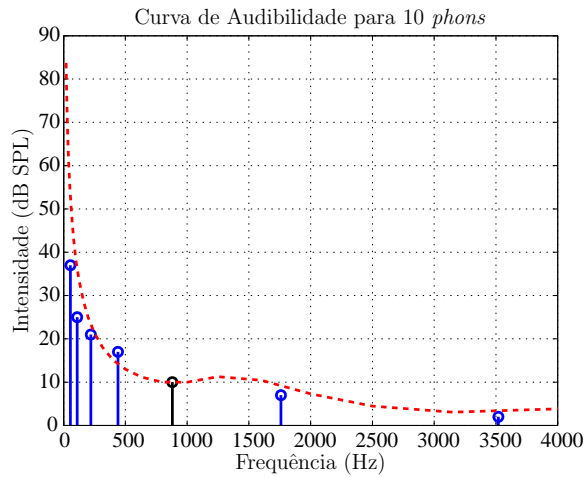
(b) Resultado para frequência de 1000 Hz e intensidade de 40 dB SPL.



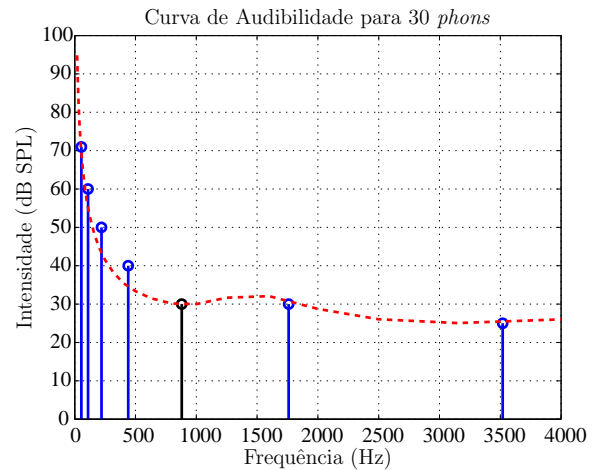
(c) Resultado para frequência de 200 Hz e intensidade de 10 dB SPL.

(d) Resultado para frequência de 200 Hz e intensidade de 40 dB SPL.

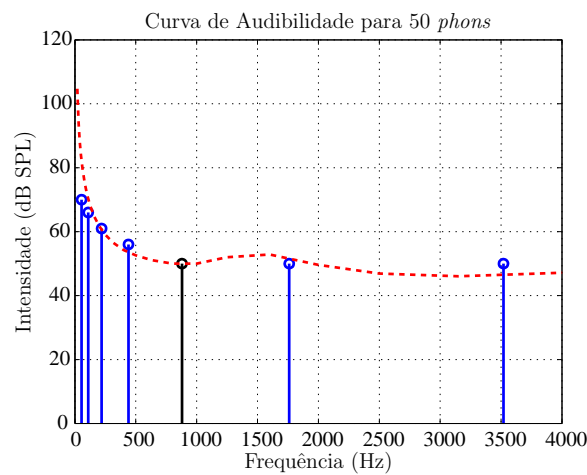
Figura 4.12: Curvas obtidas com os resultados do experimento “JND de Frequência”.



(a) Resultado para intensidade de 10 dB SPL.



(b) Resultado para intensidade de 30 dB SPL.

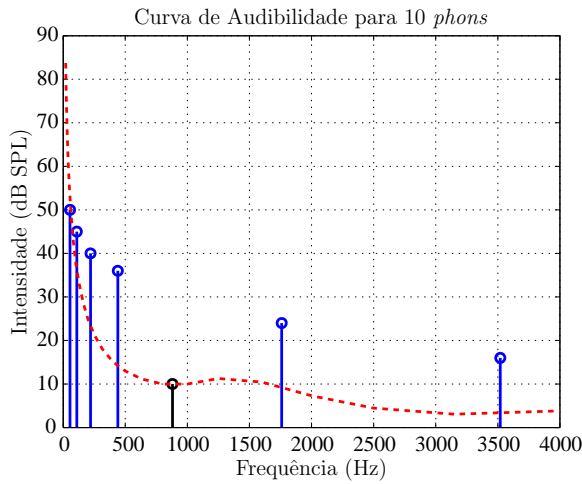


(c) Resultado para intensidade de 50 dB SPL.

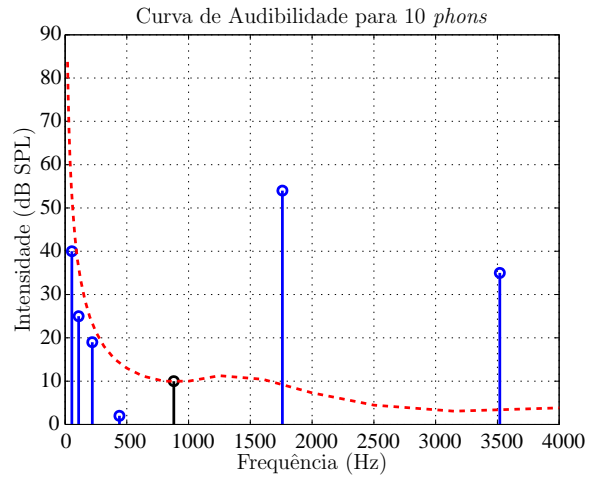
Figura 4.13: Exemplos de resultados obtidos no experimento “Curvas de Fletcher-Munson”. Em preto, o nível de calibração.

como “âncora” para a curva (vermelha) de referência, comum a todos os gráficos. As Figs. 4.13(a), 4.13(b) e 4.13(c) representam, assim, resultados bastante próximos ao esperado.

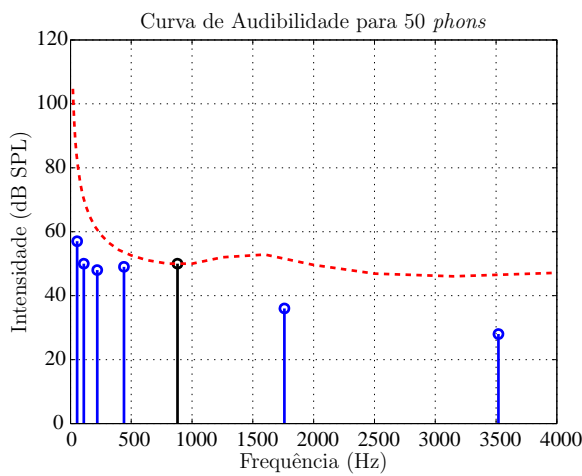
De forma geral, a maioria dos resultados observados apresentou um comportamento próximo da média. Há, no entanto, alguns casos específicos como os da Fig. 4.14. Na Fig. 4.14(a), os pontos azuis estão ligeiramente acima da curva vermelha. Isso ocorre para pessoas com menor sensibilidade. Há ainda o exemplo da Fig. 4.14(b), onde somente alguns pontos apresentam valores maiores. Nesse caso, o ouvinte possui menor sensibilidade nessas frequências. Nos casos das Figs. 4.14(c) e 4.14(d), por outro lado, as intensidades selecionadas foram abaixo da média. No



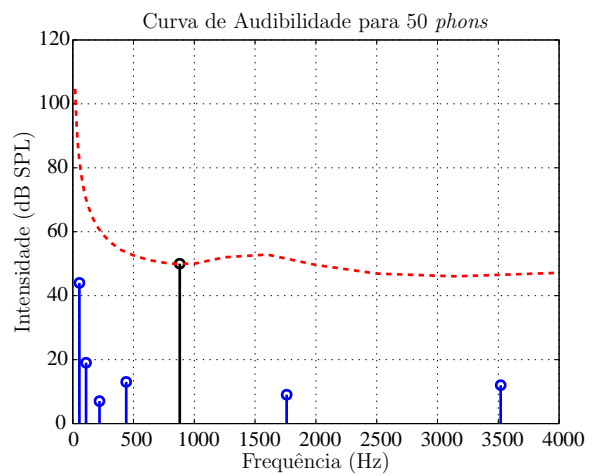
(a) Resultado acima da média.



(b) Resultado com pontos fora do padrão.



(c) Resultado abaixo da média.

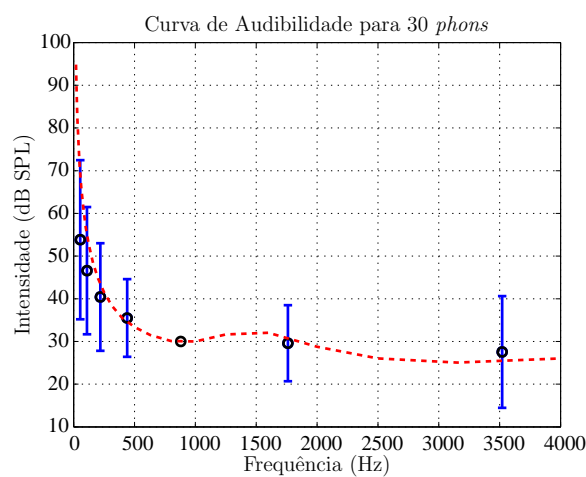
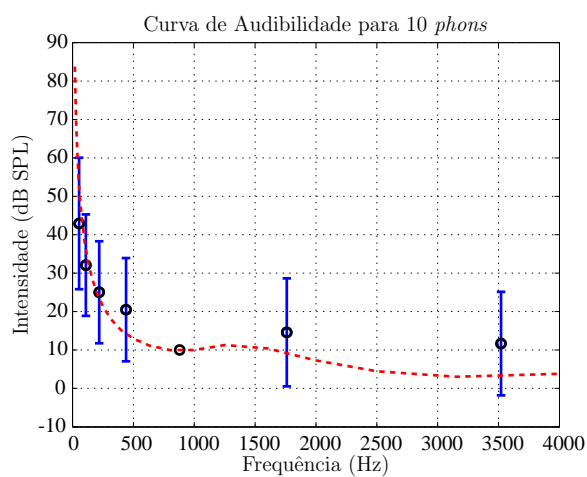


(d) Resultado muito abaixo da média.

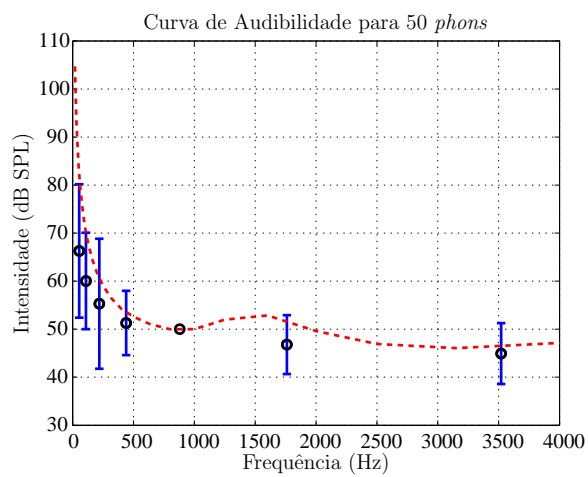
Figura 4.14: Exemplos de resultados particulares obtidos no experimento “Curvas de Fletcher-Munson”. Em preto, o nível de calibração.

segundo caso, esse fato é bastante evidente. Isso ocorre para ouvintes com uma maior sensibilidade.

É importante notar que esses efeitos percebidos ocorrem, principalmente, para as curvas de 10 e 50 dB SPL, respectivamente. O resultado disso pode ser visto nas médias de todos os ouvintes dadas em preto na Fig. 4.15. Vemos que, para uma intensidade de referência de 30 dB SPL, a média é bastante próxima à curva em vermelho.



(a) Resultado médio para intensidade de 10 dB SPL. (b) Resultado médio para intensidade de 30 dB SPL.



(c) Resultado médio para intensidade de 50 dB SPL.

Figura 4.15: Média e desvio padrão dos resultados do experimento “Curvas de Fletcher-Munson”. Sem a barra de erro, o nível de calibração.

4.5 Considerações Finais

Este capítulo apresentou os resultados obtidos com os testes subjetivos, e realizou uma análise detalhada destes. Alguns resultados individuais foram utilizados para ilustrar bons e maus exemplos, ou seja, casos próximos e distantes do esperado, e outros casos particulares.

Podemos perceber que os resultados estiveram, para a maioria dos programas selecionados, próximos ao esperado, ou seja, são compatíveis com a análise teórica feita no Cap. 2. De forma geral, um aumento no número de participantes envolvidos deveria trazer uma melhora nos resultados, sem que fosse necessário, para isso, alterar o *software* desenvolvido. No experimento “JND de Frequência”, um maior número de participantes tornaria as curvas dadas na Fig. 4.12 mais suaves, já que reduziria os erros apresentados. Porém, tomar somente essa medida pode não ser o suficiente, já que o número de resultados fora do padrão, como os das Figs. 4.10(c) e 4.10(d), foi elevado. Uma possível alteração é o aumento tanto do número de pontos utilizados—atualmente igual a 20—, como do alcance em frequência selecionado, principalmente para o caso em que a intensidade de referência é igual a 40 dB SPL.

Concluimos, com isso, que os programas selecionados podem ser utilizados tanto para a aplicação de testes subjetivos como para fins didáticos. Há, no entanto, alguns detalhes relevantes que devem ser levados em consideração na análise dos resultados obtidos. Dentre eles, estão o batimento e o limiar de audibilidade no caso do experimento “Mascaramento Freqüencial”, os erros de detecção no caso do experimento “JND de Frequência” e a calibração no caso dos experimentos de mascaramento.

Capítulo 5

Conclusões

A Psicoacústica possui grande relevância, pois está diretamente relacionada a aplicações como compressão de áudio e análise musical. Nessa área, existem ainda diversos tópicos que não possuem uma teoria totalmente bem fundamentada, como é o caso da percepção de *pitch*. Isso dá margem à continuação de pesquisas científicas nesse campo, podendo gerar resultados que contribuam positivamente com as aplicações citadas aqui.

O objetivo principal deste trabalho foi implementar um conjunto de testes psicoacústicos com duas finalidades: possibilitar a realização de testes subjetivos e servir ao uso didático. Os programas foram implementados com base nos principais conceitos de Psicoacústica, apresentados no Cap. 2. Por ser uma área estudada tanto por alunos de engenharia como por alunos de música, a criação de um tal pacote de *software* para fins didáticos abrange um grande número de pessoas interessadas.

O trabalho procurou, de forma geral, ser bastante completo e abrangente nos temas selecionados para a geração dos programas. Estes, no entanto, são um pouco genéricos no sentido de que ilustram um determinado conceito, como o mascaramento frequencial, por exemplo, e não abordam detalhes mais finos, como a influência da duração temporal ou da intensidade do pulso mascarador nas curvas de decaimento. Um trabalho futuro é, portanto, agregar opções ao usuário de forma que este possa variar outros parâmetros de entrada e analisar comportamentos mais específicos da saída. Essa alteração, no entanto, deve ser feita de forma a não prejudicar a facilidade de manuseio dos experimentos.

O *software* selecionado para a criação dos experimentos foi o Matlab[®]. Ele é de difícil acesso, em especial para os estudantes de música. Uma ideia de trabalho futuro que seria capaz de contornar esse problema é a geração de arquivos executáveis. Isso facilitaria o acesso e, portanto, a distribuição do pacote de experimentos.

Os resultados obtidos no Cap. 4 mostram que os experimentos implementados estão de acordo com a teoria apresentada aqui. Isso mostra que, além de sua funcionalidade, os experimentos apresentam um fácil manuseio. Participaram do teste 24 pessoas e a média de praticamente todos os resultados obtidos se enquadra, com alguma margem de erro, dentro dos resultados esperados. Isso indica que, apenas com uma breve explicação inicial, os ouvintes foram capazes de realizar o teste sem dificuldades.

Como conclusão final, podemos dizer que os objetivos traçados foram alcançados e que o trabalho, como um todo, deve contribuir de forma positiva para os interessados em continuar (ou começar) os seus estudos em Psicoacústica.

Referências Bibliográficas

- [1] COOK, P. R., *Music, Cognition, and Computerized Sound*. Londres, Inglaterra, The MIT Press, 2001.
- [2] HARTMANN, W. M., *Signals, Sound, and Sensation*. Nova Iorque, EUA, Springer, 1998.
- [3] HARTMANN, W. M., “Pitch, periodicity, and auditory organization”, *Journal of the Acoustical Society of America*, v. 100, n. 6, pp. 3491–3502, Outubro 1996.
- [4] JENNIFER H. JOHNSON, CHRISTOPHER W. TURNER, J. J. Z., MARGOLIS, R. H., “Just noticeable differences for intensity and their relation to loudness”, *Journal of Acoustical Society of America*, v. 93, n. 2, pp. 983–991, Fevereiro 1993.
- [5] GELFAND, S. A., *Hearing: An Introduction to Psychological and Physiological Acoustics*. Londres, Inglaterra, Informa Healthcare, 2010.
- [6] BOSI, M., GOLDBERG, R. E., *Introduction to Digital Audio Coding and Standards*. Massachusetts, EUA, Kluwer Academic Publishers, 2003.
- [7] BREEBAART, J., FALLER, C., *Spatial Audio Processing*. Chichester, Inglaterra, John Wiley & Sons Ltd, 2007.
- [8] MOURA, L. F. N. D., *Estudo de Psicoacústica e suas Aplicações*. Projeto final de graduação, DEL/Poli, Universidade Federal do Rio de Janeiro, Abril 2006.
- [9] BILSEN, F. A., RITSMA, R. J., “Repetition Pitch and Its Implication for Hearing Theory”, *Acustica*, v. 22, n. 2, pp. 63–73, 1969/70.