

UNIVERSIDADE FEDERAL FLUMINENSE
DEPARTAMENTO DE ENGENHARIA DE TELECOMUNICAÇÕES

MÓNICA JAZMIN GUERRA CARRILLO

ANÁLISE DE SENSIBILIDADE DE PARÂMETROS
EM MODELOS ESTOCÁSTICOS DE JITTER NA
PRODUÇÃO DA VOZ

NITERÓI

2014

UNIVERSIDADE FEDERAL FLUMINENSE
DEPARTAMENTO DE ENGENHARIA DE TELECOMUNICAÇÕES

MÓNICA JAZMIN GUERRA CARRILLO

ANÁLISE DE SENSIBILIDADE DE PARÂMETROS EM MODELOS ESTOCÁSTICOS DE JITTER NA PRODUÇÃO DA VOZ

Dissertação de Mestrado apresentada
ao Programa de Pós-Graduação em
Engenharia de Telecomunicações da
Universidade Federal Fluminense como
requisito para a obtenção do Grau de
Mestre em Engenharia de Telecomunicações.
Área de concentração:
Sistemas de Telecomunicações

Orientador:
EDSON CATALDO

NITERÓI

2014

MÓNICA JAZMIN GUERRA CARRILLO

ANÁLISE DE SENSIBILIDADE DE PARÂMETROS EM MODELOS
ESTOCÁSTICOS DE JITTER NA PRODUÇÃO DA VOZ

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Telecomunicações da Universidade Federal Fluminense como requisito para a obtenção do Grau de Mestre em Engenharia de Telecomunicações. Área de concentração: Sistemas de Telecomunicações

Aprovada em Maio de 2014.

BANCA EXAMINADORA

Prof. Dr. Edson Luiz Cataldo Ferreira - Orientador, Universidade
Federal Fluminense-UFF

Prof^a. Dra. Natalia Castro Fernandes - Universidade Federal
Fluminense-UFF

Prof. Dr. Leonardo Alfredo Forero Mendoza
Pontifícia Universidade Católica do Rio de Janeiro-PUC-Rio

Niterói
2014

*Dedicatória(s): Aos meus pais, Noris e Leandro, minha irmã, Ruth, minhas sobrinhas,
Maria Verónica, Maria Victoria e Maria Laura e meus sobrinhos Joseph e Jacob.*

Agradecimentos

Agradeço primeiramente a Deus por todas as bênçãos derramadas durante toda minha vida, pelas oportunidades, dificuldades, conquistas, e por ter sempre colocado pessoas em minha vida que sempre contribuem para o meu desenvolvimento.

À minha família pelo carinho, apoio e incentivo que me permitiram chegar até aqui, em especial à minha irmã Ruth, que tanto contribuiu para a minha formação técnica, moral e espiritual, sem seu apoio esse trabalho não seria possível, obrigada por tudo.

Ao meu orientador Edson Cataldo, pela oportunidade e confiança demonstrada desde o início, pela orientação e extrema disponibilidade, pelo apoio e incentivo constante.

Aos meus amigos pelo apoio incondicional e pela compreensão nos momentos de ausência.

À CAPES pelo suporte financeiro.

Ao Curso de Pós-Graduação em Engenharia de Telecomunicações da Universidade Federal Fluminense que me concedeu esta grande oportunidade de aumentar meus conhecimentos e crescer ainda mais na minha vida profissional.

Resumo

O jitter, no contexto da fala, denota pequenas perturbações involuntárias do comprimento do ciclo glotal, e é um fenômeno que pode ser observado nos sinais de voz. Essa dissertação apresenta modelos estocásticos de Jitter, fazendo uma análise de sensibilidade dos parâmetros e discutindo diferentes medidas de Jitter. O Jitter foi modelado estocasticamente usando um modelo de fita da vibração glotal e combinando modelos de perturbações de frequência instantânea, ruído e microtremor vocal, baseados em modelos de literatura. Foram apresentados cinco modelos, sendo o quinto modelo o mais geral. Baseados nos modelos, uma análise de sensibilidade dos parâmetros foi realizada indicando faixas de valores para vozes normais e vozes com características de patologia. Diferentes formas de cálculo de Jitter foram discutidas. Verificou-se que a frequência fundamental considerada, assim como o tempo de simulação, e o tipo de Jitter usados influenciaram no resultado para classificar as vozes.

Palavras chaves: Jitter, Frequência fundamental, perturbação da frequência fundamental, Análise acústica.

Abstract

Jitter in the context of speech denotes small involuntary perturbations of the glottal cycle lengths, is a phenomenon that can be observed in speech signals. This dissertation presents stochastic models of Jitter, making a sensitivity analysis of parameters and discussing different measures of Jitter. Jitter was stochastically modeled using ribbon model of the vocal fold vibrations combined with models of the disturbances of the instantaneous frequency, white noise and vocal microtremor, based on models of literature. Five models were presented and the fifth is a general model. Based on the models, a sensitivity analysis of the parameters was performed indicating the band of values for normal voices and voices with characteristics of pathology. Different ways of calculating jitter were discussed. It was verified the fundamental frequency, the simulation time and the type of jitter influenced on the results to classify voices.

Keywords: Jitter, fundamental frequency, perturbation of fundamental frequency, acoustic analysis.

Lista de Figuras

2.1	Posicionamento das cordas vocais durante a respiração (A) e durante a produção da voz (B)	5
5.1	Modelo genérico de produção da fala.	15
5.2	Diagrama de fluxo de estimação do Cepstro.	16
5.3	Cepstro de um segmento de fala [69].	17
6.1	Variação da frequência fundamental medida entre ciclos glóticos contíguos (Jitter), Sinal da largura da glote correspondente ao sexo masculino.	29
8.1	Modelo da onda de mucosa das cordas vocais [43].	41
9.1	Gráfico com os valores de <i>jitt</i> para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	54
9.2	Gráfico com os valores de <i>jitta</i> para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	54
9.3	Gráfico com os valores de <i>rap</i> para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	55
9.4	Gráfico com os valores de <i>ppq5</i> para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	55
9.5	Gráfico com os valores de <i>jitt</i> para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	56
9.6	Gráfico com os valores de <i>jitta</i> para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	56
9.7	Gráfico com os valores de <i>rap</i> para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	57
9.8	Gráfico com os valores de <i>ppq5</i> para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	57

9.9	Gráfico com os valores de <i>jitt</i> para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	58
9.10	Gráfico com os valores de <i>jitta</i> para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	58
9.11	Gráfico com os valores de <i>rap</i> para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	59
9.12	Gráfico com os valores de <i>ppq5</i> para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.	59
9.13	Gráfico com os valores de <i>jitt</i> para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	63
9.14	Gráfico com os valores de <i>jitta</i> para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	63
9.15	Gráfico com os valores de <i>rap</i> para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000, 5000 e 10000.	64
9.16	Gráfico com os valores de <i>ppq5</i> para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	64
9.17	Gráfico com os valores de <i>jitt</i> para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	66
9.18	Gráfico com os valores de <i>jitta</i> para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	66
9.19	Gráfico com os valores de <i>rap</i> para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	67
9.20	Gráfico com os valores de <i>ppq5</i> para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	67
9.21	Gráfico com os valores de <i>jitt</i> para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	69
9.22	Gráfico com os valores de <i>jitta</i> para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	69
9.23	Gráfico com os valores de <i>rap</i> para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	70

9.24	Gráfico com os valores de <i>ppq5</i> para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.	70
9.25	Gráfico com os valores de <i>jitt</i> para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	73
9.26	Gráfico com os valores de <i>jitta</i> para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	73
9.27	Gráfico com os valores de <i>rap</i> para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	74
9.28	Gráfico com os valores de <i>ppq5</i> para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	74
9.29	Gráfico com os valores de <i>jitt</i> para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	75
9.30	Gráfico com os valores de <i>jitta</i> para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	75
9.31	Gráfico com os valores de <i>rap</i> para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	76
9.32	Gráfico com os valores de <i>ppq5</i> para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	76
9.33	Gráfico com os valores de <i>jitt</i> para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	77
9.34	Gráfico com os valores de <i>jitta</i> para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	77
9.35	Gráfico com os valores de <i>rap</i> para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	78
9.36	Gráfico com os valores de <i>ppq5</i> para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.	78
9.37	Gráfico com os valores de <i>jitt</i> para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	81
9.38	Gráfico com os valores de <i>jitta</i> para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	81

9.39	Gráfico com os valores de <i>rap</i> para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	82
9.40	Gráfico com os valores de <i>ppq5</i> para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	82
9.41	Gráfico com os valores de <i>jitt</i> para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	83
9.42	Gráfico com os valores de <i>jitta</i> para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	83
9.43	Gráfico com os valores de <i>rap</i> para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	84
9.44	Gráfico com os valores de <i>ppq5</i> para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	84
9.45	Gráfico com os valores de <i>jitt</i> para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	85
9.46	Gráfico com os valores de <i>jitta</i> para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	85
9.47	Gráfico com os valores de <i>rap</i> para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	86
9.48	Gráfico com os valores de <i>ppq5</i> para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.	86
9.49	Gráfico com os valores de <i>jitt</i> para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000). . .	89
9.50	Gráfico com os valores de <i>jitt</i> para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000). . .	89
9.51	Gráfico com os valores de <i>rap</i> para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000). . .	90
9.52	Gráfico com os valores de <i>ppq5</i> para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000). . .	90

Lista de Tabelas

5.1	Grupo de locutores por tipo da voz (normal e com características patológicas) e do sexo.	20
5.2	Resultados obtidos pelo <i>Praat</i> , pelo <i>fxrapt</i> e pelo algoritmo criado com os diferentes métodos para os sinais da vogal sustentada /e/ de voz normal. .	21
5.3	Resultados obtidos pelo <i>Praat</i> , pelo <i>fxrapt</i> e pelo algoritmo criado com os diferentes métodos para os sinais da vogal sustentada /e/ de voz com presença de nódulo nas cordas vocais.	23
5.5	Resultados obtidos pelo <i>Praat</i> , pelo <i>fxrapt</i> e pelo algoritmo criado com os diferentes métodos para os sinais da vogal sustentada /e/ de voz com presença de paralisia nas cordas vocais.	23
5.6	As médias e a média das médias da frequência fundamental em sinais de voz masculinas normais.	24
5.7	As médias e a média das médias da frequência fundamental em sinais de voz femininas normais	24
5.8	As médias e a média das médias da frequência fundamental em sinais de voz masculinas com presença de paralisia.	25
5.9	As médias e a média das médias da frequência fundamental em sinais de voz femininas com presença de paralisia.	25
5.10	As médias e a média das médias da frequência fundamental em sinais de voz femininas com presença de nódulos.	26
7.1	Resultados obtidos pelo algoritmo criado e pelo <i>Praat</i> para o Jitter dos sinais da vogal sustentada /e/ de voz normal.	33
7.2	Resultados obtidos pelo algoritmo criado e pelo <i>Praat</i> para o Jitter dos sinais da vogal sustentada /e/ de voz com presença de nódulos.	34

7.4	Resultados obtidos pelo algoritmo criado e pelo <i>Praat</i> para o Jitter dos sinais da vogal sustentada /e/ de voz com presença de paralisia nas cordas vocais.	35
9.1	Valores dos parâmetros de modelagem da análises de regressão múltipla linear.	53
9.2	Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitt</i> dentro da faixa considerada normal no Modelo I.	60
9.3	Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitta</i> dentro da faixa considerada normal no Modelo I.	60
9.4	Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>rap</i> dentro da faixa considerada normal no Modelo I.	60
9.5	Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>ppq5</i> dentro da faixa considerada normal no Modelo I.	61
9.6	Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de <i>jitt</i> dentro da faixa considerada normal no Modelo II.	71
9.7	Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de <i>jitta</i> dentro da faixa considerada normal no Modelo II.	71
9.8	Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de <i>rap</i> dentro da faixa considerada normal no Modelo II.	71
9.9	Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de <i>ppq5</i> dentro da faixa considerada normal no Modelo II.	72
9.10	Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitt</i> dentro da faixa considerada normal no Modelo III.	79

9.11 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitta</i> dentro da faixa considerada normal no Modelo III.	79
9.12 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>rap</i> dentro da faixa considerada normal no Modelo III.	79
9.13 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>ppq5</i> dentro da faixa considerada normal no Modelo III.	80
9.14 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitt</i> dentro da faixa considerada normal no Modelo IV.	87
9.15 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>jitta</i> dentro da faixa considerada normal no Modelo IV.	87
9.16 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>rap</i> dentro da faixa considerada normal no Modelo IV.	87
9.17 Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de <i>ppq5</i> dentro da faixa considerada normal no Modelo IV.	88

Sumário

1	Introdução	1
	Motivação	2
	Objetivos da dissertação	2
	Objetivo Geral	2
	Objetivos Específicos	2
	Estrutura da dissertação	3
2	Princípios de produção da voz humana	5
2.1	Funções da Laringe	6
2.1.1	Função respiratória	6
2.1.2	Função deglutitória	6
2.1.3	Função fonatória	6
2.2	Oscilação das cordas vocais	6
3	Avaliação Acústica	8
3.1	Avaliação Acústica	8
3.2	O programa Praat	9
3.3	Função fxrapt	9
4	Frequência fundamental	10
4.0.1	Estatísticas da frequência fundamental	10
4.0.2	Fatores relacionados nas variações da frequência fundamental	12
4.1	Intensidade vocal	13

4.2	Qualidade vocal	13
5	Métodos para determinação da frequência fundamental de sinais de voz	15
5.1	Determinação da frequência fundamental	15
5.2	Algoritmos usados para o cálculo da F_0	15
5.2.1	Determinação de F_0 pelo Método Cepstral	15
5.2.2	Determinação da F_0 pelo método da Autocorrelação	18
5.3	Comparação entre os métodos: aplicação em um banco de vozes	19
5.4	Considerações	24
6	Jitter	28
6.1	Jitter	28
6.1.1	Medidas de parâmetros de perturbação	28
6.1.2	Medidas de Jitter	29
6.1.3	Fatores relacionados nas variações do Jitter	30
6.1.4	Determinação do Jitter	30
7	Comparação dos valores de Jitter usando a função fxrapt e o software Praat	32
7.1	Determinação do Jitter	32
7.2	Considerações	35
8	Modelos estocásticos de Jitter	38
8.1	Propriedades estatísticas de Jitter	38
8.2	Modelos estocásticos	40
8.2.1	Modelo de fita na vibração das cordas vocais	40
8.2.2	Modelos das perturbações estocásticas de frequência instantânea	43
8.2.3	Modelos estocásticos de Jitter	45
8.2.3.1	Modelo I: Passeio aleatório	45

8.2.3.2	Modelo II: Ruído filtrado	47
8.2.3.3	Modelo III: Oscilador acoplado perturbado	47
8.2.3.4	Modelo IV: Onda da corda vocal perturbada	49
8.2.3.5	Modelo V: Modelo geral estocástico de Jitter	50
9	Simulações e resultados	51
	Modelo I	53
	Considerações	61
	Modelo II	61
	Considerações	72
	Modelo III	72
	Considerações	80
	Modelo IV	80
	Considerações	88
	Modelo V	88
	Considerações	91
10	Conclusões	92
	Referências	94
	Apêndice A – Apêndice	99
	Apêndice A – Rotina para o cálculo da f_0 pelo Método Cepstral	100
	Apêndice B – Rotina para o cálculo da f_0 pelo Método da autocorrelação	101
	Apêndice C – Rotina para o cálculo do Jitter	102
	Apêndice D – Rotinas da modelagem do Jitter	104
	D.1 Modelo I	104

D.2	Modelo II	105
D.3	Modelo III	106
D.4	Modelo IV	107
D.5	Modelo V	109

Capítulo 1

Introdução

A voz é um dos principais meios de comunicação humano e, como um sinal acústico, contém informações importantes sobre algumas características dos indivíduos. A estrutura biomecânica vocal, em associação com variáveis aerodinâmicas, desempenha um papel importante na produção da voz e está ligada às mudanças na qualidade vocal [60].

A qualidade da voz depende do modo de fechamento e abertura da glote e da vibração das cordas vocais. Certas alterações laríngeas impedem que as cordas vocais tenham uma vibração glotal harmônica.

A análise acústica da voz fornece medidas que estão relacionadas ao padrão vibratório das cordas vocais, a forma do trato vocal, e suas mudanças no tempo, sendo que sua interpretação varia com a idade, o sexo, o tipo de fonação e o treino vocal. Os resultados da análise acústica se diferenciam em cada população por fatores extrínsecos e intrínsecos e podem auxiliar no pré-diagnóstico de patologias relacionadas à voz.

A análise acústica da voz é capaz de fornecer a análise da frequência fundamental, definida como o número de vibrações por segundo produzidas pelas cordas vocais, medidas de perturbação da frequência, tais como: jitter; definido como sendo a perturbação da frequência fundamental ciclo a ciclo; medidas da perturbação da amplitude, shimmer, que é a variabilidade da amplitude ciclo a ciclo e as medidas de ruído espectral como HNR (do inglês, Harmonic-to-Noise Ratio), que correlaciona o componente harmônico com o componente ruído [69].

Motivação

A literatura da área de voz apresenta muitos estudos que lançam mão da análise acústica [11], [27], [28], [16], [51],. No entanto, são poucos os que procuram modelar estocasticamente o jitter [17].

Na literatura nacional, é escasso o número de trabalhos que realizam a extração de dados acústicos das vogais do português brasileiro com uso de diferentes programas computadorizados.

O sistema fonatório não é “perfeito”, e todos os ciclos vibratórios dos falantes normais apresentam oscilações. No entanto, pode constatar-se que uma laringe com patologia produz ciclos vibratórios com maior alteração que uma sem patologia.

Modelos de jitter podem sugerir, ou confirmar, de forma matemática, que os indicadores que caracterizariam os comprimentos dos ciclos são perturbados estatisticamente ao invés de heurísticamente [59][63].

Entre outras motivações para o desenvolvimento de modelos de jitter temos que eles são parte de muitos sintetizadores de voz. Os modelos são projetados para melhorar a naturalidade ou imitar vozes roucas [57] [22] [7]. A maioria dos modelos existentes são apenas vagamente baseados em dados, e não simulam toda a gama de propriedades conhecidas jitter.

Outra aplicação interessante de modelos de jitter é a geração de sinais sintéticos para testar ou calibrar algoritmos de processamento de sinais que detectam os ciclos glóticos e as medidas de seus comprimentos [53].

Objetivos da dissertação

Objetivo Geral

- Simular modelos estocásticos de jitter e fazer a análise de sensibilidade dos parâmetros dos modelos.

Objetivos Específicos

- Elaborar e implementar algoritmos que determinam a frequência fundamental de vozes consideradas normais e com características de patologias pelo método Cepstral

e pelo método da Autocorrelação.

- Comparar os resultados dos algoritmos implementados para o cálculo da frequência fundamental pelo método Cepstral e pelo método da Autocorrelação com o programa de análise acústica *Praat* e a função *fxrapt* implementada no matlab.
- Elaborar e implementar um algoritmo que determine medidas de jitter (jitt, jitta, rap, ppq5) em vozes consideradas normais e com características de patologias.
- Comparar os resultados do algoritmo implementado para o calculo de jitter com o programa de análise acústica *Praat*.
- Simular como distúrbios das frequências instantâneas de vibração das paredes glotais causam perturbações dos comprimentos de ciclo glotal.
- Calcular comprimentos dos ciclos das perturbações quando a frequência instantânea é perturbada por um processo de ruído branco sem correlação em apenas uma parede rígida glotal.
- Calcular comprimentos dos ciclos das perturbações quando a frequência instantânea é perturbada por uma instância de microtremor artificial em apenas uma parede rígida glotal.
- Calcular comprimentos dos ciclos das perturbações quando a frequência instantânea é perturbada por um processo de ruído branco sem correlação considerando o acoplamento entre as paredes glotais esquerda e direita, não perturbadas de forma idêntica.
- Calcular comprimentos dos ciclos das perturbações quando a frequência instantânea é perturbada por um processo de ruído branco sem correlação considerando o movimento das margens superior e inferior em apenas uma parede glotal.

Estrutura da dissertação

A estrutura desta dissertação encontra-se dividida em oito capítulos:

O Capítulo I aborda os princípios de produção da voz humana, detalhando as funções da laringe e as oscilações das cordas vocais. A compreensão do aparelho fonador é importante para entender os parâmetros envolvidos na produção da voz.

O Capítulo II introduz o conceito de avaliação acústica e apresenta o programa de análise acústica *Praat* e a função *fxrapt*, um script específico desenvolvido para o programa *Matlab* para obter informações sobre os parâmetros da voz.

O Capítulo III apresenta um estudo bibliográfico, definindo a frequência fundamental f_0 , detalhando os fatores relacionados nas suas variações e suas estatísticas, que permitem distinguir entre várias categorias as vozes feminina, masculina e infantil, assim como as características que diferenciam uma voz patológica de uma voz normal.

O Capítulo IV apresenta dois algoritmos elaborados no software *Matlab* para determinação da frequência fundamental de vozes consideradas normais e com características de patologias, implementados pelo método Cepstral e pelo método da Autocorrelação, e os resultados obtidos são comparados com os encontrados pela função *fxrapt* e pelo programa *Praat*. Dentre os trabalhos estudados para esse capítulo, tem-se um artigo de Teixeira [69] de Análise acústica vocal para a determinação do jitter e shimmer para diagnóstico de patologias da fala, onde são descritas algumas patologias do aparelho fonador que provocam alterações na qualidade da fala produzida e algoritmos implementados para determinação da frequência fundamental.

O Capítulo V introduz os conceitos básicos necessários para compreensão do parâmetro jitter. Esse capítulo encontra-se segmentado em várias seções onde as estatísticas do jitter, os fatores relacionados às suas variações, os tipos de medidas do jitter e sua determinação são discutidas.

O Capítulo VI apresenta a implementação de um algoritmo no software *Matlab* para a determinação das medidas de jitter, baseando-se nas equações correspondentes a cada medida, apresentando os resultados que foram comparados com os valores do jitter determinados usando o software *Praat*.

O Capítulo VII apresenta modelos estocásticos de jitter. Na primeira seção, introduz-se o modelo geral de fita para as cordas vocais. Na segunda seção, apresentam-se os modelos dos distúrbios estocásticos das frequências instantâneas de vibração e, na última seção, combinamos os modelos de fita com os modelos de perturbação em quatro modelos de jitter. Finalmente, apresenta-se um modelo que combina as características dos quatro sub-modelos que foram introduzidos, com o modelo de microtremor vocal.

No capítulo VIII, são apresentados os resultados obtidos por meio das simulações dos algoritmos criados dos cinco modelos de jitter implementados no software *Matlab*. Também são discutidos e interpretados os resultados de cada simulação.

Finalmente, apresenta-se o capítulo das conclusões.

Capítulo 2

Princípios de produção da voz humana

A voz é produzida pelo fluxo de ar expiratório que advém dos pulmões e passa pelas cordas vocais, fazendo-as vibrar (Fig.2.1)[2].

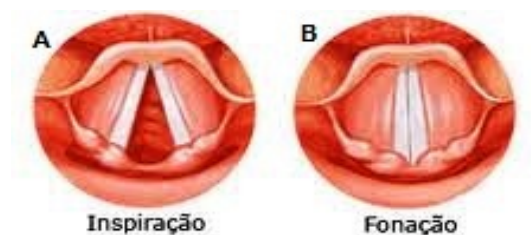


Figura 2.1: Posicionamento das cordas vocais durante a respiração (A) e durante a produção da voz (B)

<http://www.gastroweb.com.br/imgend/laringe.htm>

Os sons vocálicos são produzidos pela propagação dos pulsos de ar, quase periódicos, gerados pelo movimento cíclico das cordas vocais e que, posteriormente, são filtrados e amplificados pelo trato vocal. Os sons fricativos (/f/, /v/) são obtidos com a passagem turbulenta de ar através de alguma constrição formada no trato vocal e os sons plosivos (/d/, /t/, /p/, /b/) são produzidos pelo fechamento completo do trato vocal com consequente aumento da pressão anterior à obstrução e liberação abrupta desta. Pode-se dizer que a voz é uma forma inteligível de som, produto de uma excitação (gerador) e dos efeitos ressonantes do trato vocal (filtro). Do ponto de vista de análise do aparato vocal, este modo de geração de voz contém toda a informação fisiológica sobre o sistema[2].

2.1 Funções da Laringe

A laringe possui várias funções, das quais as mais importantes são: a respiratória, a deglutitória e a fonatória[2].

2.1.1 Função respiratória

Esta função ocupa a maior parte do tempo desse órgão, pois é a partir da abertura da laringe que são permitidas a entrada e a saída do ar dos pulmões. Durante o ato inspiratório, a laringe é ligeiramente rebaixada, passando a ser tracionada ligeiramente para cima durante a expiração. Esse deslocamento é proporcional à intensidade da respiração.

2.1.2 Função deglutitória

Primeiramente, a deglutição ocorre na boca e é voluntária, sendo seguida de uma fase involuntária (fase faríngea), onde ocorre o fechamento máximo da laringe.

2.1.3 Função fonatória

A fonação é uma habilidade adquirida pela espécie humana. Com a evolução, o ser humano aprendeu a coordenar órgãos que, fisiológica e anatomicamente, apresentavam como funções primordiais a respiração, a alimentação, e a proteção das vias aéreas inferiores. Deste modo, a voz, apesar de sua beleza e aprimoramento, é uma função adaptada e, como tal, depende da integridade e equilíbrio das estruturas do trato vocal.

Ao emitir a voz, as cordas vocais se aproximam da linha média, controlando a saída do ar que advém dos pulmões. O ar, por sua vez, ao passar pela laringe, coloca em oscilação as cordas vocais, que estão próximas entre si. As cordas vocais, neste instante, fecham-se e abrem-se em uma rápida sequência, realizando a função fonatória.

2.2 Oscilação das cordas vocais

As cordas vocais oscilam em um movimento quase periódico. Porém a oscilação pode ser alterada em decorrência de mudanças na massa e na tensão das próprias cordas vocais.

Outras alterações, como por exemplo, a paralisia dos músculos respiratórios, podem causar pressão subglótica insuficiente, alterando as forças aerodinâmicas que agem sobre

as cordas vocais o que, também, alterará sua oscilação[2].

A conceituação de vozes normais e disfônicas é negociável, visto que implica certa subjetividade. A dicotomia existente entre normalidade versus disfonia é um dos temas mais controversos na área de Fonoaudiologia. A palavra normal, que significa segundo a norma, quando aplicada ao comportamento humano, defronta-se a questões influenciáveis, como, por exemplo, gosto pessoal, modismo, fatores raciais, culturais e sexuais. Desta maneira, podem-se formular apenas padrões de vozes ditas normais.

De acordo com Tabith[67], normalidade vocal implica a dependência de vários fatores, como normalidade anatômica e funcional, boa coordenação entre respiração e fonação, boa saúde geral, ausência de problemas psicológicos relevantes e emissão agradável.

São sugeridos por Aronson[4], três questões para o julgamento de vozes normais:

1. se a voz é adequada para oferecer ao ouvinte inteligibilidade de fala;
2. se suas propriedades acústicas são aceitáveis;
3. se a voz preenche as demandas profissionais e sociais do falante. O autor acrescenta ao seu discurso que a variedade vocal é ilimitada e os padrões de adequação vocal são amplos.

Para Behlau[9], a voz deve ser produzida sem esforço, e com conforto, identificando o sexo e a faixa etária a que pertence. A autora propõe a utilização do termo voz adaptada, ao invés de voz normal, pois a produção vocal deve ser aceitável socialmente e apropriada às características sócio-histórico-culturais do falante.

Encontra-se, na literatura, uma variedade de classificações destinadas a conceituar as disfonias ([3],[12],[29],[58]).

De qualquer forma, os sintomas acústicos percebidos durante o processo patológico da laringe são mudanças na frequência fundamental da voz, sua intensidade, ou sua qualidade. Esses sintomas são indicativos de alterações orgânicas e/ou funcionais e a natureza desses sintomas varia para cada paciente e com o estágio da patologia.

Capítulo 3

Avaliação Acústica

3.1 Avaliação Acústica

Ternstrom [68] diz que o uso das medidas acústicas para a análise vocal requer um conhecimento aprofundado do sinal acústico e da forma como este é estruturado. Esta forma de análise permite determinar e quantificar, de forma não invasiva, a qualidade vocal (análise do sinal vocal, ao nível laríngeo e supralaríngeo) do indivíduo através dos parâmetros acústicos que compõem o sinal: periodicidade, amplitude, duração e composição do espectro.

De acordo com um estudo de Behlau [9], a avaliação acústica quantifica o sinal sonoro e permite algumas vantagens, levando em conta apenas a análise acústica para fins clínicos: permite uma maior compreensão do sinal vocal e associação entre as análises audio-perceptual e acústica; promove dados normativos para realidades vocais diferentes (culturais, profissionais ou patológicas); oferece uma documentação suficiente para uma linha de base da voz do indivíduo; monitoria a eficácia de um tratamento e compara resultados vocais de diferentes técnicas terapêuticas, nas diversas fases do tratamento; permite acompanhar o desenvolvimento de uma voz profissional ao longo de um período; permite fazer um diagnóstico precoce de problemas vocais e laríngeos.

Os parâmetros acústicos mais referidos na literatura são: frequência fundamental (F_0), perturbação da F_0 (Jitter), perturbação da amplitude (Shimmer) e medidas de ruído espectral (HNR, do inglês, Harmonic-to-Noise Ratio e NNE, do inglês, Normalized Noise Energy) [6], [9], [18], [32], [56].

Os Programas de análise acústica através de processamento de sinais e algoritmos são capazes de obter o traçado do formato da onda sonora, análise de frequência funda-

mental, medidas de perturbação como jitter e shimmer e medidas de ruído, permitindo descrever quase completamente a voz humana. Um dos programas usados para calcular os parâmetros acústicos é o Praat.

3.2 O programa Praat

O Praat é um *software* livre, desenvolvido por Paul Boersma e David Weenink, do Instituto de Ciências Fonéticas de Amsterdam, disponível em www.praat.org. É um programa compatível com diversos sistemas operacionais (Windows, Linux, Mac...), sobre o qual há um grupo de discussão, o que permite a identificação de erros e atualização constante, tanto que, praticamente, semanalmente é lançada uma nova versão do programa.

O Praat é bastante difundido e respeitado no meio científico como um programa de análise acústica da fala. Uma grande gama de trabalhos nas áreas de fonética e fonologia faz uso desse programa, que oferece uma série de funções, como: analisar, sintetizar, e manipular desde os segmentos até a melodia dos sons da fala e ainda é possível criar figuras de alta qualidade como espectrogramas, oscilogramas, curvas de pitch, intensidade e muito mais.

Em geral, o programa permite estudar o espectrograma, já que, por meio dessa ferramenta, obtemos os valores da frequência fundamental, dos formantes, e principais parâmetros acústicos, como jitter e shimmer.

3.3 Função fxrapt

Algoritmo robusto para o rastreamento do pitch, o RAPT (do inglês robust algorithm for pitch tracking) usa uma correlação cruzada normalizada para estimar a frequência fundamental, e programação dinâmica para remover descontinuidades no pitch. O nome da função é *fxrapt* (do inglês fundamental frequency estimation by Robust Algorithm for Pitch Tracking), basicamente, é um script específico desenvolvido para o programa *Matlab* para obter informações sobre os parâmetros da voz.

Capítulo 4

Frequência fundamental

A frequência fundamental (f_0) é definida como o número de oscilações por segundo produzidas pelas cordas vocais. De acordo com Baken e Orlikoff[6], a medição da frequência fundamental depende da periodicidade do sinal, ou seja, um sinal é periódico se é caracterizado por um conjunto de ondas que se repetem em ciclo. A f_0 , de uma dada emissão é estabelecida pelo número de ciclos que as cordas vocais conseguem produzir em um segundo, ou seja, pelo número de ciclos glóticos que se repetem [8].

A f_0 de um som vocálico é função da massa, da tensão das cordas vocais, da pressão subglótica e da configuração do trato vocal (carga acústica). De acordo com Nieto [54], a f_0 é diretamente proporcional à tensão das cordas vocais, ou à pressão subglótica, e inversamente proporcional à massa. Medida mais frequentemente em Hertz, pois é a unidade usada para descrever o parâmetro físico resultante da vibração das cordas vocais, por unidade de tempo, no comportamento vocal sustentado ou na fala encadeada [69].

A frequência fundamental reflete a eficiência do sistema fonatório, a biomecânica laríngea e a sua interação com a aerodinâmica. Qualquer ajuste que reduza os ciclos glóticos vai, conseqüentemente, reduzir a f_0 [8]. De acordo com Ternstrom[68], para sinais de fonação normal, esta medida é relativamente fácil de se obter.

4.0.1 Estatísticas da frequência fundamental

As estatísticas de f_0 mais usadas são a média, a moda e o desvio padrão [32]. Behlau [9] apresenta valores de f_0 , com uma distribuição média de 80 a 250 Hz, nos adultos jovens, sendo que nos homens as frequências situam-se entre os 80 a 150 Hz e de 150 a 250 Hz nas mulheres. As crianças apresentam valores acima de 250 Hz, atingindo os 400 Hz nos bebês. Num estudo do português falado no Brasil, os valores médios apresentados para

os homens, mulheres e crianças de 8 a 11 anos, são respectivamente 113 Hz, 205 Hz e 236 Hz [9].

A medição da f_0 pode ser efetuada de diversos modos e utilizando diferentes tipos de amostras de fala, como por exemplo, vogal sustentada, leitura, e conversação espontânea [18].

Levando em conta as diferentes amostras de fala para extração da f_0 , Guimarães [32] refere universalidade dos valores da f_0 em relação à produção sustentada das diferentes vogais. Assim, vogais “altas”, como o /i/ e o /u/, são produzidas com f_0 mais elevada, e vogais “baixas”, como o /a/, são produzidas com f_0 mais baixa. A explicação para esta universalidade tem por base o modelo teórico da articulação, no qual se explica que a f_0 varia com a postura da língua, devido às suas ligações à estrutura laríngea. Sorensen e Horii citados por Guimarães [32] verificaram esta relação da f_0 com o tipo de vogal, ou seja, o valor de f_0 para a vogal /a/ foi 198,8 Hz e para as vogais mais altas f_0 de 204,6 Hz para /u/ e 205,5 para /i/, para mulheres adultas. Para os homens adultos, os valores de f_0 foram 123,2 Hz para /u/, 125,6 para /i/ e 110,9 Hz para /a/.

Sussman e Sapienza (1994, citados por Guimarães [32]), também encontraram valores coincidentes com o modelo de articulação, ou seja, 231 Hz para a vogal /u/, 228 Hz para a vogal /i/, 215 Hz para a vogal /a/, para mulheres adultas; e 128 Hz para a vogal /u/, 122 para a vogal /i/ e 115 para a vogal /a/, para os homens adultos. Num estudo com 148 indivíduos saudáveis, de ambos os gêneros e com idades compreendidas entre os 20 e os 43 anos encontraram valores médios da f_0 para a vogal /a/, de 120 Hz para os homens e 200 Hz para as mulheres. Estes valores variaram entre os 83 e os 153 Hz para os homens e entre os 158 e os 274 Hz para as mulheres [33].

Num estudo com português do Brasil, envolvendo uma amostra de 20 homens e 20 mulheres saudáveis, com idades compreendidas entre os 20 e os 45 anos, verificou-se que a f_0 média para as mulheres situava-se nos 205,8 Hz para a vogal /a/ e 206,6 Hz para a vogal /e/. Para os homens, os valores médios da f_0 foram 119,9 Hz para a vogal /a/ e 118,9 Hz para a vogal /e/ [26].

Baken e Orlikoff [6] dizem não haver consenso no que diz respeito ao efeito das perturbações vocais na variação da f_0 constatam que habitualmente a f_0 aumenta com a presença de patologia vocal, em relação a falantes normais, tanto para os homens como para as mulheres.

4.0.2 Fatores relacionados nas variações da frequência fundamental

O gênero e a idade são os fatores referidos pela literatura como sendo os que alteram mais os valores da f_0 [9]. No entanto, existem outros fatores igualmente importantes a considerar, quando se estuda a f_0 , como os hábitos de vida (álcool e tabaco), o uso profissional da voz (falada ou cantada) e a presença de qualquer alteração vocal que perturbe a adequada vibração das cordas vocais, podendo originar assim disfonias [32]. Baken e Orlikoff [6] apresentam ainda a raça como um potencial fator de alteração, citando alguns estudos, como os de Hollien e Malcik (1962, 1967) que encontraram um valor ligeiramente inferior da f_0 em jovens africanos do sexo masculino, em relação aos seus pares europeus e americanos. Hudson e Holbrook (1981 e 1982, citados por Baken e Orlikoff [6]) também encontraram valores de f_0 menor em homens e mulheres afro-americanos que em europeus e americanos, no entanto, segundo os mesmos autores as diferenças não foram significativas.

Segundo Guimarães [32] é aceita universalmente que a f_0 difere de acordo com o sexo e se modifica em função da idade. É consensual que a f_0 das mulheres é mais elevada que a dos homens, mas os valores expostos apresentam discrepâncias. A investigação da qualidade vocal relacionada com a idade mostra, para ambos os sexos, uma diminuição do nível médio da f_0 com a idade adulta e novamente um aumento da f_0 com o início da velhice. Estas diferenças e as respectivas idades ainda não são consensuais.

Foram ainda encontrados valores baixos da f_0 associados ao número de cigarros por dia e aos hábitos tabágicos. As variáveis, número de cafés por dia, fatores psicológicos associados à disфонia e stresse foram associados a valores superiores da f_0 [8].

Num outro estudo de Guimarães e Abberton [31], com uma amostra de 20 fumantes (média de idades de 37 anos, com idades compreendidas entre os 27 e 51 anos) e 12 não fumadores (média de idades de 32 anos, com idades compreendidas entre os 20 e os 51 anos), foram encontrados valores médios da f_0 menores para os fumadores, com e sem problemas vocais em relação aos seus pares não fumadores, para todas as tarefas vocais. Dentro do grupo dos fumadores, os indivíduos com problemas vocais apresentavam valores menores da f_0 na tarefa das vogais sustentadas, que os fumadores sem problemas vocais, com exceção da vogal /u/, nas mulheres. No entanto, os resultados estatísticos mostram que as diferenças entre grupos não foram significativas para as tarefas vocais referidas anteriormente[8].

Em relação à influência dos hábitos alcoólicos na voz, Guimarães [32], refere que a investigação nesta área ainda é limitada. Klingholz et al.(1988, citados por Guimarães[32]) efetuaram um estudo com 15 homens não alcoólicos, onze desses homens consumiram álcool em diferentes quantidades (de forma a apresentar diferentes níveis de concentração de álcool no sangue), enquanto os restantes mantiveram-se sóbrios (grupo de controlo). Os resultados acústicos demonstraram um aumento da variabilidade da f_0 relacionado com os níveis de intoxicação alcoólica.

Niedzielska e Pruszeuics (1996, citados por Guimarães[32]) efetuaram um estudo com 30 adultos alcoólicos e dez adultos sóbrios, verificando um aumento significativo nas medidas acústicas da f_0 nos alcoólicos, em relação aos sóbrios.

4.1 Intensidade vocal

A intensidade vocal depende diretamente da resistência que a glote oferece à passagem de ar, ou seja, uma glote eficiente gera aumento da pressão subglótica, responsável pela variação da intensidade [9].

Fatores como velocidade da emissão do ar e quantidade de ar emitido também interferem na intensidade vocal, e são diretamente relacionados à pressão subglótica, ou seja, quanto maior for a pressão subglótica, maior será a velocidade e a quantidade do ar emitida.

A etiologia de vozes com intensidade fraca pode ser atribuída à coaptação glótica deficiente, gerando assim uma pressão subglótica insuficiente (por exemplo, em casos de paralisia dos músculos respiratórios, ou por paralisia de cordas vocais cuja boa coaptação está afetada).

A forma perceptiva da intensidade é denominada *loudness*.

4.2 Qualidade vocal

O termo qualidade vocal, antigamente conhecido como timbre, é o que caracteriza e identifica a impressão vocal. Uma degradação na qualidade vocal, geralmente definida como rouquidão, é frequentemente o primeiro, e algumas vezes o único sintoma de disfunção laringeal.

Vários são os termos utilizados para conceituar a voz, levando assim a divergências

no que concerne a nomenclatura. A imprecisão da terminologia é tão acentuada que se apropria de vocábulos inerentes a outros sentidos, como por exemplo, ao dizer que a voz é clara, utiliza-se assim do sentido da visão, ou quando se infere que a voz é rude, sendo esse um atributo do sentido tátil. No entanto, tais terminologias são corriqueiramente usadas.

Há, na literatura, registro de vinte e um tipos de vozes: voz rouca, soprosa, àspera, sussurrada, fluida, gutural, comprimida, tensa-estrangulada, bitonal, diplofônica, polifônica, monótona, trêmula, pastosa, crepitante, infantizada, virilizada, presbifônica, hipernasal, hiponasal e com nasalidade mista [10].

Capítulo 5

Métodos para determinação da frequência fundamental de sinais de voz

5.1 Determinação da frequência fundamental

Existem vários métodos que nos permitem a extração dos parâmetros do sinal de voz, incluindo a frequência fundamental, que se baseiam no modelo genérico de produção da fala (Fig.(5.1)) e na parametrização por formantes do trato vocal. Neste capítulo, dois métodos distintos para extração da frequência fundamental são discutidos: o Método Cepstral e o Método da Autocorrelação.

5.2 Algoritmos usados para o cálculo da F_0

5.2.1 Determinação de F_0 pelo Método Cepstral

A análise Cepstral de um sinal de voz permite-nos trabalhar com o sinal da glote (excitação) e do trato vocal (ressonância) separadamente, o que facilita o estudo das alterações nas cordas vocais.

O nome “cepstro” é obtido ao se inverter a ordem das primeiras quatro letras de

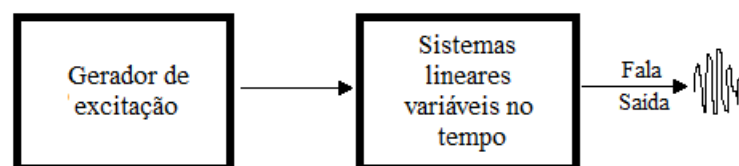


Figura 5.1: Modelo genérico de produção da fala.

“spectro”[69]. O Cepstro é uma representação do sinal de voz em que um sinal de fonte glótica, de variação temporal rápida, e a resposta do trato vocal, de variação lenta, são desacoplados e transformados em dois componentes aditivos. Este procedimento de análise Cepstral é ilustrado no diagrama da Fig.(5.2)

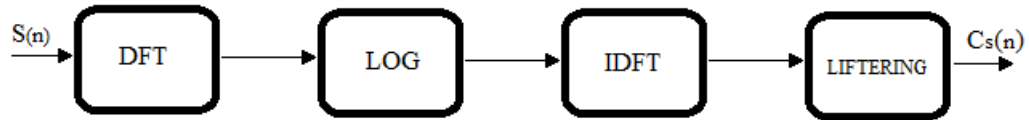


Figura 5.2: Diagrama de fluxo de estimação do Cepstro.

Onde $S(n)$ é o sinal de voz discretizado, DFT é a Transformada Discreta de Fourier (Discrete Fourier Transform), IDFT é a Transformada Inversa Discreta de Fourier (Inverse Discrete Fourier Transform), LOG é a função logarítmica e $C_s(n)$ são os coeficientes Cepstrais obtidos.

Este algoritmo resume-se na sequência de passos:

1ª etapa: O sinal de voz $s(t)$ é composto por um sinal de excitação $e(t)$ aplicado ao filtro do trato vocal, com uma resposta impulsional $v(t)$, de um ponto de vista no domínio temporal Eq.(5.1)

$$s(t) = e(t) * v(t) \quad (5.1)$$

Onde $s(t)$ representa a convolução de $e(t)$ com $v(t)$.

Aplicando-se a DFT na Eq.(5.1) temos,

$$S(w) = E(w).V(w) \quad (5.2)$$

Onde $E(w)$ e $V(w)$ representam as transformadas de $e(t)$ e $v(t)$, respectivamente.

2ª etapa: No espectro de magnitude da voz, duas componentes podem ser identificadas: uma parte lentamente variante (envoltória), devido ao sistema vocal, e uma parte rapidamente variante, devido à excitação. Estas componentes são combinadas por multiplicação Eq.(5.2) e a sua separação é possível usando a função logarítmica

Eq.(5.3):

$$\log[S(w)] = \log[E(w)] + \log[V(w)] \quad (5.3)$$

Portanto, a propriedade aditiva do espectro logarítmico representa a soma do espectro da fonte com o do filtro, o que nos fornece um caminho para os separar.

3ª etapa: Depois de separados, podemos calcular a IDFT (Inverse Discrete Fourier Transform) de sua densidade espectral produzindo a função Cepstro C_n , e o conjuntos de valores na saída desse processo são os chamados coeficientes cepstrais.

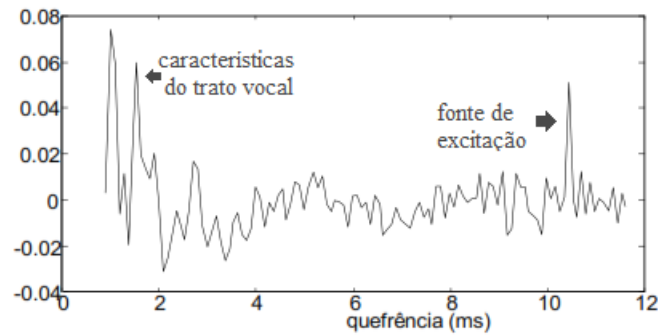


Figura 5.3: Cepstro de um segmento de fala [69].

O eixo horizontal do gráfico de cepstro (Fig.(5.3)) tem dimensões temporais e são chamados de **quefrências**. Pode, também, ser feita uma clara distinção entre a componente de excitação e a contribuição do trato vocal.

A Fig.(5.3), ilustra o Cepstro de um segmento de fala, onde o pico mais alto correspondente ao período fundamental (excitação) está próximo da quefrência de 10 ms, separado das componentes do trato vocal às baixas quefrências.

4ª etapa: Como a função de transferência do trato vocal e a função de excitação da voz aparecem em partes separadas da escala de quefrências, estas podem ser separadas ou removidas por um processo de “lifteragem”(liftering), facilitando o estudo individualizado das modificações ocorridas na excitação e na parte ressonatal.

5.2.2 Determinação da F_0 pelo método da Autocorrelação

Para um sinal $x(t)$, a autocorrelação, $R_x(\tau)$, em função do atraso τ , é definida pela Eq.(5.4):

$$R_x(\tau) = \int_{-\infty}^{+\infty} x(t)x(t+\tau)dt. \quad (5.4)$$

A F_0 de um sinal de voz $x(t)$ contínuo no tempo pode ser determinada a partir dos valores máximos locais da autocorrelação de um segmento do sinal. O algoritmo de determinação da F_0 pelo método de Autocorrelação será resumidamente explicado a seguir [69].

1ª etapa: Encontra-se uma janela representativa do sinal $x(t)$ com uma duração T , centrada em torno de um valor, digamos t_{mid} . Esta janela deve conter pelo menos dois períodos fundamentais do sinal. Subtrai-se desta janela a sua média, μ_x , e multiplica-se o resultado por uma função de janela $w(t)$ (por exemplo uma janela de Hanning, Hamming, Blackman, Kaiser, etc.), para obter o sinal, como na Eq.(5.5):

$$a(t) = (x(t_{mid} - \frac{1}{2}T + \mu_x))w(t). \quad (5.5)$$

A janela $w(t)$ é simétrica em relação a $t = (1/2)T$ e zero fora do intervalo de tempo $[0, T]$. Escolheu-se para tal a janela de Hanning, dada pela Eq.(5.6):

$$w(t) = \frac{1}{2} - \frac{1}{2} \cos \frac{2\pi t}{T}. \quad (5.6)$$

Esta janela é criada para diminuir os efeitos de "Leakage".

2ª etapa: A autocorrelação normalizada, $R_a(\tau)$, do sinal $a(t)$, é uma função par dada por Eq.(5.7):

$$R_a(\tau) = R_a(-\tau) = \frac{\int_0^{T-\tau} a(t)a(t+\tau)dt}{\int_0^T a^2(t)dt}. \quad (5.7)$$

3ª etapa: Em seguida, calcula-se a autocorrelação normalizada da janela de Hanning

$R_w(\tau)$, de modo análogo à equação correspondente à $R_a(\tau)$. A autocorrelação normalizada da janela de Hanning é dada pela Eq.(5.8):

$$R_w(\tau) = \left(1 - \frac{|\tau|}{T}\right) \left(\frac{2}{3} + \frac{1}{3} \cos \frac{2\pi\tau}{T}\right) + \frac{1}{2\pi} \sin \frac{2\pi|\tau|}{T}. \quad (5.8)$$

O resultado final da autocorrelação, $R_x(\tau)$, do sinal original determina-se dividindo o valor da autocorrelação normalizada pelo valor da autocorrelação normalizada da janela Eq.(5.9):

$$R_x(\tau) = \frac{R_a(\tau)}{R_w(\tau)}. \quad (5.9)$$

A autocorrelação foi determinada numericamente pela FFT (Fast Fourier Transform). Isto é possível graças ao fato de a autocorrelação de um sinal poder ser obtida calculando a Transformada de Fourier do sinal, como na Eq.(5.10):

$$\tilde{a}(\omega) = \int_{-\infty}^{+\infty} a(t) e^{-i\omega t} dt. \quad (5.10)$$

Em seguida, calcula-se a IDFT da densidade de potência $|\tilde{a}(\omega)|^2$, que nos leva ao domínio de atraso (τ), como na Eq.(5.11):

$$R_a(\tau) = \int_{-\infty}^{+\infty} |\tilde{a}(\omega)|^2 e^{-i\omega\tau} \frac{d\omega}{2\pi}. \quad (5.11)$$

No sinal resultante, $R_a(\tau)$, pode-se determinar o período fundamental pela medição do tempo entre dois pontos sucessivos da curva com o valor máximo. O seu inverso é a F_0 :

$$F_0 = \frac{1}{T_0}. \quad (5.12)$$

5.3 Comparação entre os métodos: aplicação em um banco de vozes

A partir de um banco de vozes construído num consultório de fonoaudiologia, os métodos para o cálculo de F_0 foram aplicados e comparados.

O banco de vozes composto por vinte e nove sujeitos, com emissões da vogal sustentada /e/, consiste de três grupos: dez indivíduos com voz considerada normal, sete indivíduos que apresentam paralisia nas cordas vocais e doze indivíduos com presença de nódulos nas cordas vocais. Na Tab.(5.1) apresenta-se a faixa etária dos locutores usados e o sexo correspondente a cada um dos três grupos.

Tipo de voz	Homens (24-66 anos)	Mulheres (24-69 anos)
Normal	6	4
Nódulos	0	12
Paralisia	5	2

Tabela 5.1: Grupo de locutores por tipo da voz (normal e com características patológicas) e do sexo.

Para a análise das amostras, foram utilizados aproximadamente de 3 a 5 segundos de emissão, sendo descartados o início, e, em algumas amostras, o final da emissão das vogais. As estatísticas analisadas da F_0 foram a média, a mediana, o desvio de padrão e o máximo e o mínimo todas calculadas dependendo do sexo, da idade, do tipo de voz (normal e com características patológicas) e da vogal /e/.

No algoritmo utilizado para a determinação da F_0 pelo Método do Cepstro implementou-se uma função que eliminava os “outliers”(valores extremos) que se encontravam muito afastados dos demais valores, afetando consideravelmente a média e o desvio padrão, eles foram desconsiderados devido que no início e no fim da gravação não se tem informação do sinal da vogal sustentada. O “outlier” é definido como um valor mais do que três desvios padrão de distância da média.

Para comparação dos resultados dos algoritmos implementados usamos um script específico desenvolvido para o programa *Matlab* para a extração do pitch, a função *fxrapt* e o software *Praat* (ambos implementam o Método da Autocorrelação para a extração da F_0 , porém com algoritmos próprios).

Cada uma das emissões sustentadas da vogal /e/ foi analisada individualmente em cada um dos programas. Ao usar a função *fxrapt*, o sinal tinha o mesmo tamanho dos sinais usados nos algoritmos anteriores. No *Praat* selecionava-se a amostra para ter aproximadamente o mesmo tamanho analisado pelos outros programas.

Apresenta-se, a seguir, os resultados nas Tab.(5.2), Tab.(5.3) e Tab.(5.5), correspondentes aos valores de F_0 relativos aos sinais de voz (normal, com presença de nódulos e

*Os sinais (1,3,5,10) correspondem a vozes femininas e os sinais (2,4,6,7,8,9) pertencem a vozes masculinas.

Voz com presença de Nódulo											
S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)	S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
1	$\tilde{x}$	190.09	190.09	198.42	198.4	2	$\tilde{x}$	159.78	159.78	163.04	162.77
	$\bar{x}$	191.12	191.58	198.31	198.32		$\bar{x}$	161.18	161.44	164.15	164.06
	σ	1.91	1.76	1.41	1.4		σ	4.63	4.55	4.58	4.61
	Min	186.86	186.86	194.64	194.5		Min	153.13	155.28	158.47	158.05
	Max	196.88	193.42	201.56	201.57		Max	177.82	172.27	175.68	178.63
3	$\tilde{x}$	208.02	208.02	217.77	217.79	4	$\tilde{x}$	157.5	157.5	162.78	162.79
	$\bar{x}$	209.71	209.87	217.68	217.8		$\bar{x}$	158.56	158.66	162.85	162.9
	σ	2.9	2.42	2.33	2.43		σ	1.38	1.34	1.22	1.28
	Min	204.17	204.17	212.61	212		Min	155.28	157.5	159.97	159.77
	Max	220.5	216.18	225.29	225.45		Max	162.13	162.13	167.17	166.96
5	$\tilde{x}$	208.02	208.02	214.41	214.46	6	$\tilde{x}$	167.05	167.05	172.87	172.88
	$\bar{x}$	206.49	207.31	214.39	214.42		$\bar{x}$	167.81	167.72	172.74	172.78
	σ	1.9	1.5	1.02	1.06		σ	1.5	1.38	1.43	1.41
	Min	204.17	204.17	211.4	211.41		Min	164.55	164.55	167.95	168.91
	Max	208.02	208.02	217.38	217.35		Max	169.62	169.62	175.62	176.05
7	$\tilde{x}$	157.5	157.5	162.22	162.06	8	$\tilde{x}$	200.45	200.45	209.84	209.96
	$\bar{x}$	157.02	157.28	161.99	161.81		$\bar{x}$	200.86	202.5	209.57	209.43
	σ	1.58	1.66	1.47	1.61		σ	3.89	3.36	4.94	4.35
	Min	153.13	153.13	157.84	157		Min	193.42	196.88	194.49	200.07
	Max	159.78	159.78	165.41	165.66		Max	208.02	208.02	217.93	218.44
9	$\tilde{x}$	186.86	186.86	193.48	193.57	10	$\tilde{x}$	193.42	193.42	200.4	200.29
	$\bar{x}$	187.14	187.3	193.52	193.64		$\bar{x}$	192.99	194.09	200.56	200.43
	σ	1.63	1.77	1.4	1.55		σ	2.38	1.37	0.8	0.91
	Min	183.75	183.75	189.96	189.97		Min	186.86	193.42	199.01	198.55
	Max	190.09	190.09	196.93	197.63		Max	196.88	196.88	202.72	203.13
11	$\tilde{x}$	175	175	180.4	180.28	12	$\tilde{x}$	200.45	200.45	209.94	209.59
	$\bar{x}$	175.52	175.68	180.12	180.1		$\bar{x}$	201.6	201.43	210.18	209.76
	σ	1.6	1.53	1.77	1.84		σ	2.02	2.12	1.62	1.95
	Min	172.27	172.27	174.68	175.4		Min	196.88	196.88	206.81	204.56
	Max	177.82	177.82	184.72	184.95		Max	204.17	204.17	214.11	214.18

Tabela 5.3: Resultados obtidos pelo *Praat*, pelo *fxrapt* e pelo algoritmo criado com os diferentes métodos para os sinais da vogal sustentada /e/ de voz com presença de nódulo nas cordas vocais.

*Todos os sinais pertencem a vozes femininas

Voz com presença de Paralisia											
S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)	S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
1	$\tilde{x}$	151.03	153.13	156.55	156.7	2	$\tilde{x}$	141.35	141.35	142.05	142.14
	$\bar{x}$	152.74	152.48	156.62	156.65		$\bar{x}$	140.87	141.3	142.63	142.63
	σ	4.19	1.93	1.79	1.9		σ	0.96	1.08	1.68	1.73
	Min	147	148.99	151.81	151.68		Min	139.56	139.56	139.08	138.95
	Max	164.55	157.5	161.23	161.79		Max	143.18	143.18	146.02	146.63
3	$\tilde{x}$	141.35	141.35	138.85	135.18	4	$\tilde{x}$	204.17	204.17	212.67	212.09
	$\bar{x}$	141.6	142.11	138.31	136.45		$\bar{x}$	206.4	206.19	213.94	213.99
	σ	2.0071	1.4633	4.942	3.7542		σ	5	5.02	7.84	6.39
	Min	139.56	139.56	128.35	130.43		Min	200.45	200.45	191.85	198.37
	Max	147	145.07	146.15	143.45		Max	212.02	212.02	248.42	232.8
5	$\tilde{x}$	220.5	220.5	228.33	228.44	6	$\tilde{x}$	77.64	158.64	81.1	81.58
	$\bar{x}$	218.84	218.91	228.3	228.33		$\bar{x}$	80.91	156.99	81.51	81.99
	σ	2.42	2.63	2.66	2.67		σ	6.65	8.08	3.55	4.51
	Min	216.18	212.02	222.04	221.4		Min	70.67	143.18	74.82	71.81
	Max	225	225	234.59	235.02		Max	98.44	167.05	92.11	112.5

7	$\tilde{x}$	180.74	183.75	188.54	187.9
	$\bar{x}$	181.34	182.83	184.9	185.63
	σ	19.75	14.58	16.47	15.38
	Min	134.45	159.78	127.8	135.83
	Max	275.63	250.57	210.46	211.79

Tabela 5.5: Resultados obtidos pelo *Praat*, pelo *fxrapt* e pelo algoritmo criado com os diferentes métodos para os sinais da vogal sustentada /e/ de voz com presença de paralisia nas cordas vocais.

*Os sinais (1,2,3,6,7) correspondem a vozes masculinas e os sinais (4,5) a vozes femininas.

5.4 Considerações

Para efeitos de comparações calculou-se a média das médias, tal como feito no estudo [24].

S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
2	$\bar{x}$	120.43	120.58	123.14	123.14
4	$\bar{x}$	99.61	99.41	101.47	101.40
6	$\bar{x}$	101.04	101.16	103.25	102.85
7	$\bar{x}$	157.94	158.01	162.56	162.55
8	$\bar{x}$	130.45	130.68	133.61	133.62
9	$\bar{x}$	121.45	121.57	124.18	124.11
		121.82	121.90	124.70	124.61

Tabela 5.6: As médias e a média das médias da frequência fundamental em sinais de voz masculinas normais.

Observamos na Tab.(5.6) que a média das médias da frequência fundamental encontrada, para a vogal /e/, em homens com voz considerada normal para o método de Cepstro (121,82 Hz) foi muito semelhante à obtida para o método da Autocorrelação (121,90). Em comparação com os atingidos pelo *Praat* (124,70 Hz) e pelo *fxrapt* (124,61 Hz), foram obtidos erros percentuais para o método de Cepstro de (2.31%) e (2.25%) e para o método de Autocorrelação (2.24%) e (2.17%) respectivamente.

S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
1	$\bar{x}$	182.84	183.10	188.62	188.63
3	$\bar{x}$	166.18	166.28	171.14	171.14
5	$\bar{x}$	184.26	184.16	190.60	190.53
10	$\bar{x}$	192.52	192.46	199.17	199.21
		181.45	181.50	187.38	187.38

Tabela 5.7: As médias e a média das médias da frequência fundamental em sinais de voz femininas normais .

Como pode-se observar na Tab.(5.7), as médias da média obtidas do mesmo parâmetro para as mulheres para o método de Cepstro (181,45 Hz) e para o método de Autocorrelação

(181,50 Hz) foram muitos semelhantes. Em comparação com os atingidos pelo *Praat* (124,70 Hz) e pelo *fxrapt* (124,61 Hz), foram obtidos erros percentuais para o método de Cepstro de (3.17%) e (3.14%) e para o método de Autocorrelação (3.16%) e (3.14%), respectivamente.

S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
1	$\bar{x}$	152.74	152.48	156.62	156.65
2	$\bar{x}$	140.87	141.30	142.63	142.63
3	$\bar{x}$	141.60	142.11	138.31	136.45
6	$\bar{x}$	80.91	156.99	81.51	81.99
7	$\bar{x}$	181.34	182.83	184.90	185.63
		139.49	155.14	140.80	140.67

Tabela 5.8: As médias e a média das médias da frequência fundamental em sinais de voz masculinas com presença de paralisia.

Para voz com presença de paralisia nas cordas vocais da vogal /e/ (Tab.(5.8)), a média das médias da F_0 em homens para o método de Cepstro (139,49 Hz) foi inferior à encontrada no método da Autocorrelação (155,14 Hz). Neste valor viu-se afetado pela média do sinal 6 (corresponde a um individuo de 66 anos de idade e apresenta paralisia pearrencial à esquerda com moderado arqueamento da borda livre, trauma de entubação), que apresentou discrepância em relação aos obtidos com os outros métodos. Talvez devido ao ruído introduzido pela patologia que pode dificultar e até mesmo impedir o cálculo da F_0 . Assim, pode haver grande dificuldade na detecção dos segmentos sonoros do sinal de voz em análise.

Em comparação com os atingidos pelo *Praat* (140,80 Hz) e pelo *fxrapt* (140,67 Hz), foram obtidos erros percentuais para o método de Cepstro de (0.93%) e (0.85%) e para o método de Autocorrelação (-10.19%) e (-10.29%), respectivamente.

S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
4	$\bar{x}$	206.40	206.19	213.94	213.99
5	$\bar{x}$	218.84	218.91	228.30	228.33
		212.62	212.55	221.12	221.16

Tabela 5.9: As médias e a média das médias da frequência fundamental em sinais de voz femininas com presença de paralisia.

Observamos na Tab.(5.9) que a média das médias da F_0 encontrada em mulheres com o método de Cepstro (212,62 Hz) foi muito semelhante com o método da Autocorrelação

(212,55 Hz). Em comparação com os atingidos pelo *Praat* (221.12 Hz) e pelo *fxrapt* (221.16 Hz), foram obtidos erros percentuais para o método de Cepstro de (3.84%) e (3.86%) e para o método de Autocorrelação (3.88%) e (3.89%), respectivamente.

S*	M	C(Hz)	A(Hz)	Pt(Hz)	fx(Hz)
1	$\bar{x}$	191.12	191.58	198.31	198.32
2	$\bar{x}$	161.18	161.44	164.15	164.06
3	$\bar{x}$	209.71	209.87	217.68	217.80
4	$\bar{x}$	158.56	158.66	162.85	162.90
5	$\bar{x}$	206.49	207.31	214.39	214.42
6	$\bar{x}$	167.81	167.72	172.74	172.78
7	$\bar{x}$	157.02	157.28	161.99	161.81
8	$\bar{x}$	200.86	202.50	209.57	209.43
9	$\bar{x}$	187.14	187.30	193.52	193.64
10	$\bar{x}$	192.99	194.09	200.56	200.43
11	$\bar{x}$	175.52	175.68	180.12	180.10
12	$\bar{x}$	201.60	201.43	210.18	209.76
		184.17	184.57	190.51	190.45

Tabela 5.10: As médias e a média das médias da frequência fundamental em sinais de voz femininas com presença de nódulos.

Como pode-se observar na Tab.(5.10), a média das médias da F_0 , em mulheres com presença de nódulos nas cordas vocais para o método de Cepstro (184,17 Hz) foi muito semelhante à obtida com o método da Autocorrelação (184,57 Hz). Em comparação com os atingidos pelo *Praat* (190,51 Hz) e pelo *fxrapt* (190,45 Hz), foram obtidos erros percentuais para o método de Cepstro de (3.33%) e (3.30%) e para o método de Autocorrelação (3.11%) e (3.09%), respectivamente.

Baseando-se nos resultados obtidos, as médias das medias para cada uns dos grupos, conseguiu-se demonstrar que habitualmente a F_0 aumenta com a presença de patologia vocal, em relação a locutores normais, tanto para os homens como para as mulheres, coincidindo com as revisões de outros estudos, por exemplo, Baken e Orlikoff[6]. Similarmente validou-se que nas mulheres os valores da média de F_0 para a vogal sustentada /e/ são maiores do que nos homens.

A diferença significativa nos valores médios da F_0 , em função do sexo, encontrada no presente estudo, era esperada, pois é influenciada pelo comprimento das cordas vocais, que é maior no sexo masculino. Esta diferença já foi exaustivamente apontada na literatura [24].

Basicamente, existem dois tipos de abordagens para o cálculo de frequência fundamental. Uma categoria é no domínio do tempo, e a outra categoria é no domínio da frequência. No domínio do tempo implementou-se o método da Autocorrelação, e no domínio da frequência o Método Cepstral. Estes dois tipos de métodos têm suas vantagens e desvantagens, respectivamente; por exemplo, o método no domínio da frequência geralmente teve maior precisão que o método no domínio do tempo, mas ele precisou de mais cálculos.

Capítulo 6

Jitter

6.1 Jitter

O Jitter é uma medida de curto termo (medida de ciclo a ciclo), que mede a variabilidade na frequência fundamental não voluntária, permitindo determinar o grau de estabilidade do sistema fonatório [8]. O Jitter (perturbação de frequência) é um dos índices usados para indicar patologias vocais e pode ser de fácil medição.

O Jitter indica a variação da frequência fundamental a curto prazo, medida entre ciclos glóticos contíguos. Expressa o quanto um período é diferente do anterior e do seu sucessor imediato [9]. A irregularidade da vibração das cordas vocais pode ser manifestada por diferenças no tempo de vibração ou na amplitude de vibração [18].

Os ciclos vibratórios dos falantes normais apresentam oscilações. No entanto, pode constatar-se que uma laringe com patologia produz ciclos vibratórios (fonação) com maior alteração que uma sem patologia. É aceitável que uma voz apresente características instáveis perante fatores de ordem neurológica, emocional e biomecânica. Sendo previsível a presença de um pequeno grau de perturbação e irregularidade no sinal vocal[32].

Os valores de Jitter, em indivíduos “normais”, podem representar uma pequena variação na massa ou na tensão das cordas vocais, seja pela presença de muco ou por alterações momentâneas na atividade neuronal ou muscular[6].

6.1.1 Medidas de parâmetros de perturbação

Os parâmetros de perturbação representam medidas para quantificar as variações dos valores dos parâmetros ao longo do tempo, como, por exemplo, o Jitter. Há outros parâme-

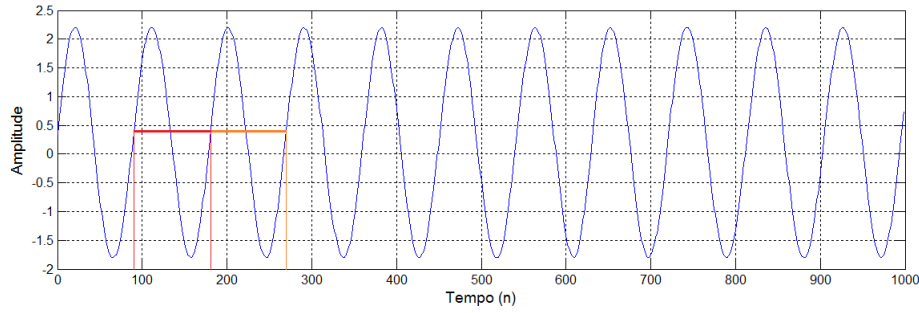


Figura 6.1: Variação da frequência fundamental medida entre ciclos glóticos contíguos (Jitter), Sinal da largura da glote correspondente ao sexo masculino.

tros como Shimmer e HNR que não serão tratados aqui. De forma geral, pode-se definir a perturbação de um parâmetro como:

$$V_p = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |P(i+1) - P(i)|}{\frac{1}{N} \sum_{i=1}^{N-1} P(i)}$$

Onde $P(i)$ é o valor do parâmetro para o período i , N é o número de períodos e V_p é o valor da perturbação do parâmetro.

6.1.2 Medidas de Jitter

A extração dos valores do Jitter pode ser feita através de medidas absolutas ou relativas. O Jitter absoluto expressa a diferença média absoluta entre os períodos sequenciais medidos através de uma fonação sustentada. É expresso em segundos, milissegundos ou micro-segundos. As medidas relativas do Jitter correlacionam-se com a média da F_0 . Habitualmente, as frequências altas apresentam menor perturbação, ou seja, maiores diferenças entre períodos ciclo a ciclo estão associadas com períodos de F_0 mais elevados[8].

O valor “normal” para o Jitter situa-se entre os 0,5% e os 1%. Os fatores que podem influenciar os valores do Jitter podem ser o nível da F_0 , o gênero, a idade e a presença de patologia vocal[8]. Também, existe um limite de normalidade em cada língua. Assim mesmo, parece haver uma relação da F_0 e do Jitter com o fumo, sendo que a F_0 é mais baixa nos fumantes e o Jitter mais altos nessa população[19].

6.1.3 Fatores relacionados nas variações do Jitter

Alguns autores consideram que as mulheres adultas (pelo menos em algumas das vogais) apresentam valores de Jitter superiores aos homens, no entanto, outros referem o inverso [8].

Um maior valor de Jitter nos homens indica menos controle sobre a voz e alguns autores referem que isto pode estar relacionado ao fato dos homens fazerem menos uso vocal do que as mulheres, evidenciando, então, menos treino vocal e consequente maior instabilidade. O Jitter aparece correlacionado à presença de ruído na voz ou à rouquidão, em determinados estudos. Portanto, quando os achados de valores de Jitter aparecem maiores nos homens do que nas mulheres, os valores de medidas de ruído, também tendem a ser maiores no sexo masculino[19].

6.1.4 Determinação do Jitter

Existem vários tipos de medidas do Jitter, discriminadas a seguir:

Jitter (local) : Representa a diferença média absoluta entre dois períodos consecutivos, dividido pelo período médio, Eq.(6.1). Chama-se a esse parâmetro **jitt**, e o valor limite para detectar patologias é de 1.040%.

$$jitt = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T(i) - T(i-1)|}{\frac{1}{N} \sum_{i=1}^{N-1} T(i)} * 100 \quad (6.1)$$

Onde T_i é o valor do parâmetro para o período i , N é o número de períodos e $jitt$ é o valor da perturbação de Jitter.

Jitter (local, absoluto) : Representa a diferença média absoluta entre dois períodos consecutivos e chama-se a este parâmetro **jitta**, Eq. (6.2). Este tem como umbral para patologias o valor de 83.200 μs .

$$jitta = \frac{1}{N-1} \sum_{i=1}^{N-1} |T(i) - T(i-1)| \quad (6.2)$$

Onde T_i é o valor do parâmetro para o período i , N é o número de períodos e $jitta$ é o valor da perturbação de Jitter.

Jitter (rap) (relative average perturbation): Representa a média relativa da perturbação, isto é, a diferença média absoluta entre um período e a média desse mesmo

período com os seus dois vizinhos, dividido pelo período médio, Eq.(6.3). Este tem como limite para detectar patologias o valor de 0.680%.

$$rap = \frac{\frac{1}{N-1} \sum_{i=2}^{N-1} |T(i) - (\frac{1}{3} \sum_{n=i-1}^{i+1} T_n)|}{\frac{1}{N} \sum_{i=1}^{N-1} T(i)} * 100 \quad (6.3)$$

T_i é o valor do parâmetro para o período i , N é o número de períodos e rap é o valor da perturbação de Jitter.

Jitter (ppq5) (Five Point Period Perturbation Quotient): Representa o quociente de perturbação num período de cinco pontos, ou seja, a diferença média absoluta entre um período e a média desse mesmo período com os seus quatro vizinhos, dividido pelo período médio, Eq.(6.4). Este tem como umbral para patologias o valor de 0,840%.

$$ppq5 = \frac{\frac{1}{N-1} \sum_{i=3}^{N-2} |T(i) - (\frac{1}{5} \sum_{n=i-2}^{i+2} T_n)|}{\frac{1}{N} \sum_{i=1}^{N-1} T(i)} * 100 \quad (6.4)$$

Onde T_i é o valor do parâmetro para o período i , N é o número de períodos e $ppq5$ é o valor da perturbação de Jitter.

Capítulo 7

Comparação dos valores de Jitter usando a função *fxrapt* e o software Praat

7.1 Determinação do Jitter

Para a determinação das medidas de Jitter, implementou-se um algoritmo no software *Matlab* considerando os diferentes tipos de medidas *jitt*, *jitta*, *rap*, *ppq5*, baseando-se nas equações correspondentes a cada medida (Eq.(6.1), Eq.(6.2), Eq.(6.3), e Eq.(6.4)). A rotina pode ser vista no Apêndice C.

No fato de o Jitter refletir a variação nos “períodos consecutivos” do sinal de voz ou, equivalentemente, no sinal glotal, calcularam-se os comprimentos de cada período do sinal, pela função *fxrapt*, que extrai a frequência fundamental. Cada intervalo de tempo considerado é o inverso da frequência fundamental correspondente. O banco de vozes utilizado foi o mesmo usado para o cálculo da F_0 . Os resultados do algoritmo implementado foram comparados pelo obtido pelo programa *Praat*.

Cada uma das emissões sustentadas da vogal /e/ foi analisada individualmente em cada um dos programas. No *Praat*, selecionava-se a amostra para ter aproximadamente o mesmo tamanho analisado no algoritmo implementado em *Matlab*.

As Tab.(7.1), Tab.(7.2) e Tab.(7.4) apresentam os valores de Jitter obtidos (considerada normal, com presença de nódulos e com paralisia nas cordas vocais respectivamente).

Voz Normal							
Sinal	Jitter	Algoritmo-Jitter	Praat	Sinal	Jitter	Algoritmo-Jitter	Praat
1	Jitt (%)	0.31	0.293	2	Jitt (%)	0.284	0.316
	Jitta (s)	1.64E-05	1.55E-05		Jitta (s)	2.31E-05	2.57E-05
	rap (%)	0.135	0.155		rap (%)	0.154	0.181
	ppq5 (%)	0.166	0.166		ppq5 (%)	0.171	0.185
3	Jitt (%)	0.236	0.238	4	Jitt (%)	0.457	0.444
	Jitta (s)	1.38E-05	1.39E-05		Jitta (s)	4.51E-05	4.38E-05
	rap (%)	0.122	0.131		rap (%)	0.233	0.23
	ppq5 (%)	0.131	0.151		ppq5 (%)	0.28	0.279
5	Jitt (%)	0.309	0.357	6	Jitt (%)	0.26	0.269
	Jitta (s)	1.62E-05	1.87E-05		Jitta (s)	2.53E-05	2.61E-05
	rap (%)	0.149	0.212		rap (%)	0.111	0.121
	ppq5 (%)	0.18	0.202		ppq5 (%)	0.146	0.154
7	Jitt (%)	0.165	0.125	8	Jitt (%)	0.209	0.169
	Jitta (s)	1.01E-05	7.68E-06		Jitta (s)	1.56E-05	1.26E-05
	rap (%)	0.061	0.053		rap (%)	0.104	0.082
	ppq5 (%)	0.098	0.07		ppq5 (%)	0.12	0.1
9	Jitt (%)	0.256	0.251	10	Jitt (%)	0.194	0.274
	Jitta (s)	2.06E-05	2.02E-05		Jitta (s)	9.75E-06	1.38E-05
	rap (%)	0.142	0.141		rap (%)	0.095	0.164
	ppq5 (%)	0.15	0.151		ppq5 (%)	0.118	0.155

Tabela 7.1: Resultados obtidos pelo algoritmo criado e pelo *Praat* para o Jitter dos sinais da vogal sustentada /e/ de voz normal.

*Os sinais (1,3,5,10) correspondem a vozes femininas e os sinais (2,4,6,7,8,9) pertencem a vozes masculinas.

Voz com presença de Nódulo							
Sinal	Jitter	Algoritmo-Jitter	Praat	Sinal	Jitter	Algoritmo-Jitter	Praat
1	Jitt (%)	0.182	0.173	2	Jitt (%)	0.59	0.883
	Jitta (s)	9.17E-06	8.73E-06		Jitta (s)	3.60E-05	5.39E-05
	rap (%)	0.075	0.093		rap (%)	0.345	0.538
	ppq5 (%)	0.097	0.098		ppq5 (%)	0.345	0.532
3	Jitt (%)	0.454	0.559	4	Jitt (%)	0.306	0.37
	Jitta (s)	2.08E-05	2.57E-05		Jitta (s)	1.88E-05	2.27E-05
	rap (%)	0.259	0.325		rap (%)	0.171	0.221
	ppq5 (%)	0.282	0.352		ppq5 (%)	0.175	0.22
5	Jitt (%)	0.235	0.349	6	Jitt (%)	0.275	0.3
	Jitta (s)	1.10E-05	1.63E-05		Jitta (s)	1.59E-05	1.74E-05
	rap (%)	0.11	0.206		rap (%)	0.149	0.169
	ppq5 (%)	0.144	0.202		ppq5 (%)	0.166	0.184
7	Jitt (%)	0.135	0.11	8	Jitt (%)	0.17	0.156
	Jitta (s)	8.338E-06	6.79E-06		Jitta (s)	1.27E-05	1.17E-05
	rap (%)	0.05	0.048		rap (%)	0.081	0.077
	ppq5 (%)	0.083	0.064		ppq5 (%)	0.097	0.09
9	Jitt (%)	0.468	0.47	10	Jitt (%)	0.233	0.341
	Jitta (s)	2.42E-05	2.43E-05		Jitta (s)	1.16E-05	1.70E-05
	rap (%)	0.257	0.249		rap (%)	0.126	0.2
	ppq5 (%)	0.301	0.325		ppq5 (%)	0.149	0.215
11	Jitt (%)	0.417	0.619	12	Jitt (%)	0.209	0.227
	Jitta (s)	2.31E-05	3.44E-05		Jitta (s)	9.99E-06	1.08E-05
	rap (%)	0.245	0.368		rap (%)	0.11	0.127
	ppq5 (%)	0.254	0.373		ppq5 (%)	0.132	0.139

Tabela 7.2: Resultados obtidos pelo algoritmo criado e pelo *Praat* para o Jitter dos sinais da vogal sustentada /e/ de voz com presença de nódulos.

* Todos os sinais pertencem a vozes femininas

Voz com presença de Paralisia							
Sinal	Jitter	Algoritmo-Jitter	Praat	Sinal	Jitter	Algoritmo-Jitter	Praat
1	Jitt (%)	0.498	0.487	2	Jitt (%)	0.432	0.464
	Jitta (s)	3.17E-05	3.10E-05		Jitta (s)	3.03E-05	3.26E-05
	rap (%)	0.298	0.291		rap (%)	0.245	0.267
	ppq5 (%)	0.29	0.302		ppq5 (%)	0.239	0.28
3	Jitt (%)	0.646	0.653	4	Jitt (%)	1.178	1.184
	Jitta (s)	4.73E-05	4.77E-05		Jitta (s)	5.47E-05	5.43E-05
	rap (%)	0.325	0.353		rap (%)	0.664	0.705
	ppq5 (%)	0.366	0.39		ppq5 (%)	0.763	0.642
5	Jitt (%)	0.302	0.429	6	Jitt (%)	4.506	5.298
	Jitta (s)	1.33E-05	1.88E-05		Jitta (s)	5.51E-04	6.52E-04
	rap (%)	0.145	0.253		rap (%)	2.488	2.982
	ppq5 (%)	0.178	0.252		ppq5 (%)	2.906	3.5

7	Jitt (%)	0.765	0.724
	Jitta (s)	4.10E-05	3.90E-05
	rap (%)	0.408	0.422
	ppq5 (%)	0.418	0.438

Tabela 7.4: Resultados obtidos pelo algoritmo criado e pelo *Praat* para o Jitter dos sinais da vogal sustentada /e/ de voz com presença de paralisia nas cordas vocais.

*Os sinais (1,2,3,6,7) correspondem a vozes masculinas e os sinais (4,5) a vozes femininas.

7.2 Considerações

Para efeitos de comparações, calculou-se a média das medidas de Jitter. Observamos que as médias de jitt, rap, ppq5 referente à vogal /e/ em homens com voz considerada normal, encontrada no algoritmo implementado, foram respectivamente 0,260%, 0,128% e 0,154% e foram muito semelhantes aos calculados pelo *Praat* 0,258%, 0,133% e 0,154%. A média das mesmas medidas para as mulheres encontrada no algoritmo implementado, foram respectivamente 0,262%, 0,125% e 0,149% foram inferiores aos calculados pelo *Praat*

0,291%, 0,166%, 0,169%.

A medida de Jitter analisada num estudo pelo Zampieri [77] teve como base o seguinte critério: medidas relativas de Jitter PPQ (period perturbation quotient), que emprega uma média de 5 pontos e com valor limite de normalidade de 0,5%, estes foram de 0,23% em homens e 0,31% em mulheres com voz considerada normal, ambos foram muito semelhantes às médias de ppq5 obtida no algoritmo implementado.

Para voz com presença de paralisia nas cordas vocais da vogal /e/, as médias em homens de jitt, rap, ppq5 foram respectivamente 1,369%, 0,753% e 0,844% foram inferiores aos obtidos pelo *Praat* 1,525%, 0,863% e 0,982%. A média das médias de jitt, rap, ppq5 encontrada em mulheres foram respectivamente 0,740%, 0,405%, 0,476%, os valores das medidas de jitt e rap foram inferiores aos calculados pelo *Praat* 0,807% e 0,479% mas a média de ppq5 foi maior ao obtido com o *Praat* 0,447%.

As médias das medidas de Jitter (jitt, rap, ppq5), em mulheres com presença de nódulos nas cordas vocais foram respectivamente 0,363%, 0,190% e 0,219%, os valores das medidas de jitt e rap foram inferiores as obtidas pelo *Praat* 0,435% e 0,249%, enquanto o média da medida ppq5 ficou muito próximo ao calculado pelo *Praat* 0,265.

Não temos referência para discutirmos os valores encontrados de outros estudos para a vogal /e/ de voz com presença com patologia em homens e mulheres da mesma faixa etária ao banco de dados utilizado no estudo.

As médias de Jitter em função do sexo não foram significativamente diferentes, embora o sexo feminino tenha apresentado valor menor que o masculino para voz com presença de paralisia nas cordas vocais. Em vozes consideradas normais obtiveram-se valores menores nos homens que nas mulheres. De acordo com De Felipe [24] e outros autores como Behlau e Tosi (1985, citados por De Felipe [24]) resultados semelhantes foram encontrados e, também, julgaram difícil lançar hipóteses sobre quais seriam os motivos para este melhor controle da voz apresentado pelo sexo feminino. Estes autores levantam a possibilidade de ser pelo maior uso que as pessoas do sexo feminino fazem da voz, o que serviria como treino.

Os valores de Jitter (jitt) obtidos com presença de patologia foram maiores que os obtidos em voz considerada normal, constando-se que uma laringe com patologia produz ciclos vibratórios com maior alteração que uma sem patologia.

A pouca diferença entre nossos resultados e os obtidos pelo *Praat* deve-se, talvez, à maneira dos programas calcularem os parâmetros. Isto são fatores que geram variação nos

valores dos parâmetros acústicos. Devemos considerar também as variações dos tamanhos analisados das amostras. Estas foram selecionadas para que foram aproximadamente iguais.

Capítulo 8

Modelos estocásticos de Jitter

O objetivo deste capítulo é de apresentar e simular modelos estocásticos de Jitter, baseados nos modelos propostos por Schoentgen [61]. São propostos modelos estocásticos que simulam Jitter através dos distúrbios das frequências instantâneas correspondentes as oscilações das cordas vocais.

8.1 Propriedades estatísticas de Jitter

Para simular o Jitter através do modelo de vibração das cordas vocais e do modelo das perturbações estocásticas da frequência instantânea de vibração, deve-se levar em consideração as seguintes características:

- I A distribuição de probabilidade das perturbações relativas ao comprimento de cada ciclo glotal é Gaussiana [39],[49],[35],[59].
- II O valor do Jitter relativo (local) para vozes normais deve-se situar entre os 0,1% a 1% da média dos comprimentos dos ciclos [38], [39], [35].
- III os comprimentos das perturbações dos ciclos vizinhos estão positivamente correlacionados. Isto significa que a perturbação atual depende das perturbações passadas [42], [40], [63]. Schoentgen [61] confirmou estas observações através do modelo auto-regressivo linear para séries temporais na duração das perturbações do ciclo. As ordens dos modelos indicam o número de perturbações passadas que devem ser levados em conta para explicar a perturbação atual.
- IV O maior motivo provável da correlação observada entre os comprimentos das perturbações dos ciclos vizinhos é o microtremor vocal, que é definido como variações da

frequência fundamental dos sons da voz, cuja frequência de modulação varia entre 7 e 10 Hz, podendo ser frequentemente menor ou maior. Os dados quantitativos sobre microtremor são escassos. De acordo com um estudo realizado por Winholtz e Ramig [76], a frequência de modulação pode variar de 4 a 10 Hz com valores medianos de 6 Hz para os falantes do sexo masculino e 5 Hz para os falantes do sexo feminino. O nível de modulação varia de 0,5% a 2,8%. Pode-se dizer, de forma simples, que a frequência de modulação determina a extensão temporal da correlação entre as perturbações e o nível de modulação da dimensão das perturbações.

- V O Jitter parece ser um fenômeno estocástico genuíno. No estudo [63], os autores demonstraram que nenhuma relação não-linear pôde ser encontrada nas séries temporais das perturbações, onde a correlação linear tinha sido removida previamente. Isto implica que os modelos de série temporais lineares são adequados para a representação de Jitter. Resulta que o Jitter deve ser modelado de forma estocástica, porque nos sistemas lineares apenas obtém-se sinais aleatórios na saída quando eles são conduzidos por uma entrada estocástica. Isto é indiretamente confirmado por um modelo fisiológico que casualmente liga as propriedades estatísticas dos espasmos musculares para alguns dos distúrbios da frequência instantânea na vibração [72], [73].
- VI Em média, os comprimentos das perturbações do ciclo aumentam com a duração do ciclo, isto é, quanto maior é a duração do ciclo, maior será o Jitter [38],[47], [39] .
- VII Tem sido repetidamente confirmado o aumento do Jitter na presença de pelo menos uma das patologias da laringe [47], [23], [50], [42], [65].
- VIII Uma hipótese feita implicitamente pela maioria dos autores é que a série temporal dos comprimentos dos ciclos é ergódica. Isto significa que as séries temporais da perturbação devem ser estacionárias e não decompostas em diferentes processos estocásticos que não compartilham propriedades estatísticas semelhantes [75].
- IX Perturbações não processadas e decorrelatadas têm a mesma ordem de grandeza. Isto significa que o desvio padrão das perturbações decorrelatadas é tipicamente maior do que 0,1% e menor do que o desvio padrão das perturbações normais [25].
- X Existem vários modelos de perturbações do ciclo glotal. A maioria dos modelos reproduzem um pequeno subconjunto das propriedades que foram indicadas anteriormente. Alguns dos modelos não tomam em conta o que se conhece sobre o Jitter

e geram perturbações sintéticas com características que não foram observadas. Nenhum dos modelos procura explicar propriedades de Jitter, com uma possível exceção do modelo de Titze [72], [73]. Este modelo, no entanto, não parece se preocupar com as perturbações dos comprimentos dos ciclos da glote como tal. Em vez disso, ele tenta explicar algumas propriedades dos distúrbios da frequência instantânea de vibração.

8.2 Modelos estocásticos

Na primeira seção, apresenta-se o modelo geral da fita das cordas vocais. Na segunda seção, apresentam-se os modelos dos distúrbios estocásticos das frequências instantâneas de vibração, e, na última seção, combinamos os modelos de fita com os modelos de perturbação em quatro modelos.

O objetivo é aproveitar as soluções de forma fechada de modelos simples para fornecer informações sobre o comportamento dos modelos mais complicados. Finalmente, apresenta-se um modelo que combina as características dos quatro submodelos que foram introduzidos previamente, com o modelo de microtremor vocal.

8.2.1 Modelo de fita na vibração das cordas vocais

Titze [70],[71] propôs um modelo para o sincronismo periódico da largura da glote. Este modelo é conhecido, por alguns, como modelo de fita [20]. Titze e Childers [20] têm demonstrado, por meio de simulações, que o sincronismo da largura da glote concorda com o sincronismo da área da glote, com a área de contato das cordas vocais, e com a impedância elétrica do pescoço.

Os comprimentos de ciclo glotal e suas perturbações são, de fato, frequentemente medidos por meio do electrograma [1].

O modelo de fita consiste em considerar uma largura fixa na glote e uma componente senoidal cinemática, constituindo a largura glotal normalizada h , como ilustra-se na Fig.(8.1). A normalização significa que a largura da glote foi dividida pela amplitude, constante de vibração de uma parede da glote:

$$h(y, z, t) = h_0 + \sin\left(\frac{\pi y}{L}\right) \sin\left[2\pi f_0\left(t - \frac{z}{c}\right)\right]. \quad (8.1)$$

Onde h , y e z designam as coordenadas espaciais do movimento das cordas vocais, nas direções ântero-posterior, inferior-superior e laterais; t é o tempo, c é a velocidade da onda na superfície da da corda vocal, f_0 é a frequência de vibração de uma das cordas vocais, e L e h_0 designam o comprimento na vibração das cordas vocais e a semi-largura glotal pré-fonatória normalizada respectivamente.

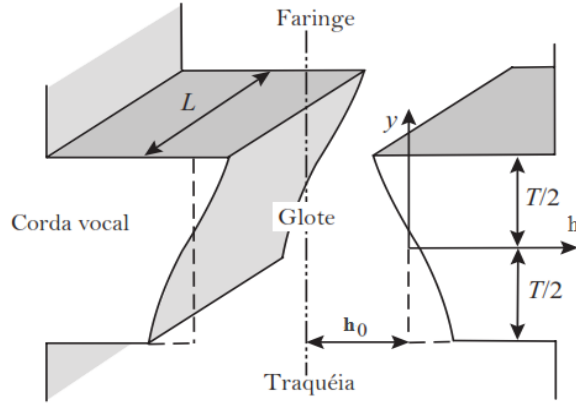


Figura 8.1: Modelo da onda de mucosa das cordas vocais [43].

No contexto do modelo de fita, os movimentos interpenetrantes esquerdo e direito das paredes glotais são periódicos e simétricos. O atraso de fase ϕ entre os movimentos das margens superior e inferior da glote é dado por:

$$\phi = \frac{2\pi f_0 T}{c}. \quad (8.2)$$

Sendo T a espessura das cordas vocais em vibração.

O comprimento das cordas vocais e sua espessura dependem da frequência de vibração f_0 através das seguintes expressões:

$$L = 0,038 L_0^{1,6} f_0^{0,5}, \quad T = T_0 L_0/L.$$

Os valores das constantes L_0 e T_0 são, em média, 1,6 cm e 0,45 cm para homens e 1,0 centímetros e 0,35 cm para as mulheres. A Eq.(8.1) pode ser interpretada como uma solução parametrizada de um modelo de muitos graus de liberdade das cordas vocais em vibração. Além de duas constantes anatômicas (T_0 e L_0), o modelo é caracterizado por três parâmetros, que são a frequência f_0 , a fase de atraso ϕ , e, implicitamente, o acoplamento entre os movimentos da esquerda e direita das paredes glotais.

Os modelos são baseados nas seguintes particularidades:

- (a) O modelo mais simples a ser considerado é obtido partindo-se do princípio que a glote é um prisma reto, as paredes glotais são rígidas, e os movimentos das paredes esquerda e direita são simétricas. Isso significa que o atraso de fase ϕ é zero e a semi-largura da glote não depende da profundidade da glote y . Consequentemente, a largura normalizada h da glote pode ser descrita por uma única fase θ . Podemos escrever:

$$h(\theta) = 2h_0 + 2 \sin(\theta). \quad (8.3)$$

- (b) Aprimorando um pouco mais, pode-se considerar que os movimentos mais comuns das cordas vocais consistem em uma superposição de um movimento da esquerda para a direita na direção lateral e um movimento convergente-divergente na direção inferior-superior. Isto significa que as margens da glote dos movimentos inferior (índice i) e superior (índice s) devem ser descritos distintamente:

$$h_i(\theta_i) = 2h_0 + 2 \sin(\theta_i), \quad h_s(\theta_s) = 2h_0 + 2 \sin(\theta_s). \quad (8.4)$$

- (c) Como é improvável que os distúrbios das paredes glotais esquerda e direita sejam exatamente simétricos, as Eq.(8.3) e Eq.(8.4) devem ser reescritas considerando diferentes perturbações à esquerda (índice l) e à direita (índice r).

$$h(\theta) = 2h_0 + \sin(\theta_r) - \sin(\theta_l + \pi), \quad (8.5a)$$

$$h(\theta) = 2h_0 + 2 \cos\left(\frac{\theta_r - \theta_l}{2}\right) \sin\left(\frac{\theta_r + \theta_l}{2}\right), \quad (8.5b)$$

$$h(\theta) = \frac{\theta_r + \theta_l}{2}. \quad (8.5c)$$

A expressão 8.5b é obtida a partir da expressão 8.5a por meio de relações trigonométricas elementares. A Eq.(8.5c) deriva da Eq.(8.5b). Ele estabelece a ligação entre as fases da esquerda e da direita e da fase da largura da glote.

- (d) Quando são levados em conta as diferenças entre os movimentos esquerda e direita e as margens inferior e superior das cordas vocais, então, as Eq.(8.4) e Eq.(8.5) são

combinadas em Eq.(8.6):

$$\begin{aligned} h_i(\theta_{hi}) &= 2h_0 + \sin(\theta_{i,r}) - \sin(\theta_{i,l} + \pi), \\ h_s(\theta_{hs}) &= 2h_0 + \sin(\theta_{s,r}) - \sin(\theta_{s,l} + \pi), \\ \theta_{hi} &= \frac{\theta_{i,r} + \theta_{i,l}}{2}, \quad \theta_{hs} = \frac{\theta_{s,r} + \theta_{s,l}}{2}. \end{aligned} \quad (8.6)$$

Equações da (8.3)-(8.6) do modelo de fita relatam as fases das vibrações de pontos selecionados das paredes glotais para a fase de vibração da largura da glote.

Por definição, a frequência instantânea é a taxa de variação da fase e a duração do ciclo glotal é o intervalo de tempo durante o qual a fase aumenta 2π . Consequentemente, as relações entre frequência instantânea f , a duração do ciclo glotal τ_0 e frequência fundamental f_0 são as seguintes:

$$\frac{d\theta}{dt} = 2\pi f, \quad \int_{\theta_{start}}^{\theta_{start}+2\pi} d\theta = 2\pi = \int_{t_{start}}^{t_{start}+\tau_0} 2\pi f dt, \quad f_0 = \frac{1}{\tau_0}.$$

8.2.2 Modelos das perturbações estocásticas de frequência instantânea

O movimento cíclico das cordas vocais deve ser descrito por meio da frequência instantânea, a qual concorda com a frequência fundamental apenas quando o movimento é periódico.

Nesta seção, discutiremos as hipóteses sobre as perturbações estocásticas das frequências instantâneas de vibração nas paredes glotais.

Processos não correlatos ou processos de ruído branco são sequências de variáveis aleatórias não correlacionadas e_n que são identicamente distribuídas. Um dos processos mais simples de ruído branco é o processo de dois pontos (Eq.(8.7)):

$$e_n = \sqrt{\Delta} \begin{cases} +1, & p = 0.5, \\ -1, & p = 0.5. \end{cases} \quad (8.7)$$

Onde p designa a probabilidade e Δ uma constante, que mais adiante será interpretada como um passo de tempo.

Pode-se facilmente verificar que a média M do processo de dois pontos é igual a 0 e a variância é igual a Δ . Na revisão teórica, indicamos que em vogais sustentadas por indivíduos saudáveis sem vibrato, as propriedades de Jitter são condicionadas pelo

chamado microtremor. O microtremor deve ser diferenciado de tremor patológico que caracteriza as vozes de indivíduos que sofrem de doenças neurológicas, como a doença de Parkinson ou tremor essencial.

Winholtz e Ramig[76] publicaram os espectros de sinais de microtremor. Para indivíduos saudáveis, esses espectros exibiram um único pico espectral. A posição do pico é a frequência do microtremor e a largura indica a largura de banda do microtremor.

Filtrando o ruído branco, por meio de um modelo auto-regressivo linear de segunda ordem, o qual é um método direto para a geração de sinais, como os sinais de microtremor, caracterizados por um espectro de potência de um único pico. O símbolo y indica o sinal de microtremor sintético, a_i as constantes dos parâmetros do filtro, e é a entrada (ruído branco), e b o valor máximo do pico do ruído branco. O índice n é o intervalo de tempo. Os Símbolos Ψ e B , finalmente, indicam a frequência de microtremor e a largura de banda em radianos por amostra. O coeficiente a_0 normaliza o ganho do filtro, de modo que este não depende da frequência ou da largura de banda. Os coeficientes do filtro são os seguintes [66]:

$$\begin{aligned} y_n &= a_1 y_{n-1} + a_2 y_{n-2} + a_0 b e_n, \\ R &= 1 - \frac{B}{2}, \cos \varphi = \frac{2R}{1+R^2} \cos \Psi, \\ a_0 &= (1 - R^2) \sin \varphi, a_1 = \frac{4R^2}{1+R^2} \cos \Psi, a_2 = -R^2. \end{aligned} \quad (8.8)$$

Para incorporar o microtremor, será considerado um modelo linear de primeira ordem Eq.(8.9), o qual sempre posiciona em 0 Hz o pico espectral. O coeficiente a_1 corrige a queda do contorno espectral com o incremento de frequência. O parâmetro b é o valor do pico do processo de ruído branco de entrada:

$$y_n = a_1 y_{n-1} + b e_n. \quad (8.9)$$

Interpretações de expressões Eq.(8.8) e Eq.(8.9) são as seguintes:

- Quando a entrada do filtro (Eq.(8.9)) é um ruído branco (Eq.(8.7)) processo de dois pontos, a saída é uma soma de decomposições de curvas exponenciais cujos valores iniciais alternam estocasticamente entre $\pm b$.
- Da mesma forma, quando o processo branco de dois pontos é a entrada para o filtro linear de segunda ordem, a saída é uma soma estocástica de funções cosseno que se decompõem de forma exponencial.

8.2.3 Modelos estocásticos de Jitter

Nesta seção, combinamos o modelo de fita com o modelo dos distúrbios estocásticos que foram discutidos na seção anterior. O objetivo é explicitar as relações que existem entre as perturbações estocásticas das fases de vibração das paredes glotais, por um lado, e as perturbações estocásticas dos comprimentos do ciclo glotal. Examinamos a seguir cinco modelos de Jitter. Os objetivos do estudo dos cinco modelos são os seguintes:

1. **Modelo I:** considera uma perturbação, não correlata da frequência instantânea da vibração, de uma parede rígida da glote. Os distúrbios da frequência instantânea causam perturbações no comprimento do ciclo glotal. Na prática, a duração do ciclo das perturbações é estacionário e ergódico.
2. **Modelo II:** a perturbação não correlata é substituída por um microtremor artificial Eq.(8.9). O propósito é mostrar que o microtremor correlaciona os ciclos vizinhos das perturbações glotais.
3. **Modelo III:** simula o acoplamento entre as paredes glotais esquerda e direita, as quais não são perturbadas de forma idêntica. Procura mostrar que o acoplamento sincroniza os movimentos das paredes glotais esquerda e direita e impede o afastamento de suas fases. Mostra, também, que o acoplamento assimétrico aumenta a fase das perturbações. O acoplamento assimétrico é, portanto, uma possibilidade, através do qual patologias da laringe podem influenciar o valor do Jitter.
4. **Modelo IV:** inclui um atraso entre os limites glotais inferiores e superiores. O objetivo é o de mostrar que o atraso correlaciona os comprimentos dos ciclos adjacente das perturbações glotais.
5. **Modelo V:** combina as propriedades dos quatro modelos anteriores com o modelo de microtremor.

8.2.3.1 Modelo I: Passeio aleatório

Nesse modelo, assume-se que a glote é um prisma reto que evolui devido ao movimento das paredes da glote. Considera-se o movimento de apenas uma das paredes. Isto significa que a largura da glote depende do tempo e sua evolução será modelada como senoidal. A largura da glote e a parede glotal estão em fase. A evolução temporal da fase de

perturbação deve ser escrita como:

$$\frac{d\theta}{dt} = 2\pi f. \quad (8.10)$$

Onde f designa a frequência instantanea e θ a fase.

Considerando pequenos valores de Jitter a Eq.(8.10) pode ser rescrita:

$$\frac{d\theta}{dt} = 2\pi(f_0 + b\xi). \quad (8.11)$$

Onde f_0 refere-se à frequência instantânea imperturbável de vibração, b é uma constante e ξ é o ruído branco que simula distúrbios estocásticos.

A Eq.(8.11) é uma heurística só porque não existe nenhuma função ordinaria cuja derivada é um processo de ruído branco. Eq.(8.11) deve, portanto, ser transformada em uma equação diferencial estocástica Eq.(8.12) [30].

$$\frac{d\theta}{dt} = 2\pi(f_0 dt + b dW). \quad (8.12)$$

O dt é um diferencial ordinário e dW é um diferencial estocástico, que tem propriedades matemáticas diferentes a partir de um diferencial comum.

A Eq.(8.12) pode ser transformado em uma equação de diferença estocástica que pode ser resolvido numericamente ou algebricamente. O símbolo Δ designa o passo de tempo e n o índice. O símbolo Δ_W é uma diferença estocástica. Resolvendo as equações temos:

$$\begin{aligned} \theta_{n+1} &= \theta_n + 2\pi(f_0\Delta + b\Delta_W), \\ \theta_{n+1} &= \theta_n + 2\pi f_0\Delta + 2\pi b\sqrt{\Delta} \begin{cases} +1, & p = 0.5, \\ -1, & p = 0.5. \end{cases} \end{aligned} \quad (8.13)$$

No trabalho [46], foi provado que a diferença estocástica pode ser aproximada por meio de um processo de dois pontos Eq.(8.7). A condição é que as propriedades estatísticas da solução da Eq.(8.13), em tempo discreto, são necessárias para aproximar as propriedades estatísticas da solução da Eq.(8.12) em tempo contínuo. A Eq.(8.13) pode ser resolvida algebricamente por substituições repetidas:

$$\theta_n = 2\pi f_0 n\Delta + 2\pi b \sum_{i=0}^n e_n, \quad \theta_N = 2\pi. \quad (8.14)$$

8.2.3.2 Modelo II: Ruído filtrado

Nesta seção, calcula-se comprimentos dos ciclos das perturbações quando a frequência instantânea é perturbada por uma instância de microtremor artificial em vez de um ruído não correlato. Para simplificar, consideramos microtremor sintético, Eq.(8.9), cuja autocorrelação é um declínio exponencial e o pico espectral está sempre posicionada em 0 Hz.

$$\begin{aligned}\theta_{n+1} &= \theta_n + 2\pi f_0 n \Delta + 2\pi y_n, \\ y_{n+1} &= ay_n + be_n, \quad 0 < a < 1, \\ \theta_n &= \theta_0 + 2\pi f_0 n \Delta + 2\pi \sum_{i=1}^n y_{i-1}.\end{aligned}\tag{8.15}$$

As Eq.(8.15) podem ser resolvidas através de substituições repetidas. Como no Modelo I, a duração do ciclo perturbado é obtida através da inserção de 2π para a fase θ , N para o índice n , e zerar θ_0 na solução de θ_n .

8.2.3.3 Modelo III: Oscilador acoplado perturbado

A maior parte dos modelos mecânicos das cordas vocais simulam o movimento de apenas uma delas. A hipótese é que o movimento das cordas vocais é simétrico.

Como não se pode assumir que os distúrbios estocásticos das frequências instantâneas esquerda e direita são simétricos, inclui-se um acoplamento entre as fases da parede glotal esquerda e direita. O Símbolo K é a constante de acoplamento. O termo de acoplamento é a função seno da subtração entre a fase esquerda da fase direita. Uma das razões para a inclusão da função seno é que uma mudança de 2π de uma fase não tem consequências observáveis.

Levando-se em conta que $\sin x \sim x$ quando x é pequeno, vê-se que quando a fase esquerda é maior do que a fase direita o termo de acoplamento diminui a fase esquerda e a fase direita aumenta:

$$\begin{aligned}\theta_{l,n+1} &= \theta_{l,n} + 2\pi f_0 \Delta + K \sin(\theta_{l,n} - \theta_{r,n}) + 2\pi be_{l,n}, \\ \theta_{r,n+1} &= \theta_{r,n} + 2\pi f_0 \Delta + K \sin(\theta_{r,n} - \theta_{l,n}) + 2\pi be_{r,n}.\end{aligned}\tag{8.16}$$

No âmbito do Modelo III, assume-se que os distúrbios não são nem espacialmente nem temporalmente correlacionados. Os símbolos e_l e e_r , portanto, designam variáveis aleatórias não correlacionadas que são distribuídos de acordo com a lei de probabilidade

(Eq.(8.7)).

A Eq.(8.17a) mostra que o momento da largura da glote depende do produto do cosseno com a função seno. O argumento da função cosseno é a diferença entre as fases esquerda e direita e o argumento da função seno é a soma de ambas fases. As equações 8.16, portanto, são reescritas como uma soma e uma diferença:

$$h_n = 2h_0 + \cos\left(\frac{\theta_{r,n} - \theta_{l,n}}{2}\right) \sin\left(\frac{\theta_{r,n} + \theta_{l,n}}{2}\right), \quad (8.17a)$$

$$\frac{\theta_{r,n+1} + \theta_{l,n+1}}{2} = \frac{\theta_{r,n} + \theta_{l,n}}{2} + 2\pi f_0 \Delta + 2\pi b \frac{e_{l,n} + e_{r,n}}{2}, \quad (8.17b)$$

$$\frac{\theta_{r,n+1} - \theta_{l,n+1}}{2} = \frac{\theta_{r,n} - \theta_{l,n}}{2} - K \sin(\theta_{r,n} - \theta_{l,n}) + 2\pi b \frac{e_{r,n} - e_{l,n}}{2}. \quad (8.17c)$$

A soma das fases é o argumento da função seno na largura glote Eq.(8.17a). Eq.(8.17b) mostra que a soma não depende do acoplamento quando o acoplamento é simétrico. A Eq.(8.17b) é formalmente idêntico ao Modelo I livre de acoplamento.

A diferença das fases é o argumento da função cosseno de largura glote Eq.(8.17a). Eq.(8.17c) mostra que a diferença depende do acoplamento entre a esquerda e a direita. Quaisquer contrastes entre os modelos acoplados e desacoplados são, portanto, devido à função cosseno. Examinamos a seguir em que condições a influência da função cosseno é pequena. Formalmente, Eq.(8.17c) e sua solução são as seguintes, tendo em conta que $\sin x \sim x$ quando x é pequeno:

$$\alpha_n = \frac{\theta_{r,n} - \theta_{l,n}}{2}, \quad \alpha_{n+1} = \alpha_n(1 - 2K) + 2\pi b \frac{e_{l,n} - e_{r,n}}{2},$$

$$\alpha_n = \alpha_0(1 - 2K)^n + \left[\sum_{i=1}^n (1 - 2K)^{n-i} 2\pi b \frac{e_{r,i} - e_{l,i}}{2} \right].$$

A solução α_n mostra que, quando o modelo é livre de perturbação, a solução de estado estacionário é $\theta_l = \theta_r$ contanto que a constante de acoplamento é positivo e inferior a 0,5. Quando os distúrbios são finitos, o lado direito da solução deve ser tomada em consideração, o qual é uma soma ponderada das perturbações anteriores e actuais. A soma diminui quando o acoplamento aumenta constantemente. No modelo, o argumento da função cosseno, por conseguinte é, menor do forte acoplamento e menor dos distúrbios.

A generalização do modelo anterior é aquele para o qual as constantes de acoplamento esquerda e direita não são idênticos. A soma e a diferença das fases esquerda e direita,

são então como se mostra na seguinte Eq.(8.18), assumindo que $\sin x \sim x$. As Equações 8.18 mostram que $\theta_l = \theta_r$ é uma solução estável, quando as interferências externas são zero. A condição é que a soma das constantes de acoplamento seja inferior a um:

$$\begin{aligned} \frac{\theta_{l,n+1} + \theta_{r,n+1}}{2} &= \frac{\theta_{l,n} + \theta_{r,n}}{2} + 2\pi f_0 \Delta - (k_l - K_r) \frac{\theta_{l,n} - \theta_{r,n}}{2} + 2\pi b \frac{e_{l,n} + e_{r,n}}{2}, \\ \frac{\theta_{l,n+1} - \theta_{r,n+1}}{2} &= \frac{\theta_{l,n} - \theta_{r,n}}{2} + 2\pi f_0 \Delta - (k_l + K_r) \frac{\theta_{l,n} - \theta_{r,n}}{2} + 2\pi b \frac{e_{l,n} - e_{r,n}}{2}. \end{aligned} \quad (8.18)$$

A soma e a diferença das fases determinam o timing da largura da glote via relação da Eq.(8.17a). Para o acoplamento assimétrico, a soma contém um termo adicional, que depende da diferença entre as constantes de acoplamento. Uma vez que o termo cosseno na largura da glote é quase constante quando o acoplamento é forte e as perturbações fracas, o acoplamento deverá ter uma influência considerável sobre o Jitter unicamente no caso assimétrico.

8.2.3.4 Modelo IV: Onda da corda vocal perturbada

No modelo da fita os movimentos das margens superior e inferior da glote podem estar fora de fase. O modelo IV simula as consequências dos movimentos fora de fase, assumindo que as perturbações externas das margens superior e inferior são idênticas e livre de correlação. Símbolo D designa o atraso no número de intervalos de tempo entre as fases das margens inferior e superior que são indexados i e s :

$$\begin{aligned} \theta_{i,n+1} &= \theta_{i,n} + 2\pi f_0 \Delta + 2\pi b e_n, \\ \theta_{s,n+1} &= \theta_{i,n-D} + 2\pi f_0 \Delta + 2\pi b e_n. \end{aligned} \quad (8.19)$$

As soluções Eq.(8.20) da Eq.(8.19) são as seguintes:

$$\begin{aligned} \theta_{i,n} &= 2\pi f_0 \Delta + 2\pi b \sum_{j=1}^n e_{j-1}, \\ \theta_{s,n} &= 2\pi f_0 (n - D) \Delta + 2\pi b \sum_{k=1}^{n-D} e_{k-1} + 2\pi b \sum_{l=1}^n e_{l-1}. \end{aligned} \quad (8.20)$$

Convencionalmente, os comprimentos do ciclo são medidas do sinal de voz no fechamento da glote por marcadores monitoradas. Na maioria das vezes, a margem superior da glote abre e fecha depois da margem inferior. Por conseguinte, definir a duração do ciclo, como o número de intervalos de tempo necessários para aumentar 2π a fase de margem superior. A duração do ciclo τ é, portanto, obtido por inserção $N = n - D$ e $\theta = 2\pi$ na

equação que descreve o movimento da margem superior:

$$\tau = \tau_0 - b\tau_0 \sum_{k=0}^N e_k - b\tau_0 \sum_{l=D}^{N+D} e_l. \quad (8.21)$$

O comprimento τ mostra que os comprimentos de ciclo são perturbados por dois passeios aleatórios. O primeiro é a perturbação direta que é o mesmo para as margens superior e inferior. A segunda é a perturbação da margem inferior, que está atrasada.

8.2.3.5 Modelo V: Modelo geral estocástico de Jitter

O modelo a seguir combina as propriedades dos modelos que foram discutidos até agora com as Eq.(8.8) do modelo do microtremor vocal. De cima para baixo, as Eq.(8.22) descrevem as fases inferior e superior e as margens glótica esquerda e direita.

As duas últimas equações correspondem ao microtremor sintético esquerdo e direito, que é modelado por meio do modelo linear autoregressivo de segunda ordem Eq.(8.8). Os distúrbios completos são a soma do microtremor sintético e ruído não correlato que simula distúrbios estocásticos diferentes do microtremor. Os valores de pico do ruído e microtremor sintético livre de correlação são controlados de forma independente:

$$\begin{aligned} \theta_{i,l,n+1} &= \theta_{i,l,n} + 2\pi f_0 \Delta - K_l \sin(\theta_{i,l,n} - \theta_{i,r,n}) + y_{l,n} + 2\pi b_2 e_{2,l,n}, \\ \theta_{i,r,n+1} &= \theta_{i,r,n} + 2\pi f_0 \Delta - K_r \sin(\theta_{i,r,n} - \theta_{i,l,n}) + y_{r,n} + 2\pi b_2 e_{2,r,n}, \\ \theta_{s,l,n+1} &= \theta_{i,l,n-D} + 2\pi f_0 \Delta - K_l \sin(\theta_{s,l,n} - \theta_{s,r,n}) + y_{l,n} + 2\pi b_2 e_{2,l,n}, \\ \theta_{s,r,n+1} &= \theta_{i,r,n-D} + 2\pi f_0 \Delta - K_r \sin(\theta_{s,r,n} - \theta_{s,l,n}) + y_{r,n} + 2\pi b_2 e_{2,r,n}, \\ y_{l,n+1} &= a_1 y_{l,n} + a_2 y_{l,n-1} + 2\pi a_0 b_1 e_{1,l,n+1}, \\ y_{r,n+1} &= a_1 y_{r,n} + a_2 y_{r,n-1} + 2\pi a_0 b_1 e_{1,r,n+1}. \end{aligned} \quad (8.22)$$

Capítulo 9

Simulações e resultados

Para simular os modelos estocásticos de Jitter, baseados nos modelos propostos por Schoentgen [61], implementou-se um algoritmo para cada modelo, no software *Matlab*. Cada algoritmo de modelagem de Jitter baseou-se nas diversas hipóteses descritas no capítulo anterior e os modelos foram resolvidos numericamente. As rotinas correspondentes aos modelos (I, II, III, IV, V) podem ser vistas no Apêndice D.

Os sinais da largura da glote obtidos em cada uma das simulações dos algoritmos implementados no software *Matlab* foram analisados no programa *Praat*. Para cada simulação foram calculados os valores de Jitter e suas diferentes formas: *jitt*, *jitta*, *rap* e *ppq5*.

Em linhas gerais, o objetivo é o de variar os parâmetros, dos diversos modelos, considerando diferentes valores de frequência fundamental e diferentes tempo de simulação. Com isso, pretende-se descobrir que valores de parâmetros vão originar sinais com características de patologia e se os diferentes valores de frequência fundamental e tempo de simulação influencia nos resultados.

Os modelos I e II consideram um ruído branco não correlato e o microtremor, respectivamente, para gerar distúrbios na frequência instantânea de vibração em apenas uma parede glotal.

O modelo III leva em conta o acoplamento entre ambas paredes glotais, esquerda e direita, as quais não são perturbadas de forma idêntica, pelo ruído branco não correlato.

O modelo IV envolve distúrbios estocásticos não correlato das margens glotais superiores e inferiores defasadas, em apenas uma parede da glote.

Finalmente, o modelo V combina as propriedades dos quatro modelos anteriores e o

modelo de microtremor.

O objetivo, em cada caso, foi simular os comprimentos do ciclo glotal perturbados. Os modelos I, III e IV consideram um ruído branco não correlato que perturba a frequência instantânea, a qual causa distúrbios nos comprimentos do ciclo glotal. O modelo II considera um microtremor artificial, ao invés de um ruído branco.

O ruído branco que gera as perturbações estocásticas na frequência instantânea envolve uma constante b , que está relacionada com o valor máximo do distúrbio. Para gerar o microtremor no modelo II foi usado um ruído branco, filtrado por meio de um modelo linear de primeira ordem, e para o modelo V foi considerado um modelo auto-regressivo linear de segunda ordem.

Foram considerados cinco valores para a constante b : 0.05, 0.10, 0.15, 0.20 e 0.25 a fim de obter valores de Jitter dentro da faixa considerada normal e verificar a partir de que valor obtém-se patologia, exceto para o modelo II, que foram considerados apenas três valores: 0.05, 0.10, 0.25, além da variação do coeficiente do filtro de primeira ordem, a_1 , com três valores 0.25, 0.5 e 0.8, para cada valor de b .

Em cada simulação dos modelos I, III, IV e V foram assumidos três valores para os intervalos de tempo, 1000, 5000 e 10000, enquanto que, para o modelo II foram assumidos apenas dois valores para os intervalos de tempo, 5000 e 10000, com a finalidade de observar as variações dos resultados de Jitter considerando o tamanho do sinal, menor ou maior, para verificar se a quantidade de períodos fundamentais do sinal afetava os resultados.

Foram assumidos três valores de frequência fundamental: 110 Hz para homens (normalmente, nos homens as frequências situam-se entre 80 e 150 Hz), 200 Hz para mulheres (normalmente, as frequências encontram-se entre 150 e 250 Hz) e o terceiro valor, 400 Hz, normalmente, é um valor atingindo nos bebês.

Em total foram 198 simulações, distribuídas da seguinte maneira:

- 45 simulações no Modelo I, particionado em 15 simulações para cada valor de frequência fundamental.
- 54 simulações no Modelo II, particionado em 18 simulações para cada valor de frequência fundamental.
- 45 simulações no Modelo III, particionado em 15 simulações para cada valor de frequência fundamental.
- 45 simulações no Modelo IV, particionado em 15 simulações para cada valor de

frequência fundamental.

- 9 simulações no Modelo V, particionado em 3 simulações para cada valor de frequência fundamental.

Os valores dos parâmetros: passo do tempo Δ , frequência do microtremor Ψ , largura de banda do microtremor B , constantes de acoplamento k_r e k_l , valor máximo do distúrbio não correlato b_2 , valor máximo do ruído b_1 e velocidade da onda c foram fixados em cada simulação assumindo-se os valores apresentados nos modelos propostos por Schoentgen [61], mostrados na Tab.(9.1).

Parâmetro		Valor padrão	Faixa
Constante de acoplamento (direita)	k_r	0.1	0-0.1
Diferencia (direita-esquerda)	$\text{abs}(k_r \text{ e } k_l)$	0	0-0.2
Velocidade de onda (cm/s)	c	200	200-400
Frequência do microtremor (Hz)	Ψ	6	1-8
Largura de banda do microtremor (Hz)	B	4	2-6
Passo do tempo (s)	Δ	10^{-5}	
Valor máximo do distúrbio não correlato	(b_2)	0.05	0.001-0.1
Valor máximo do ruído do Microtremor	(b_1)	0.25	0.1-0.5

Tabela 9.1: Valores dos parâmetros de modelagem da análises de regressão múltipla linear.

A seguir, detalha-se para cada modelo, os resultados obtidos com os valores dos parâmetros fixados em cada simulação.

Modelo I

(i) $f_0 = 110\text{Hz}$; intervalos de tempo 1000, 5000 e 10000:

- Os valores de $jitt$ foram normais com b entre 0.05 e 0.10. Para valores dentro do limite para detectar patologias (1.04%) foram atingidos a partir de $b \geq 0.15$, como observa-se na Fig.(9.1), valores acima da linha tracejada apresentam patologia.

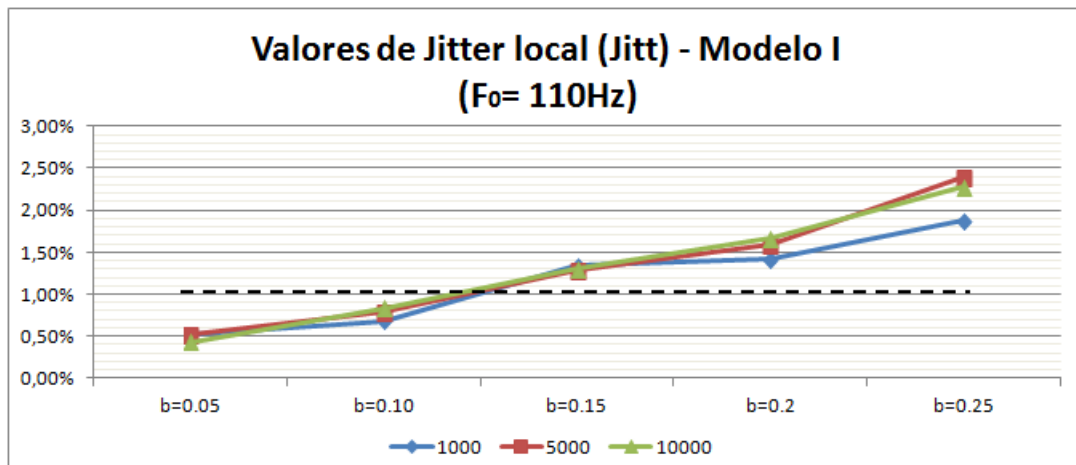


Figura 9.1: Gráfico com os valores de *jitt* para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *jitta* dentro do limite para detectar patologias ($83.200\mu s$) foram atingidos a partir de $b \geq 0.10$, exceto para o intervalo de tempo fixado em 1000. Com $b = 0.05$, para os três intervalos de tempo, foram normais, como mostra a Fig.(9.2).

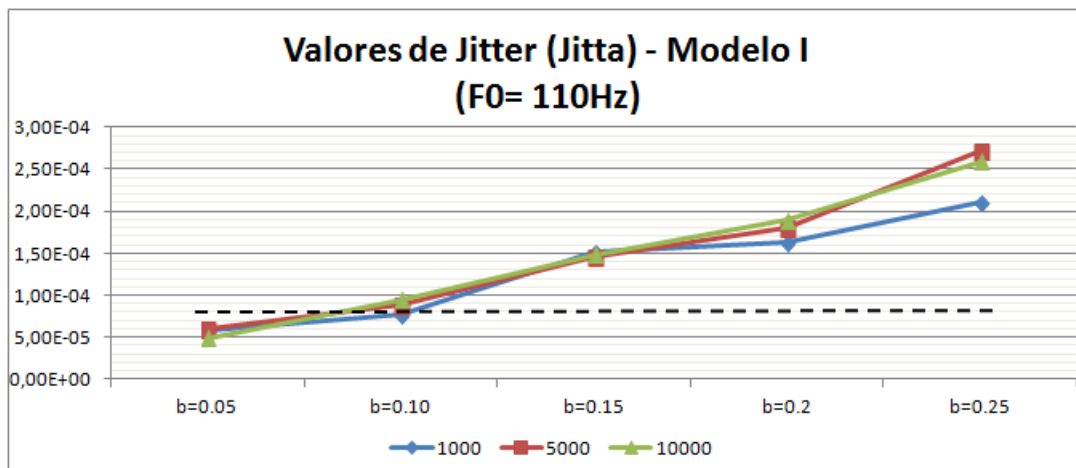


Figura 9.2: Gráfico com os valores de *jitta* para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *rap* dentro do limite para detectar patologias (0.680%) foram alcançados com $b \geq 0.15$, exceto para o intervalo de 5000 com $b = 0.15$. Com valores de b 0.05 e 0.10 foram normais, para os três intervalos de tempo, como pode ser observado na Fig.(9.3).

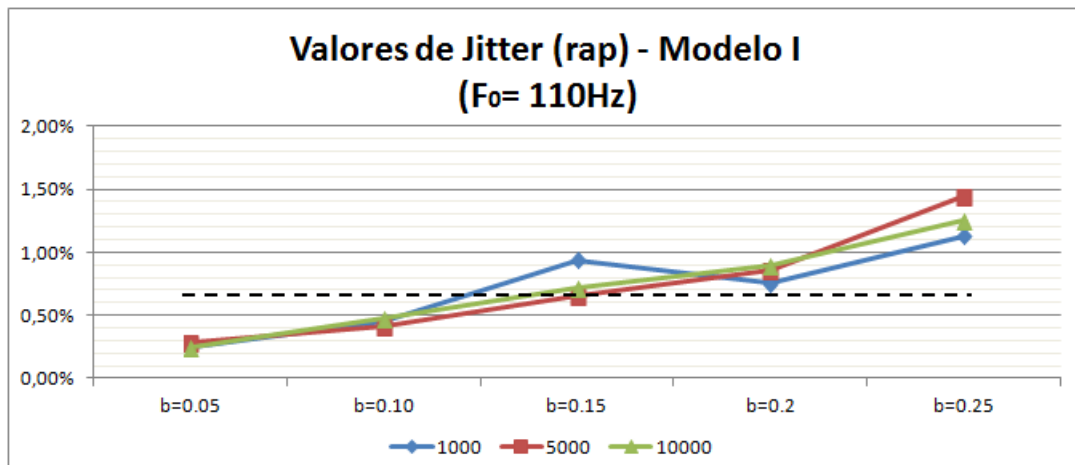


Figura 9.3: Gráfico com os valores de *rap* para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *ppq5* foram normais com valores de b 0.05 e 0.10. Os valores dentro do limite para detectar patologias (0.840%) foram atingidos com $b \geq 0.15$, como podem ser observados na Fig.(9.4), valores acima da linha tracejada apresentam patologia.

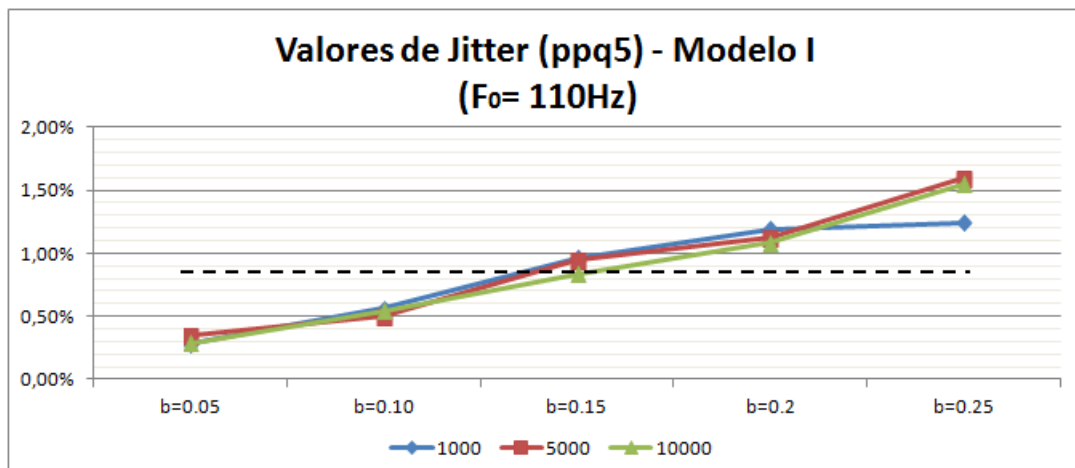


Figura 9.4: Gráfico com os valores de *ppq5* para $f_0 = 110Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

(ii) $f_0 = 200Hz$; intervalos de tempo 1000, 5000 e 10000:

- As medidas de *jitt* normais foram obtidos com valores de b entre 0.05, 0.10 e 0.15. Para $b \geq 0.2$ os valores obtidos representam patologia, como pode ser observados na Fig.(9.5).

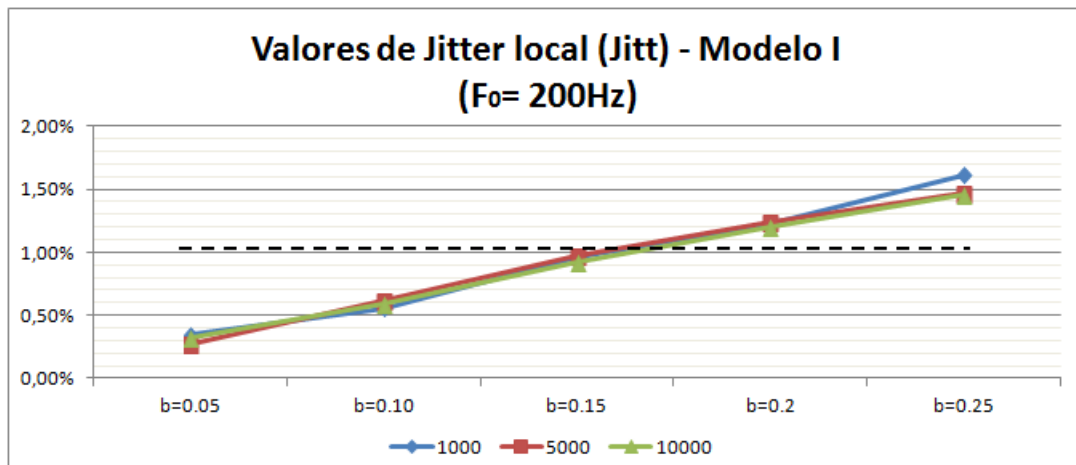


Figura 9.5: Gráfico com os valores de *jitt* para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *jitta* normais foram alcançados com valores de b 0.05, 0.10, 0.15 e 0.20. Com $b=0.25$, foram obtidos valores correspondentes a patologia como podem ser observados acima da linha tracejada na Fig.(9.6).

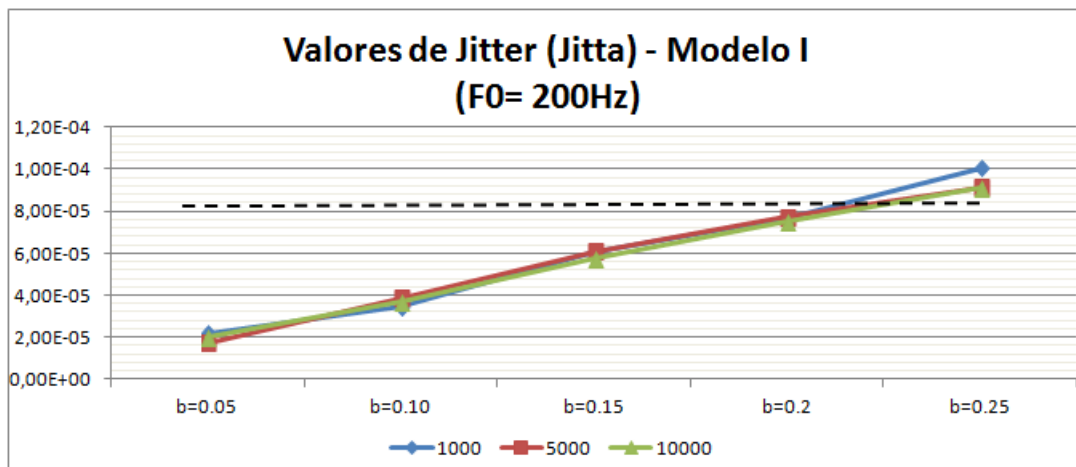


Figura 9.6: Gráfico com os valores de *jitta* para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *rap* normais foram obtidos com valores de b 0.05, 0.10, 0.15 e 0.20. Para $b = 0.20$ e intervalo de tempo 10000 foi obtido um valor dentro da faixa para detectar patologias (0,680%), como podem ser observados na Fig.(9.7).

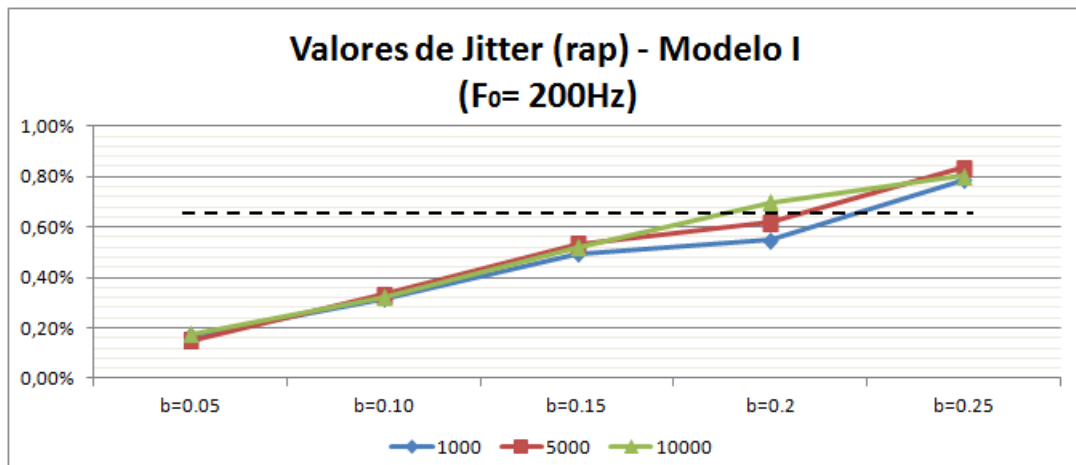


Figura 9.7: Gráfico com os valores de *rap* para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *ppq5* foram normais com valores de b 0,05, 0,10, 0,15 e 0,20. Para $b = 0,20$ e intervalo de tempo 5000 foi obtido um valor limite para detectar patologias (0,840%), como pode ser observados na Fig.(9.8).

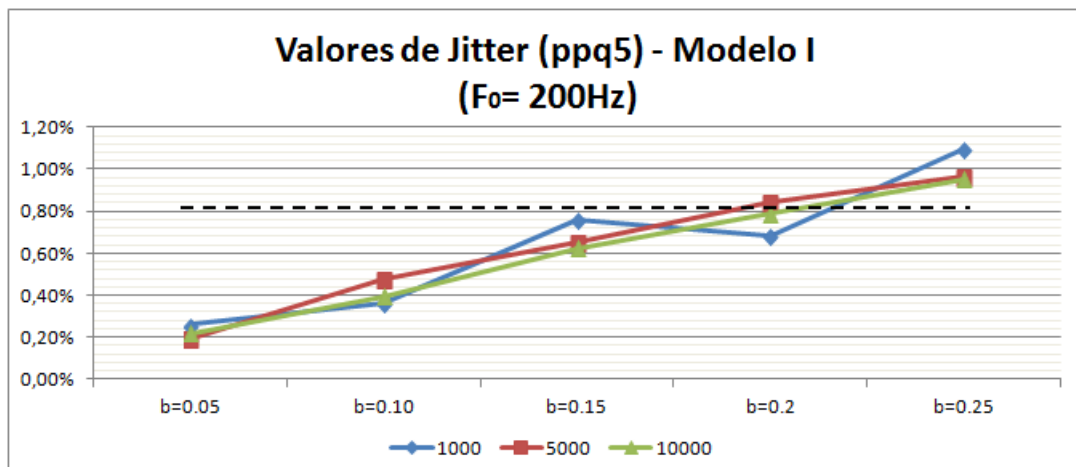


Figura 9.8: Gráfico com os valores de *ppq5* para $f_0 = 200Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

(iii) $f_0 = 400Hz$; intervalos de tempo 1000, 5000 e 10000:

- Os valores de *jitt* foram normais com valores de b 0,05 e 0,20. Com $b = 0,25$ foram obtidos valores correspondentes a patologia, como podem ser observados na Fig.(9.9).

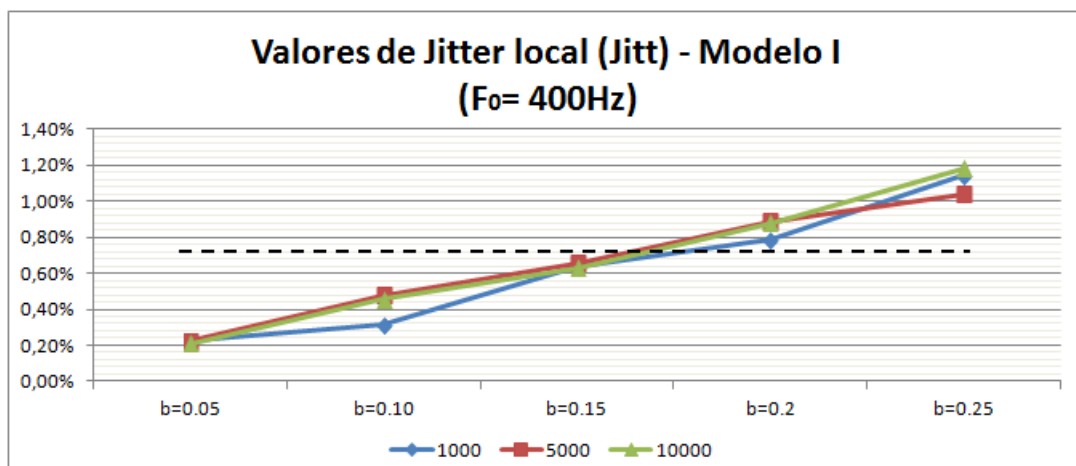


Figura 9.9: Gráfico com os valores de *jitt* para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *jitta* dentro do faixa considerada normal para todos os valores assumidos de b , como observa-se na Fig.(9.10).

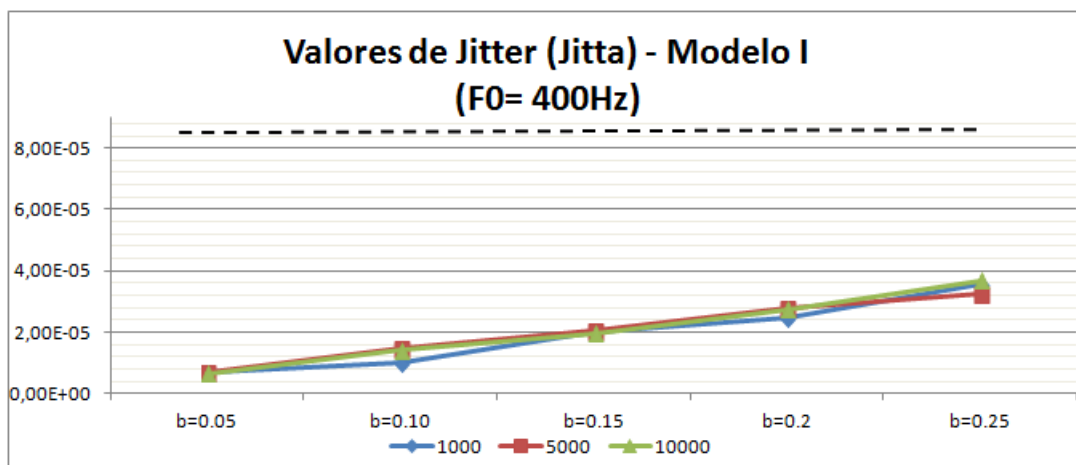


Figura 9.10: Gráfico com os valores de *jitta* para $f_0 = 400Hz$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *rap* para quase todos os intervalos de tempo 1000, 5000 e 10000 e valores de b foram normais, exceto para o intervalo de tempo 10000 e $b = 0.25$ que apresenta o um valor limite para detectar patologias (0,680%), como pode-se ver na Fig.(9.11).

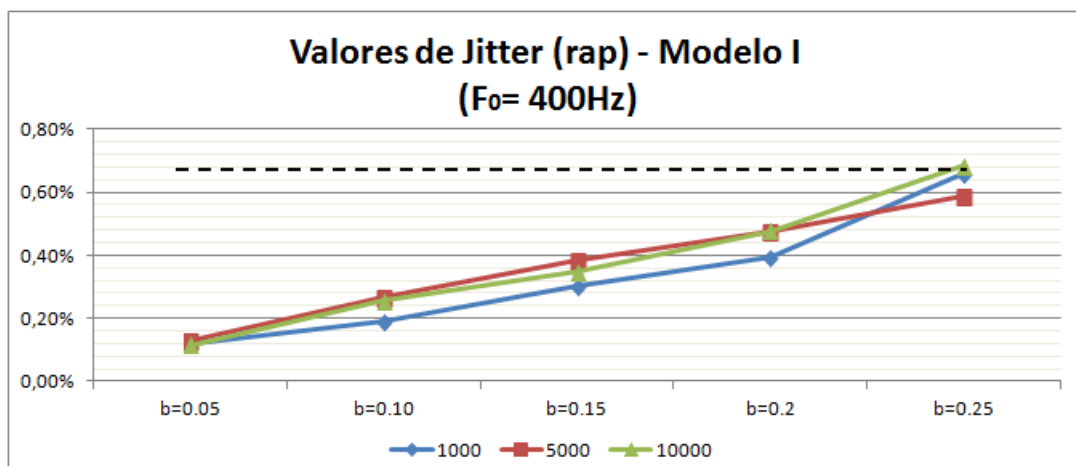


Figura 9.11: Gráfico com os valores de *rap* para $f_0 = 400\text{Hz}$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *ppq5* dentro da faixa considerado normal foram atingidos para todos os valores assumidos de *b*, como pode-se notar na Fig.(9.12).

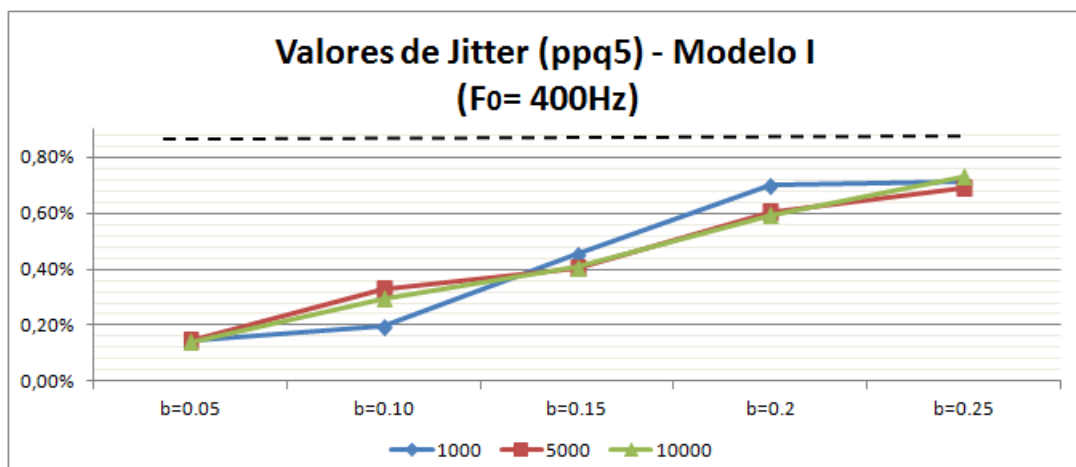


Figura 9.12: Gráfico com os valores de *ppq5* para $f_0 = 400\text{Hz}$ do Modelo I, com intervalos de tempo 1000, 5000 e 10000.

Nota-se que assumindo maior frequência fundamental mais valores de *jitt*, *jitta*, *rap* e *ppq5* encontram-se dentro da faixa considerada normal para uma maior variação de valores de *b*.

Tabela 9.5: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *ppq5* dentro da faixa considerada normal no Modelo I.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	6			11			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	-	-	-	x	x	x	x	x	x
b=0.2	-	-	-	x	-	x	x	x	x
b=0.25	-	-	-	-	-	-	x	x	x

Considerações

Nas simulações foram encontradas certas discrepâncias nos resultados obtidos de uma mesma simulação. De quarenta e cinco simulações, nove delas apresentaram contrariedades nas medidas de Jitter, três em simulações com $f_0 = 110Hz$, três com $f_0 = 200Hz$ e três simulações com $f_0 = 400Hz$.

(i) $f_0 = 110Hz$:

- Para os intervalos de tempo 5000 e 10000 ambos com o valor de $b = 0.10$.
- Para o intervalo de tempo 5000 com $b = 0.15$.

(ii) $f_0 = 200Hz$:

- Para todos os intervalos de tempo com $b = 0.20$.

(iii) $f_0 = 400Hz$:

- Para todos intervalos de tempo com $b = 0.25$.

Modelo II

(i) $f_0 = 110Hz$; intervalos de tempo 5000 e 10000:

- $a_1 = 0.25$:** - Os valores de *jitt* foram normais com apenas o valor de $b = 0.05$. Para resultados correspondentes a patologia ($\geq 1.04\%$) foram atingidos a partir do parâmetro $b \geq 0.10$, como podem ser observados na Fig.(9.13), valores acima da linha tracejada correspondem a patologia.

- Os valores de *jitta* foram normais com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias ($83.200\mu s$) foram alcançados a partir do parâmetro $b \geq 0.10$, como podem ser observados na Fig.(9.14), os valores acima da linha tracejada apresentam patologia.
- Os valores de *rap* foram normais para valores de b entre 0.05 e 0.10. Foram atingidos valores que detectam patologias (0.680%) com $b = 0.25$, como podem ser observados acima da linha tracejada da Fig.(9.15).
- Os valores de *ppq5* foram normais com valores de b 0.05 e 0.10. Os valores dentro do limite para detectar patologias (0.840%) foram obtidos com valores de $b = 0.25$, como pode ser observado na Fig.(9.16), valores acima da linha tracejada apresentam patologia.

- $a_1 = 0.5$:**
- Os valores de *jitt* foram normais com apenas $b = 0.05$. Para resultados dentro do limite para detectar patologias (1.04%) foram atingidos a partir de $b \geq 0.10$, como podem ser observados na Fig.(9.13), os valores acima da linha tracejada apresentam patologia.
 - As medidas de *jitta* foram encontrados para todos os valores de b dentro do limite para detectar patologias ($83.200\mu s$), como observa-se na Fig.(9.14).
 - Os valores de *rap* foram normais com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias (0.680%) foram atingidos a partir $b \geq 0.10$, como podem ser observados na Fig.(9.15), os valores acima da linha tracejada apresentam patologia.
 - Os valores de *ppq5* dentro do limite normal foram alcançados com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias (0.840%) foram obtidos com $b \geq 0.10$, como podem ser observados acima da linha tracejada na Fig.(9.16).

- $a_1 = 0.8$:**
- Foram obtidas em todas as medidas de Jitter (*jitt*, *jitta*, *rap* e *ppq5*), valores dentro do limite para detectar patologias para todos os valores de b , como pode-se observar nas Fig.(9.13), Fig.(9.14), Fig.(9.15) e Fig.(9.16).

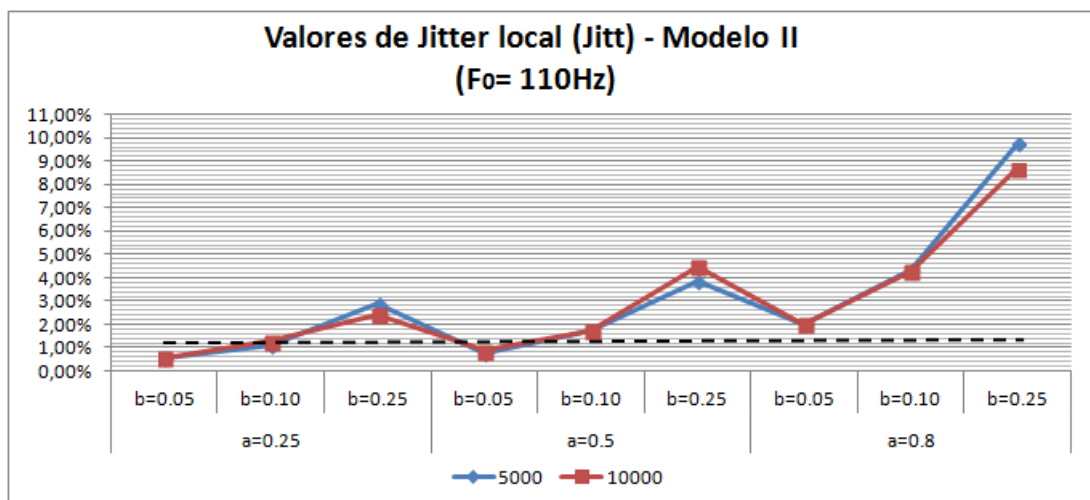


Figura 9.13: Gráfico com os valores de *jitt* para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

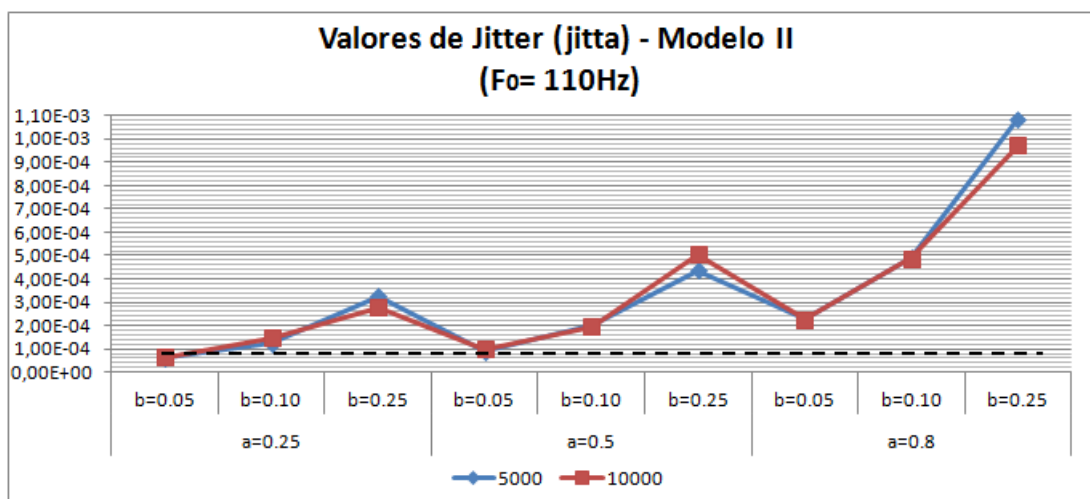


Figura 9.14: Gráfico com os valores de *jitta* para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

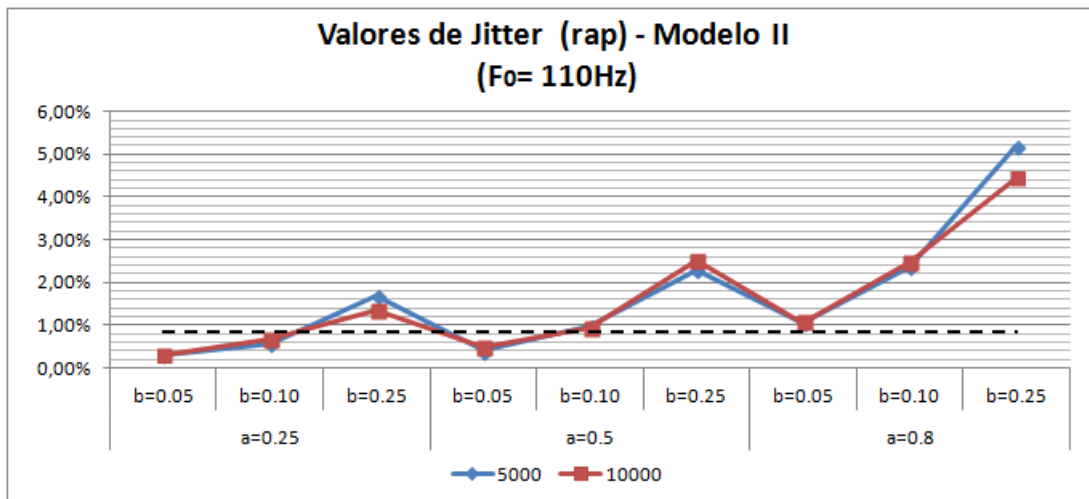


Figura 9.15: Gráfico com os valores de *rap* para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000, 5000 e 10000.

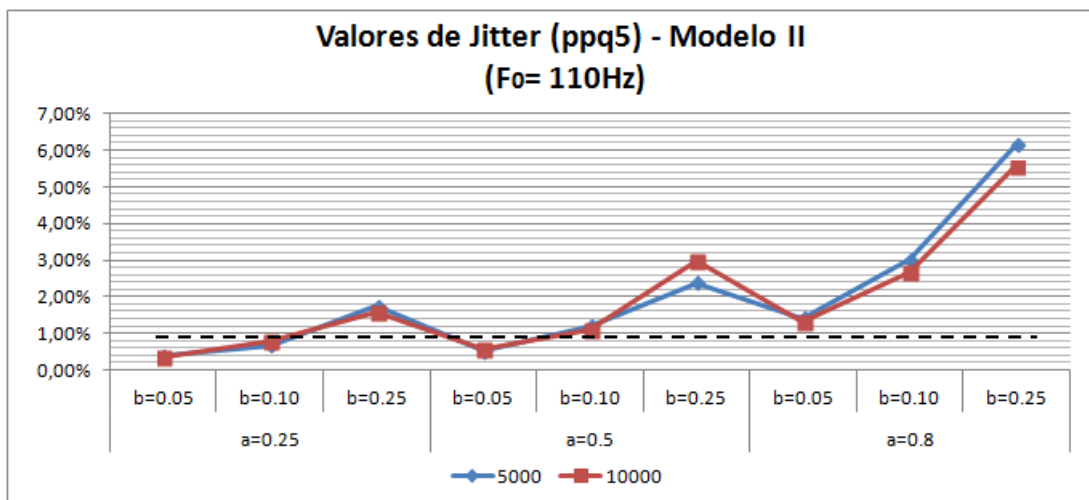


Figura 9.16: Gráfico com os valores de *ppq5* para $f_0 = 110Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

(ii) $f_0 = 200Hz$; intervalos de tempo 5000 e 10000:

$a_1 = 0.25$: - Foram atingidas em todas as medidas de Jitter (*jitt*, *jitta*, *rap* e *ppq5*) valores de voz normal com valores de *b* entre 0.05 e 0.10. Os valores dentro do limite para detectar patologias foram alcançados com valor de $b = 0.25$, como podem ser observado nas Fig.(9.17), Fig.(9.18), Fig.(9.19) e Fig.(9.20).

- $a_1 = 0.5$:**
- Os valores de *jitt* foram obtidos dentro do limite normal com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias (1.04%) foram alcançados a partir de $b \geq 0.10$, como podem ser observados na Fig.(9.17), os valores acima da linha tracejada corresponde a patologia.
 - Os valores de *jitta* foram normais para valores de b 0.05 e 0.10. Os valores dentro do limite para detectar patologias ($83.200\mu s$) foram obtidos com valores de $b = 0.25$, como pode ser observado acima da linha tracejada na Fig.(9.18).
 - As medidas de *rap* foram normais para valores de b 0.05 e 0.10, exceto para o intervalo de tempo fixado em 5000 e $b = 0.10$ que apresenta um valor dentro da faixa do limite para detectar patologias (0,680%). Foram atingidos valores correspondentes a patologia com $b = 0.25$, como pode-se observar na Fig.(9.19).
 - Os valores de *ppq5* foram normais para valores de b 0.05 e 0.10. Os valores dentro do limite para detectar patologias (0.840%) foram obtidos com $b = 0.25$, como pode ser observado na Fig.(9.18), valores acima da linha tracejada correspondem a patologia.
- $a_1 = 0.8$:**
- As medidas de *jitt* dentro do limite para detectar patologias (1.04%) para todos os valores de b , como observa-se na Fig.(9.17).
 - Os valores de *jitta* dentro do limite normal com apenas o valor de $b = 0.05$, enquanto que com $b \geq 0.10$ foram obtidos valores que detectam patologias ($83.200\mu s$), como mostra-se na Fig.(9.18), os os valores acima da linha tracejada correspondem a patologia.
 - As medidas de *rap* foram normais para valores de b 0.05 e 0.10, exceto para o intervalo de tempo fixado em 10000 e $b = 0.10$. Foram alcançados valores que detectam patologia com $b = 0.25$, como pode-se observar na Fig.(9.19).
 - Os valores de *ppq5* foram obtidos dentro do limite normal com apenas o valor de $b = 0.05$, enquanto que para resultados dentro do limite para detectar patologias (0.840%) foram atingidos com $b \geq 0.10$, como pode ser observados acima da linha tracejada na Fig.(9.20).

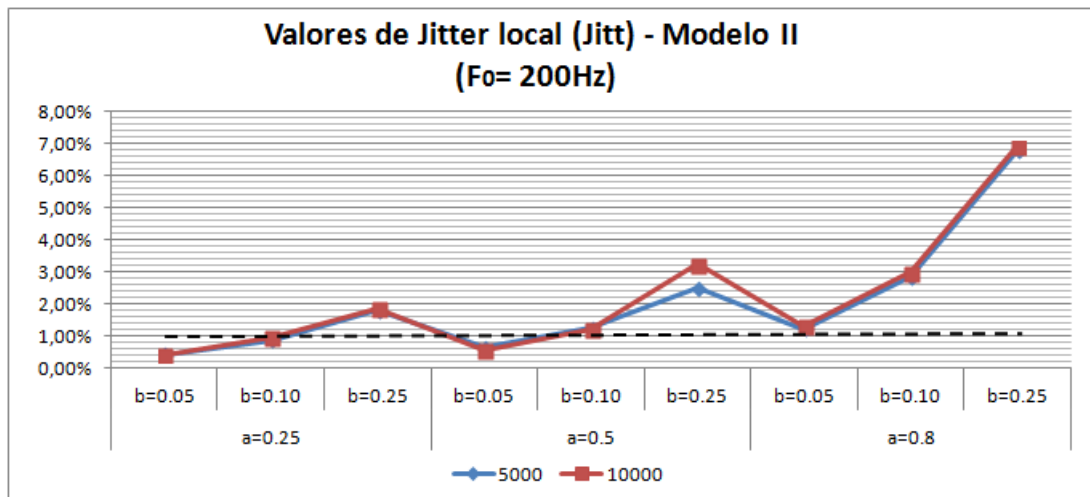


Figura 9.17: Gráfico com os valores de $jitt$ para $f_0 = 200\text{Hz}$ do Modelo II, com intervalos de tempo 1000 e 5000.

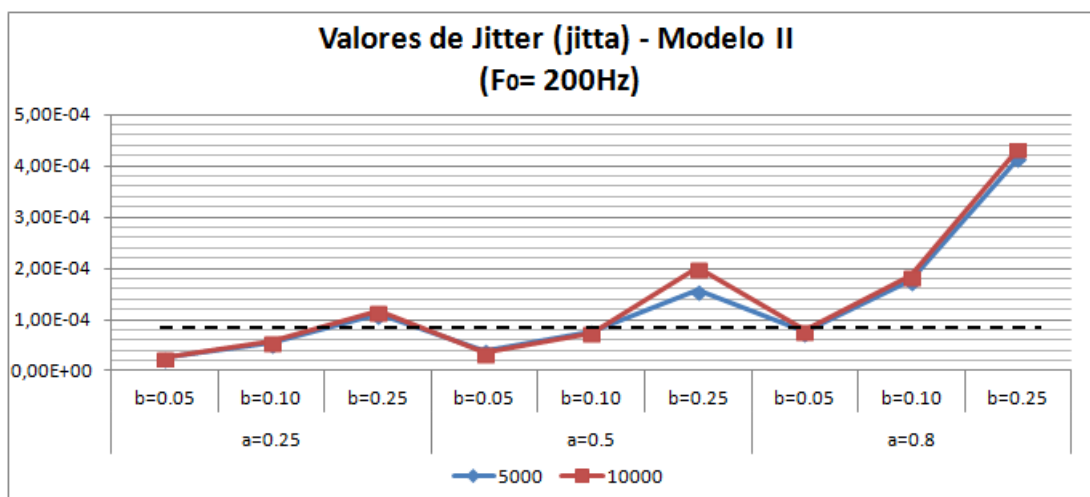


Figura 9.18: Gráfico com os valores de $jitta$ para $f_0 = 200\text{Hz}$ do Modelo II, com intervalos de tempo 1000 e 5000.

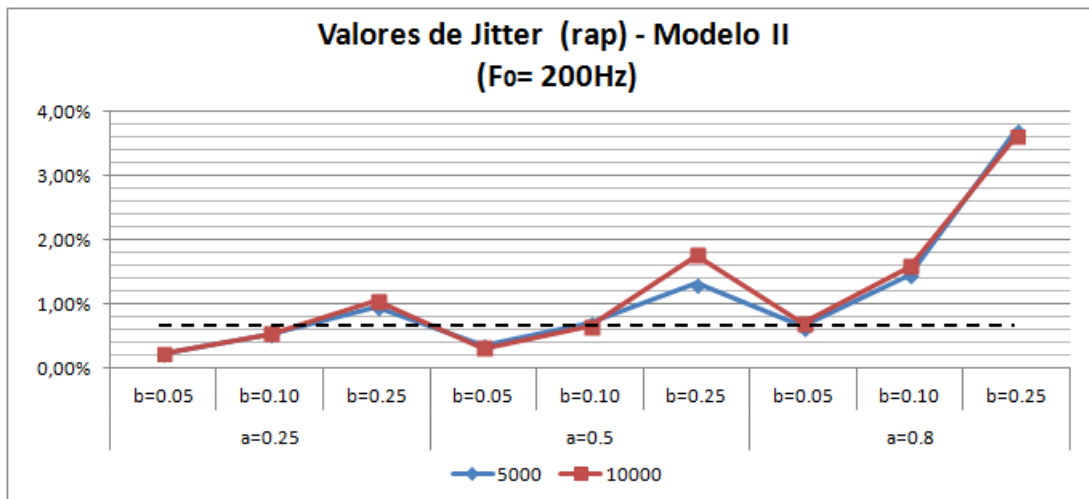


Figura 9.19: Gráfico com os valores de *rap* para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

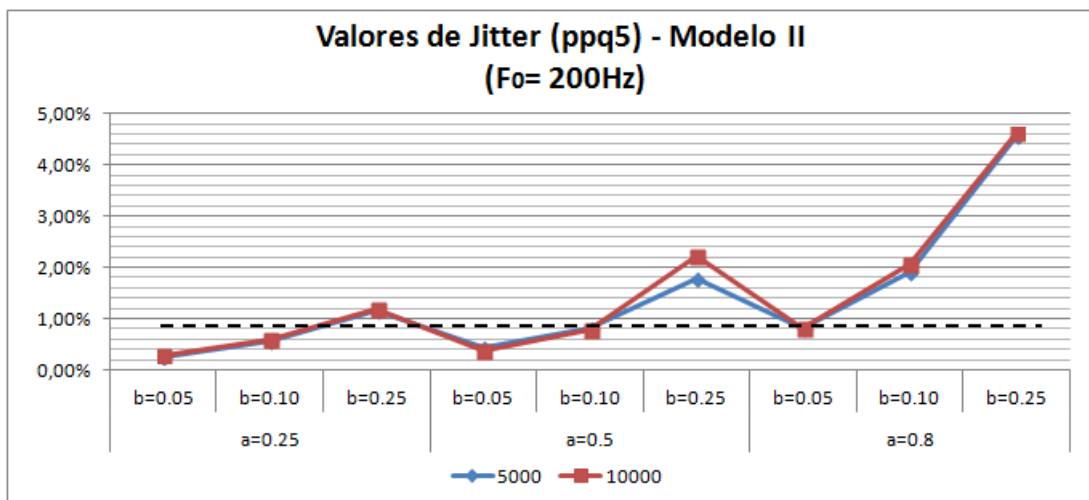


Figura 9.20: Gráfico com os valores de *ppq5* para $f_0 = 200Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

(i) $f_0 = 400Hz$; intervalos de tempo 5000 e 10000:

- $a_1 = 0.25$:**
- Os valores de *jitt* foram normais para valores de b entre 0.05 e 0.10. Para $b \geq 0.25$ foram obtidos valores que detectam patologia (1.04%), como podem ser observados na Fig.(9.21).
 - Os valores de *jitta* foram atingidos dentro do limite para detectar patologias ($83.200\mu s$) para todos os valores de b , como observa-se na Fig.(9.22).

- As medidas de *rap* foram normais para valores de b entre 0.05 e 0.10. Enquanto que com $b \geq 0.25$ foram obtidos valores que detecta patologia (0.680%), como pode-se observar na Fig.(9.23).
 - Os valores de *ppq5* foram alcançados dentro do limite normal com valores de b 0.05 e 0.10. Para $b \geq 0.25$ foram obtidos valores que detectam patologia (0.840%), como podem ser observados na Fig.(9.22), valores acima da linha correspondem a patologia.
- $a_1 = 0.5$:**
- Os valores de *jitt* foram normais para valores de b entre 0.05 e 0.10, enquanto com $b = 0.25$ e acima desse valor, foram obtidos dentro do limite que detecta patologia (1.04%), como podem ser observados na Fig.(9.21), valores acima da linha tracejada apresentam patologia.
 - Os valores de *jitta* foram obtidos dentro do limite para detectar patologias ($83.200\mu s$) para todos os valores de b , como observa-se na Fig.(9.22).
 - As medidas de *rap* dentro da faixa considerada normal com valores para b entre 0.05 e 0.10. Enquanto que com $b \geq 0.25$ foram obtidos valores que detecta patologia (0.680%), como pode-se observar na Fig.(9.23).
 - Os valores de *ppq5* foram normais para valores de b entre 0.05 e 0.10. Foram obtidos valores correspondente a patologia com $b = 0.25$, como pode ser observado na Fig.(9.22).
- $a_1 = 0.8$:**
- Os valores de *jitt* foram normais com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias (1.04%) foram atingidos a partir do parâmetro b fixado em 0.10 e em valores superiores a este valor, como podem ser observados na Fig.(9.21).
 - Os valores de *jitta* dentro do limite normal foram obtidos com valores para b entre 0.05 e 0.10. Enquanto que valores correspondentes a patologia foram alcançados com $b = 0.25$, como pode ser observado na Fig.(9.22).
 - Os valores de *rap* dentro do limite considerado normal foram atingidos com apenas o valor de $b = 0.05$. Para resultados dentro do limite para detectar patologias (0.680%) foram alcançados com $b \geq 0.10$, como podem ser observados na Fig.(9.23).
 - Os valores de *ppq5* foram normais com apenas o valor de $b = 0.05$. Para $b \geq 0.10$ foram obtidos valores que detectam patologia (0.840%), como podem ser observados na Fig.(9.24).

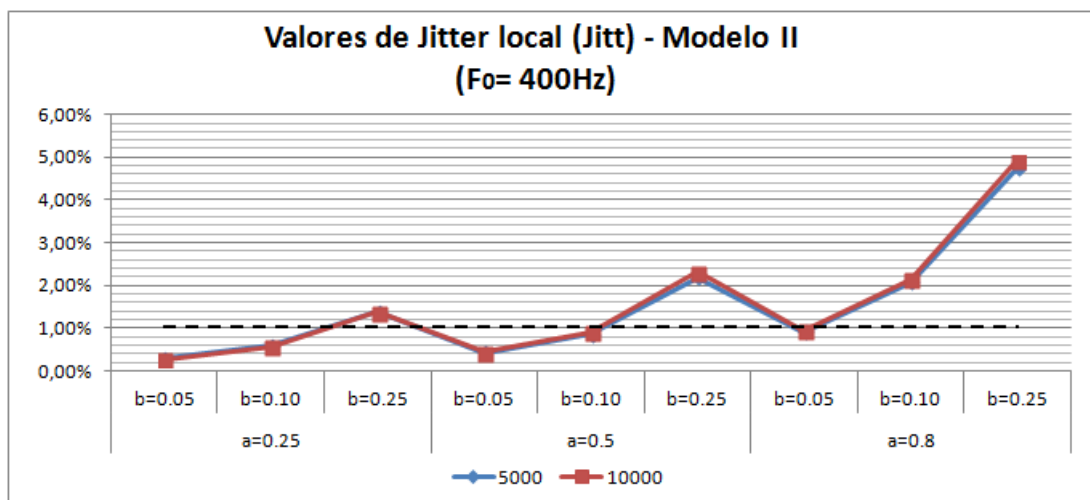


Figura 9.21: Gráfico com os valores de $jitt$ para $f_0 = 400\text{Hz}$ do Modelo II, com intervalos de tempo 1000 e 5000.

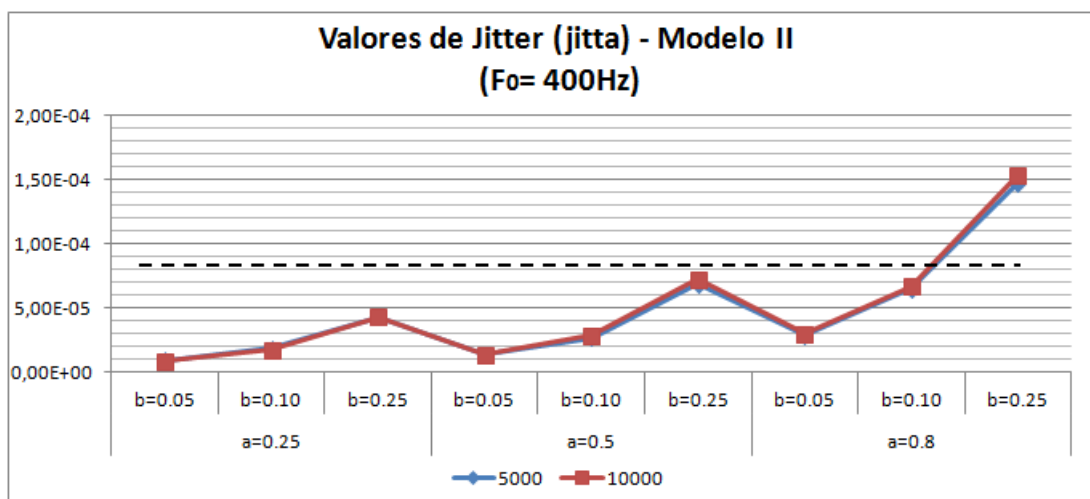


Figura 9.22: Gráfico com os valores de $jitta$ para $f_0 = 400\text{Hz}$ do Modelo II, com intervalos de tempo 1000 e 5000.

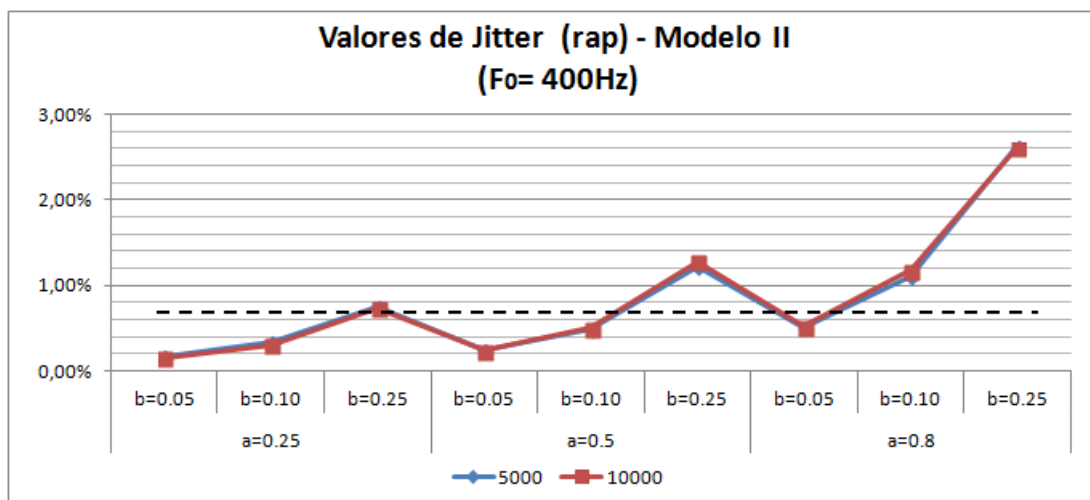


Figura 9.23: Gráfico com os valores de *rap* para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

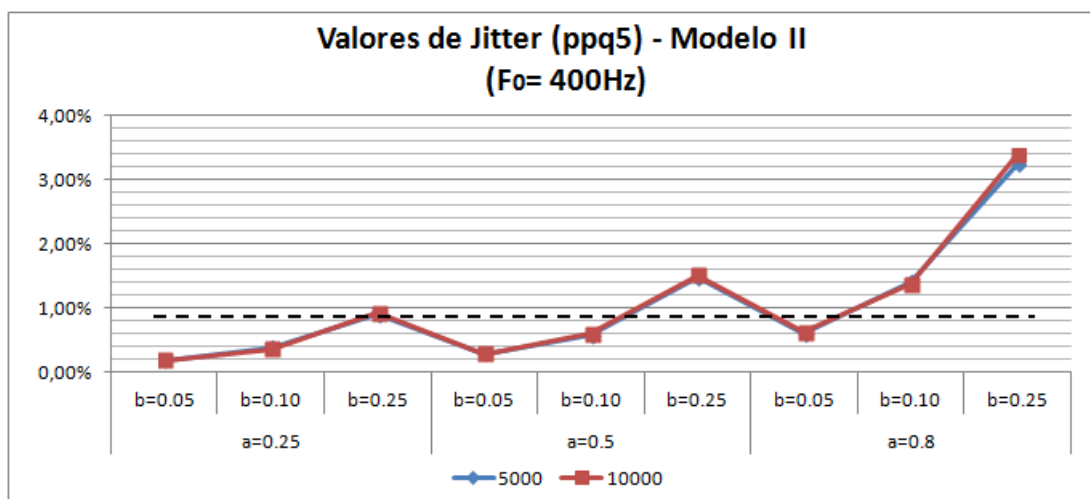


Figura 9.24: Gráfico com os valores de *ppq5* para $f_0 = 400Hz$ do Modelo II, com intervalos de tempo 1000 e 5000.

Nota-se que assumindo maior frequência fundamental mais valores de *jitt*, *jitta*, *rap* e *ppq5* encontram-se dentro da faixa normal para uma maior variação de valores de a_1 para o parâmetro b .

Tabela 9.6: Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de *jitt* dentro da faixa considerada normal no Modelo II.

		$f_0 = 110Hz$		$f_0 = 200Hz$		$f_0 = 400Hz$	
		5000	10000	5000	10000	5000	10000
		4		6		10	
b=0.05	a=0.25	x	x	x	x	x	x
	a=0.5	x	x	x	x	x	x
	a=0.8	-	-	-	-	x	x
b=0.10	a=0.25	-	-	x	x	x	x
	a=0.5	-	-	-	-	x	x

Tabela 9.7: Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de *jitta* dentro da faixa considerada normal no Modelo II.

		$f_0 = 110Hz$		$f_0 = 200Hz$		$f_0 = 400Hz$	
		5000	10000	5000	10000	5000	10000
		2		10		16	
b=0.05	a=0.25	x	x	x	x	x	x
	a=0.5	-	-	x	x	x	x
	a=0.8	-	-	x	x	x	x
b=0.10	a=0.25	-	-	x	x	x	x
	a=0.5	-	-	x	x	x	x
	a=0.8	-	-	-	-	x	x
b=0.25	a=0.25	-	-	-	-	x	x
	a=0.5	-	-	-	-	x	x

Tabela 9.8: Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de *rap* dentro da faixa considerada normal no Modelo II.

		$f_0 = 110Hz$		$f_0 = 200Hz$		$f_0 = 400Hz$	
		5000	10000	5000	10000	5000	10000
		6		8		10	
b=0.05	a=0.25	x	x	x	x	x	x
	a=0.5	x	x	x	x	x	x
	a=0.8	-	-	x	-	x	x
b=0.10	a=0.25	x	x	x	x	x	x
	a=0.5	-	-	-	x	x	x

Tabela 9.9: Distribuição da quantidade de valores de b e a_1 para cada valor de frequência fundamental dos resultados obtidos de *ppq5* dentro da faixa considerada normal no Modelo II.

		$f_0 = 110Hz$		$f_0 = 200Hz$		$f_0 = 400Hz$	
		5000	10000	5000	10000	5000	10000
		6		10		10	
b=0.05	a=0.25	x	x	x	x	x	x
	a=0.5	x	x	x	x	x	x
	a=0.8	-	-	x	x	x	x
b=0.10	a=0.25	x	x	x	x	x	x
	a=0.5	-	-	x	x	x	x

Considerações

Nas simulações foram encontrados certas discrepâncias nos resultados obtidos de uma mesma simulação, de cinquenta e quatro simulações, catorze delas apresentaram contrariedades nas medidas de Jitter, quatro em simulações com $f_0 = 110Hz$, quatro com $f_0 = 200Hz$ e seis simulações com $f_0 = 400Hz$.

(i) $f_0 = 110Hz$:

- Para todos intervalos de tempo com $b = 0.10$ e $a_1 = 0.25$.
- Para todos intervalos de tempo com $b = 0.05$ e $a_1 = 0.5$.

(ii) $f_0 = 200Hz$:

- Para todos intervalos de tempo com $b = 0.10$ e $a_1 = 0.5$.
- Para todos intervalos de tempo com $b = 0.05$ e $a_1 = 0.8$.

(iii) $f_0 = 400Hz$:

- Para todos os intervalos de tempo com $b = 0.25$ e $a_1 = 0.25, 0.5$.
- Para todos intervalos de tempo com $b = 0.10$ e $a_1 = 0.8$.

Modelo III

(i) $f_0 = 110Hz$; intervalos de tempo 1000, 5000 e 10000:

- Os valores de *jitt* foram normais com valores de b entre 0.05, 0.10 e 0.15. Foram obtidos dentro do limite para detectar patologias (1.04%) com $b \geq 0.2$, como podem ser observados na Fig.(9.25), os valores acima da linha tracejada representam patologia.

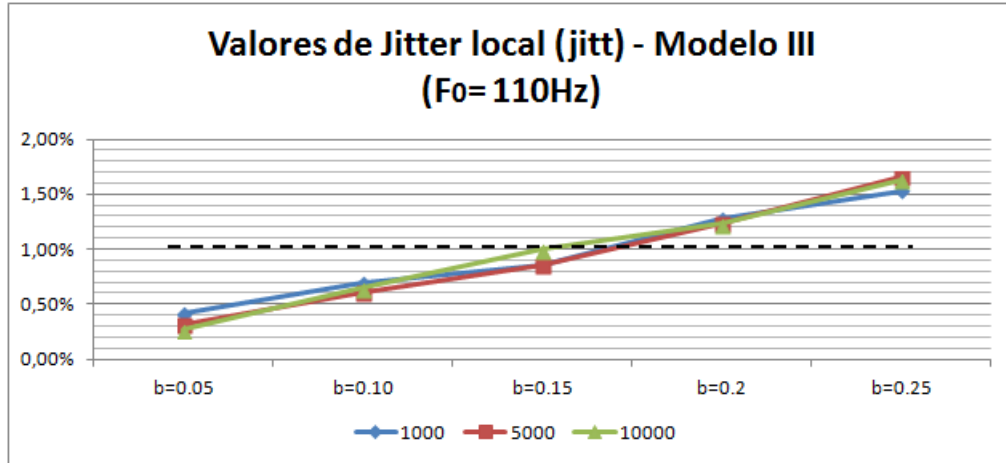


Figura 9.25: Gráfico com os valores de *jitt* para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *jitta* dentro do limite para detectar patologias ($83.200\mu s$) foram atingidos com $b \geq 0.15$. Com valores de b entre 0.05 e 0.10 foram alcançados valores normais, como podem ser observados na Fig.(9.26).

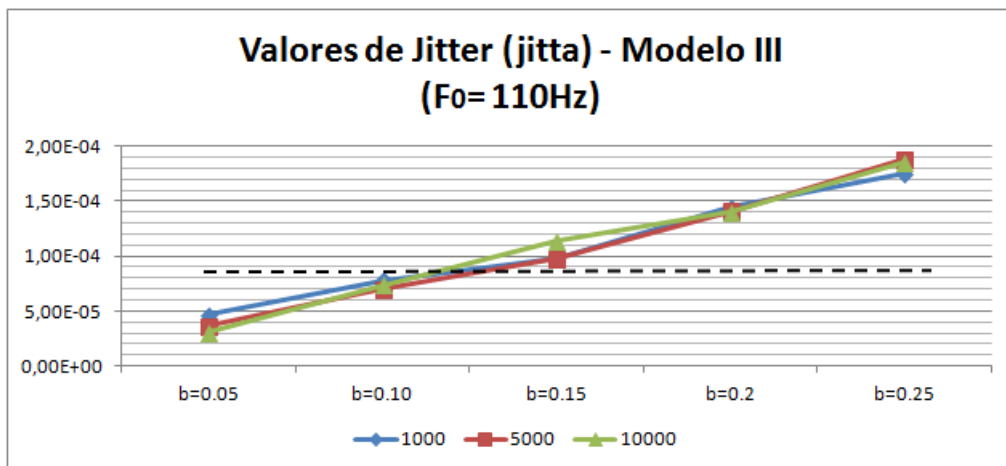


Figura 9.26: Gráfico com os valores de *jitta* para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *rap* correspondentes a patologias (0.680%) foram obtidos com $b \geq 0.2$, enquanto que para resultados dentro da faixa normal foram alcança-

dos com valores de b entre 0.05, 0.10 e 0.15, como podem ser observados na Fig.(9.27).

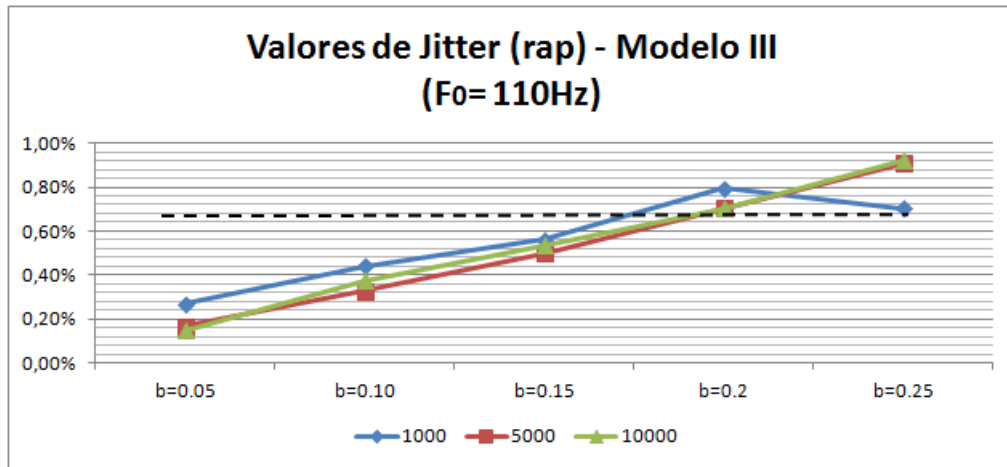


Figura 9.27: Gráfico com os valores de rap para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de $ppq5$ foram normais com valores de b entre 0.05, 0.10, 0.15 e 0.2, exceto para o intervalo de tempo fixado em 10000 e $b = 0.2$. Para $b = 0.25$ foram atingidos valores que detectam patologia (0.840%), como podem ser observados os valores acima da linha tracejada na Fig.(9.28) representam patologia.

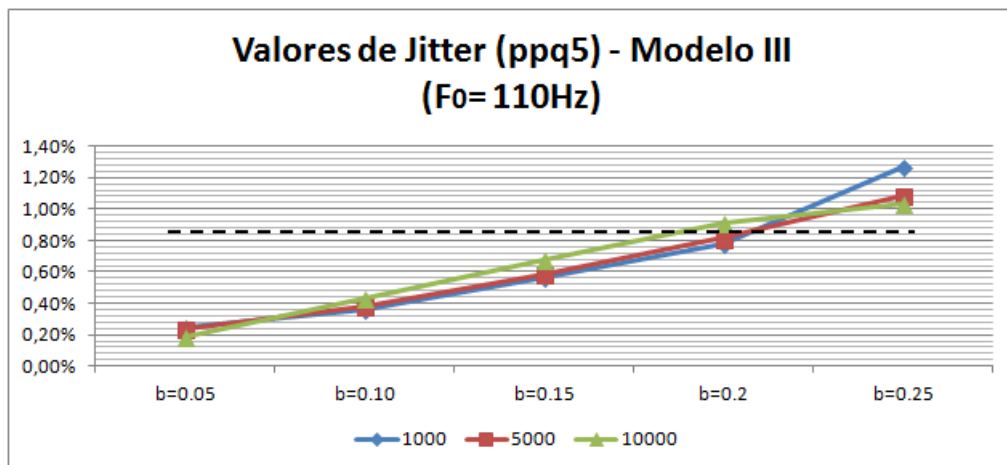


Figura 9.28: Gráfico com os valores de $ppq5$ para $f_0 = 110Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

(ii) $f_0 = 200Hz$; intervalos de tempo 1000, 5000 e 10000:

- As medidas de *jitt* normais foram obtidos com valores de b entre 0.05, 0.10 e 0.15. Com ($b \geq 0.2$) foram alcançados valores correspondentes a patologia, exceto para os intervalos de tempo 5000 e 10000 com $b = 0.25$ e para o intervalo de tempo 1000 com $b = 0.2$ que encontram dentro da faixa normal como pode ser observado na Fig.(9.29).

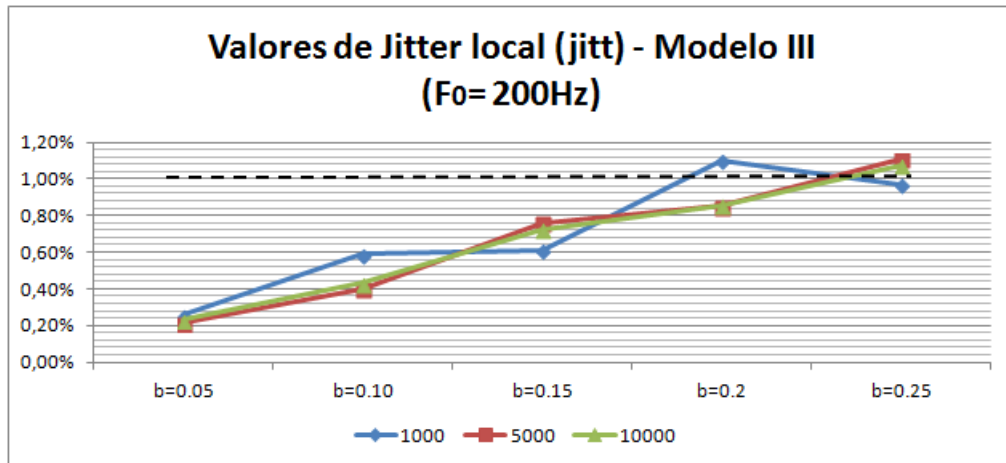


Figura 9.29: Gráfico com os valores de *jitt* para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

- Todos os valores de *jitta*, *rap ppq5* foram normais para todos os valores de b , como podem ser observados na Fig.(9.30), Fig.(9.31) e na Fig.(9.32) respectivamente.

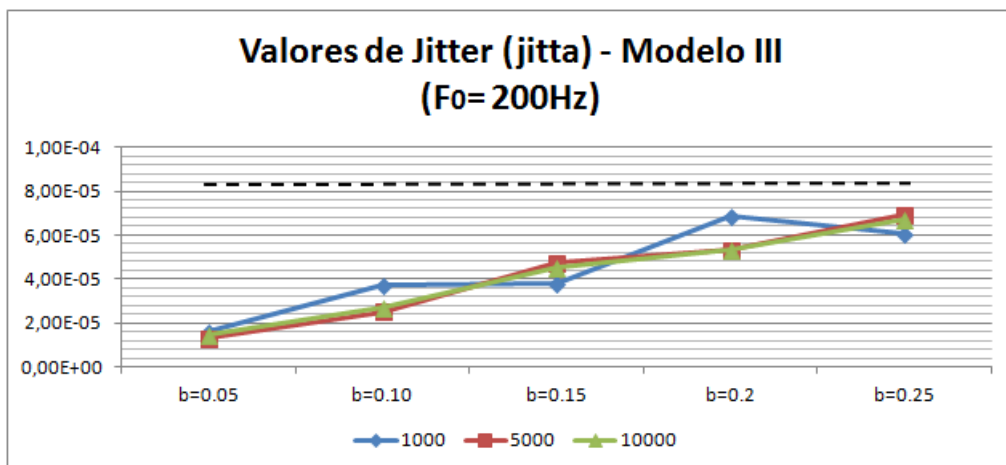


Figura 9.30: Gráfico com os valores de *jitta* para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

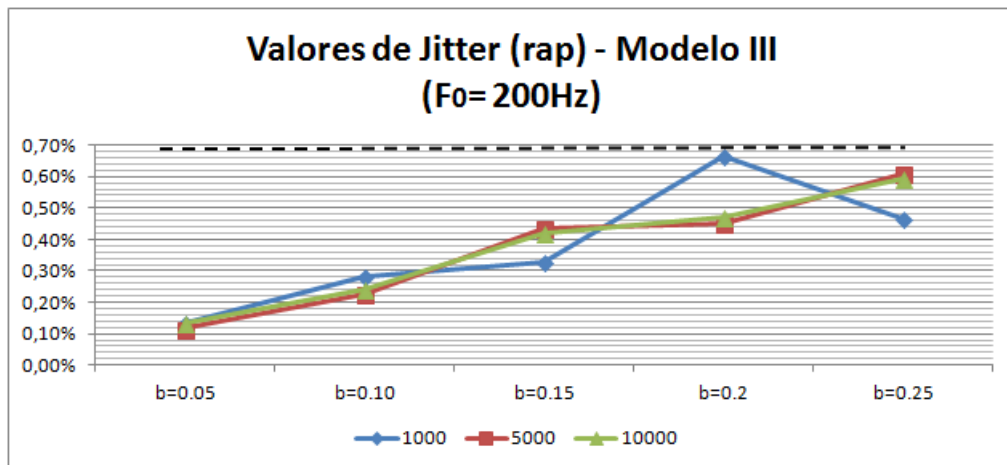


Figura 9.31: Gráfico com os valores de *rap* para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

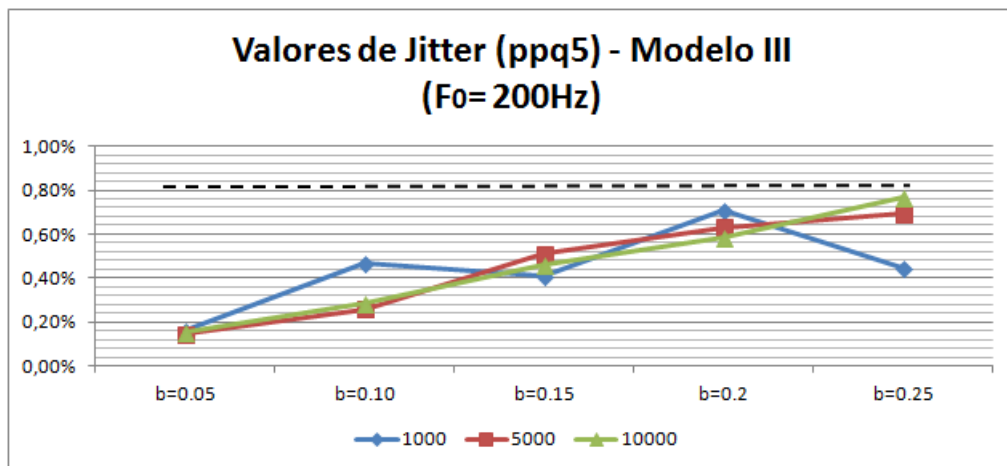


Figura 9.32: Gráfico com os valores de *ppq5* para $f_0 = 200Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

(iii) $f_0 = 400Hz$; intervalos de tempo 1000, 5000 e 10000:

- Todos os valores de *jitt*, *jitta*, *rap* *ppq5* foram normais, como podem ser observados na Fig.(9.33), Fig.(9.34), Fig.(9.35) e na Fig.(9.36) respectivamente.

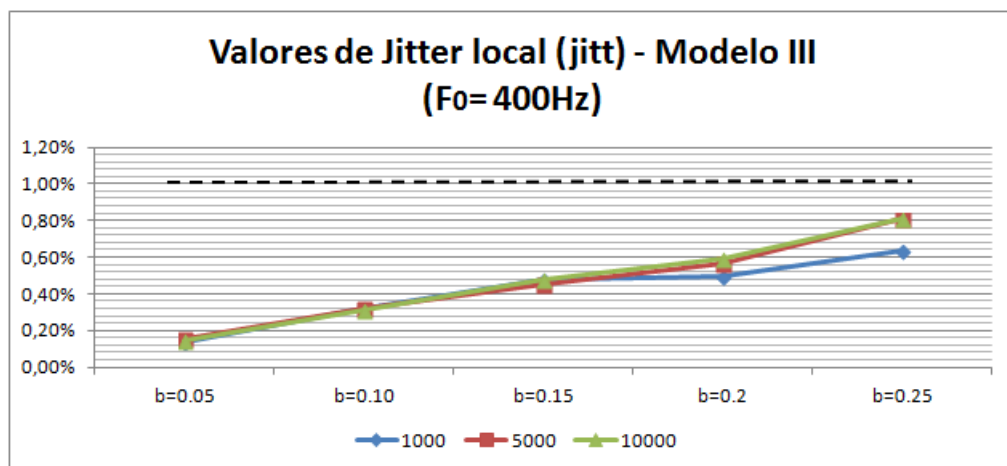


Figura 9.33: Gráfico com os valores de *jitt* para $f_0 = 400\text{Hz}$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

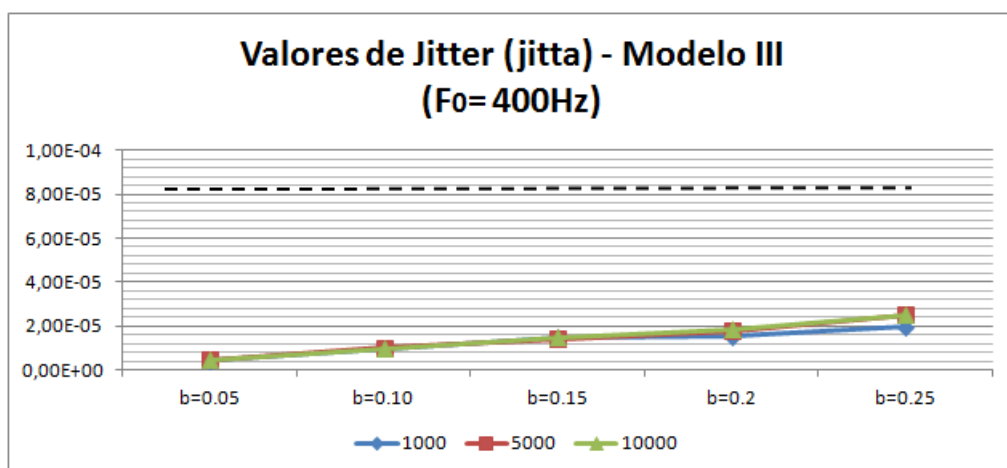


Figura 9.34: Gráfico com os valores de *jitta* para $f_0 = 400\text{Hz}$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

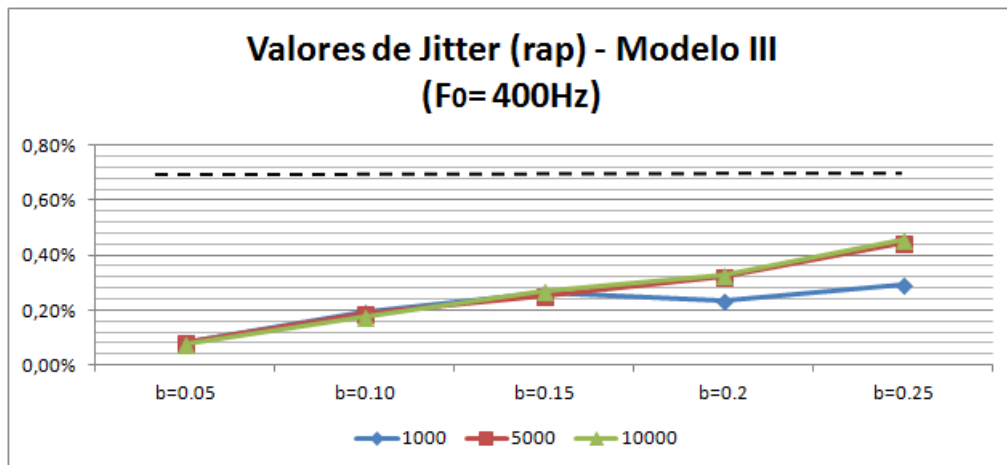


Figura 9.35: Gráfico com os valores de *rap* para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

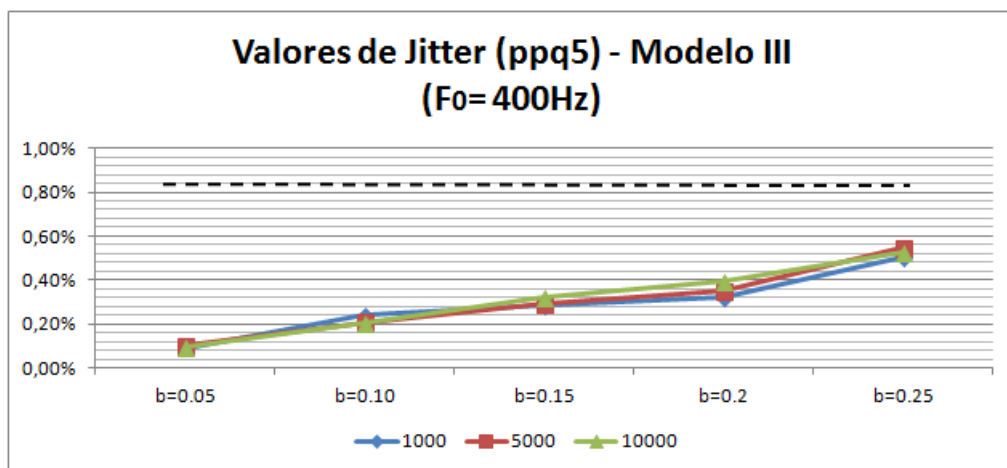


Figura 9.36: Gráfico com os valores de *ppq5* para $f_0 = 400Hz$ do Modelo III, com intervalos de tempo 1000, 5000 e 10000.

Nota-se que assumindo maior frequência fundamental mais valores de *jitt* encontram-se dentro da faixa normal com uma faixa maior de valores de *b*, como ilustra-se na Tab.(9.10), no entanto para as outras medidas *jitta*, *rap* e *ppq5* foram normais para todos os valores de *b*, pelo que só observou-se um incremento de valores ao aumentar da $f_0 = 110Hz$ a $f_0 = 200Hz$.

Tabela 9.10: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *jitt* dentro da faixa considerada normal no Modelo III.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	9			12			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	x	x	x	x	x	x	x	x	x
b=0.2	-	-	-	-	x	x	x	x	x
b=0.25	-	-	-	x	-	-	x	x	x

Tabela 9.11: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *jitta* dentro da faixa considerada normal no Modelo III.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	6			15			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	-	-	-	x	x	x	x	x	x
b=0.2	-	-	-	x	x	x	x	x	x
b=0.25	-	-	-	x	x	x	x	x	x

Tabela 9.12: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *rap* dentro da faixa considerada normal no Modelo III.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	9			15			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	x	x	x	x	x	x	x	x	x
b=0.2	-	-	-	x	x	x	x	x	x
b=0.25	-	-	-	x	x	x	x	x	x

Tabela 9.13: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *ppq5* dentro da faixa considerada normal no Modelo III.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	11			15			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	x	x	x	x	x	x	x	x	x
b=0.2	x	x	-	x	x	x	x	x	x
b=0.25	-	-	-	x	x	x	x	x	x

Considerações

Nas simulações foram encontradas certas discrepâncias nos resultados obtidos de uma mesma simulação. De quarenta e cinco simulações, oito delas apresentaram contrariedades nas medidas de Jitter, cinco em simulações com $f_0 = 110Hz$ e três com $f_0 = 200Hz$.

(i) $f_0 = 110Hz$:

- Para todos intervalos de tempo com $b = 0.15$.
- Para os intervalos de tempo (1000 e 5000) com $b = 0.2$.

(ii) $f_0 = 200Hz$:

- Para o intervalo de tempo (1000) com $b = 0.2$.
- Para os intervalos de tempo 5000 e 10000 com $b = 0.25$.

Modelo IV

(i) $f_0 = 110Hz$; intervalos de tempo 1000, 5000 e 10000:

- Os valores de *jitt* dentro do limite para detectar patologias (1.04%) foram obtidos com $b \geq 0.2$. Para valores de b entre 0.05, 0.10 e 0.15 foram normais, como podem ser observados na Fig.(9.37).

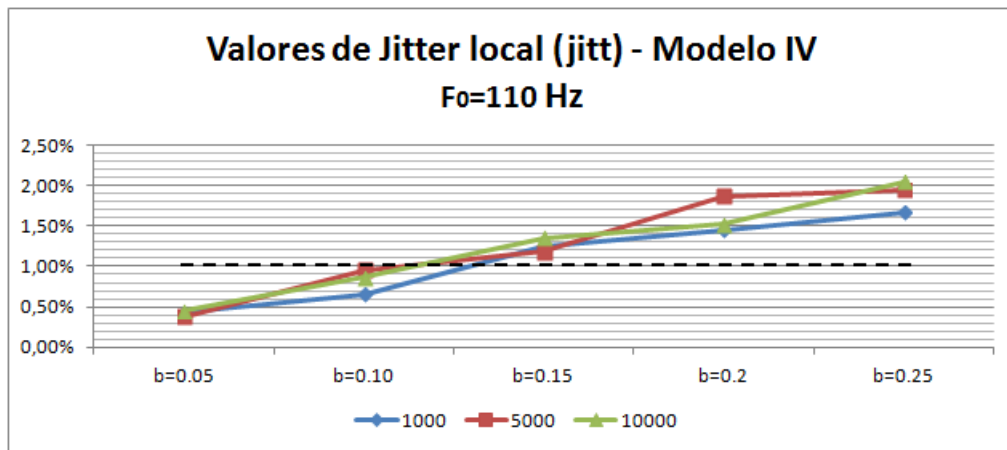


Figura 9.37: Gráfico com os valores de *jitt* para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *jitta* foram normais com valores de b entre 0,05 e 0,10. Para $b \geq 0,15$ foram obtidos valores correspondentes a patologia, como podem ser observados acima da linha tracejada na Fig.(9.38).

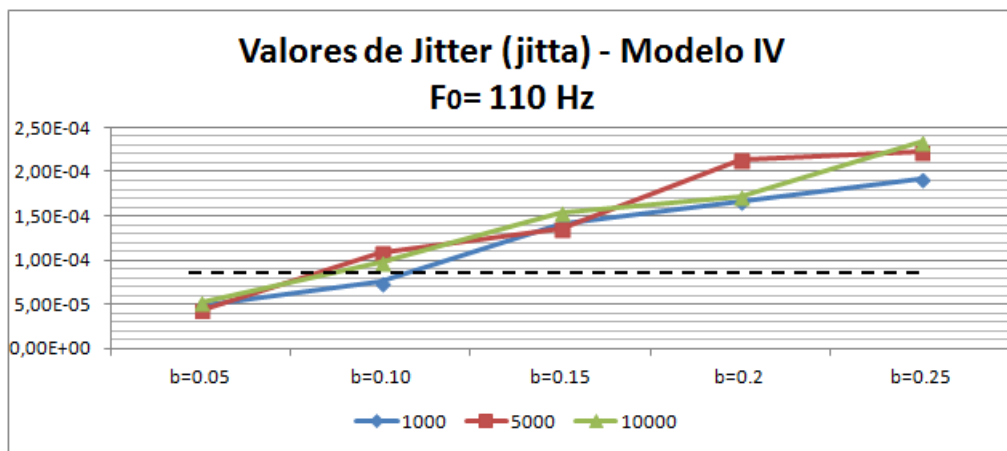


Figura 9.38: Gráfico com os valores de *jitta* para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

- As medidas de *rap* dentro do limite para detectar patologias (0,680%) foram obtidos com $b \geq 0,2$. Para valores de b entre 0,05, 0,10 e 0,15 foram normais, como podem ser observados na Fig.(9.39).

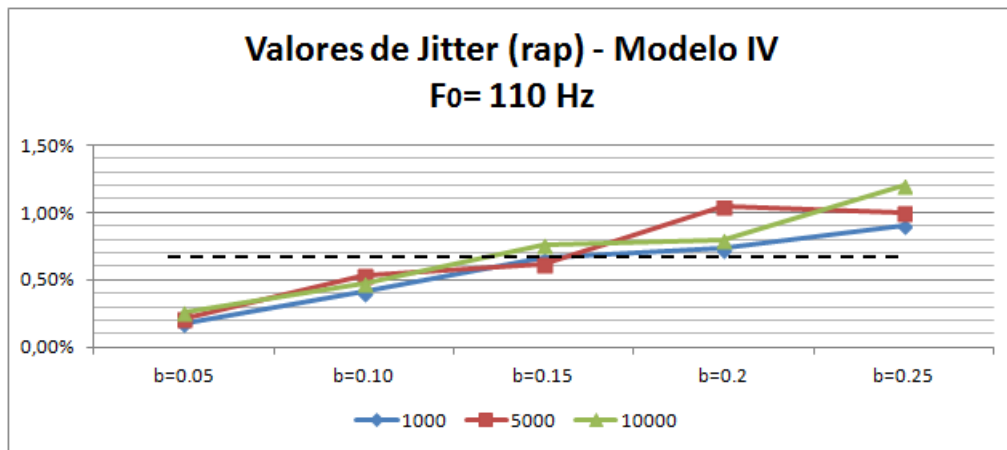


Figura 9.39: Gráfico com os valores de *rap* para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

- Os valores de *ppq5* foram normais com valores de b entre 0.05, 0.10, 0.15 e 0.2, exceto para $b = 0.2$ para o intervalo de tempo 10000 que corresponde a patologia. Para $b = 0.25$ foram obtidos valores que detectam patologia, como podem ser observados acima da linha tracejada na Fig.(9.40).

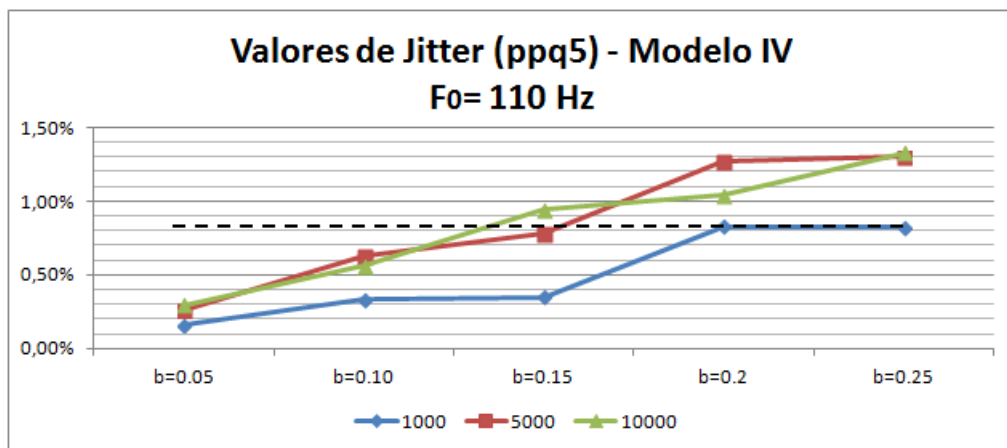


Figura 9.40: Gráfico com os valores de *ppq5* para $f_0 = 110Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

(ii) $f_0 = 200Hz$; intervalos de tempo 1000, 5000 e 10000:

- As medidas de *jitt* foram obtidas dentro do limite para detectar patologia com valores de $b = 0.2$ (para o intervalo de tempo 1000) e $b = 0.25$ (para os intervalos de tempo 5000 e 10000), os valores restantes foram normais, como podem ser observados na Fig.(9.41).

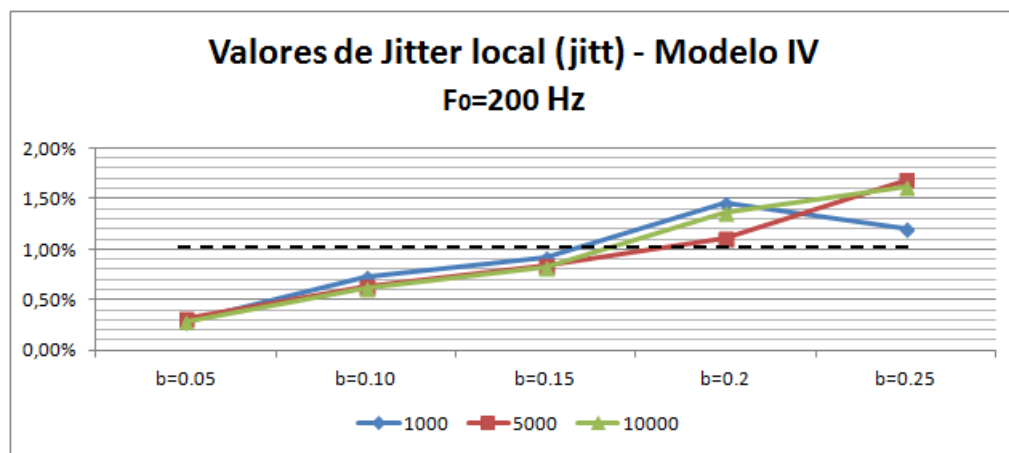


Figura 9.41: Gráfico com os valores de *jitt* para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

- Todos os valores de *jitta*, *rap* *ppq5* foram normais, como podem ser observados na Fig.(9.42), Fig.(9.42), Fig.(9.43) e na Fig.(9.44) respectivamente.

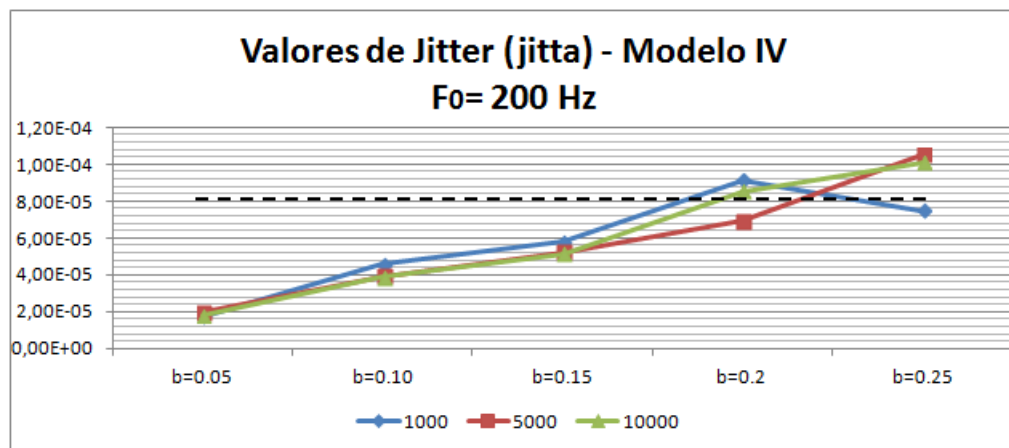


Figura 9.42: Gráfico com os valores de *jitta* para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

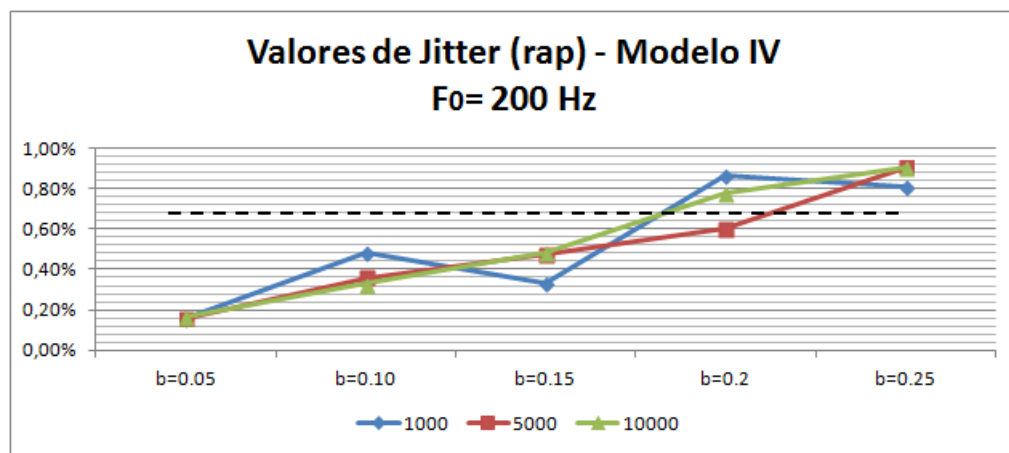


Figura 9.43: Gráfico com os valores de *rap* para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

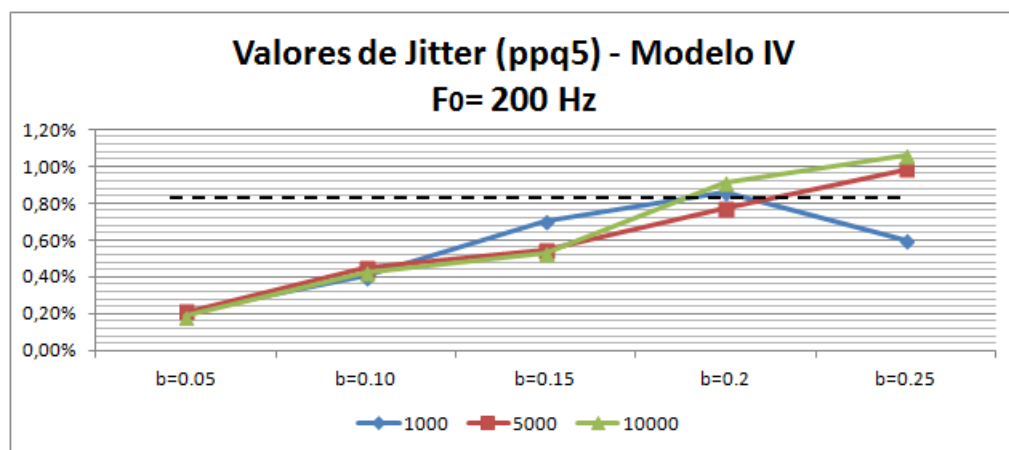


Figura 9.44: Gráfico com os valores de *ppq5* para $f_0 = 200Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

(iii) $f_0 = 400Hz$; intervalos de tempo 1000, 5000 e 10000:

- Todos os valores de *jitt*, *jitta*, *rap* *ppq5* foram normais, como podem ser observados na Fig.(9.33), Fig.(9.46), Fig.(9.47) e na Fig.(9.48) respectivamente.

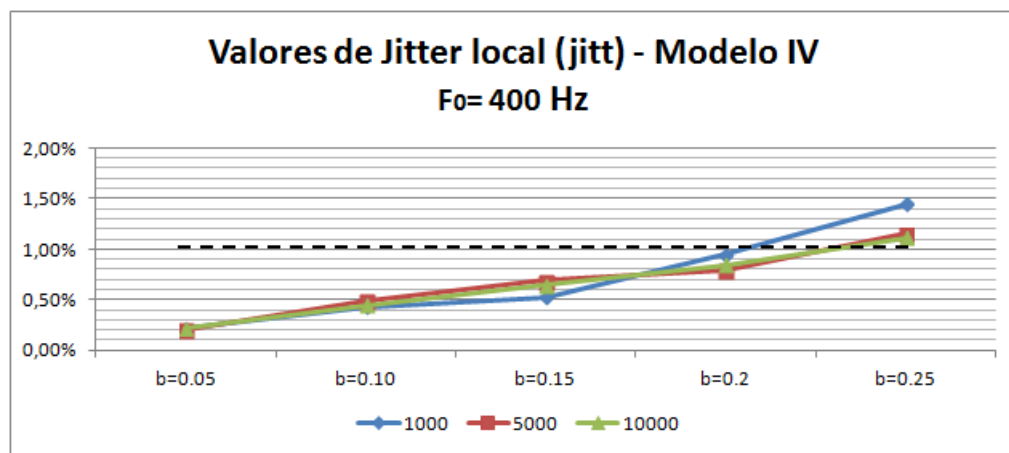


Figura 9.45: Gráfico com os valores de $jitt$ para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

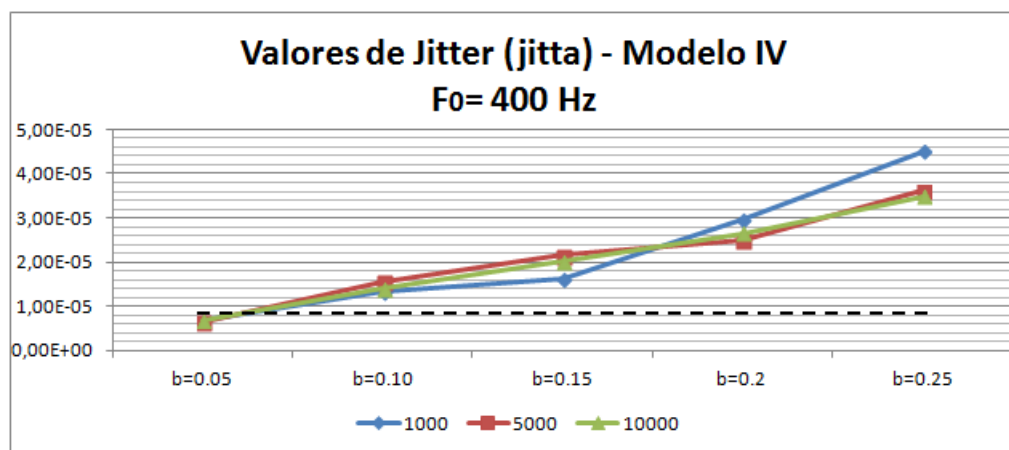


Figura 9.46: Gráfico com os valores de $jitta$ para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

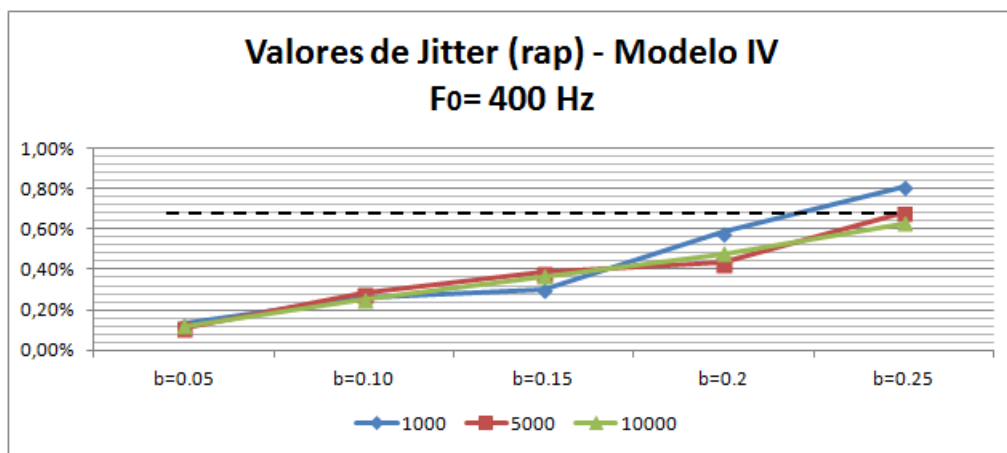


Figura 9.47: Gráfico com os valores de *rap* para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

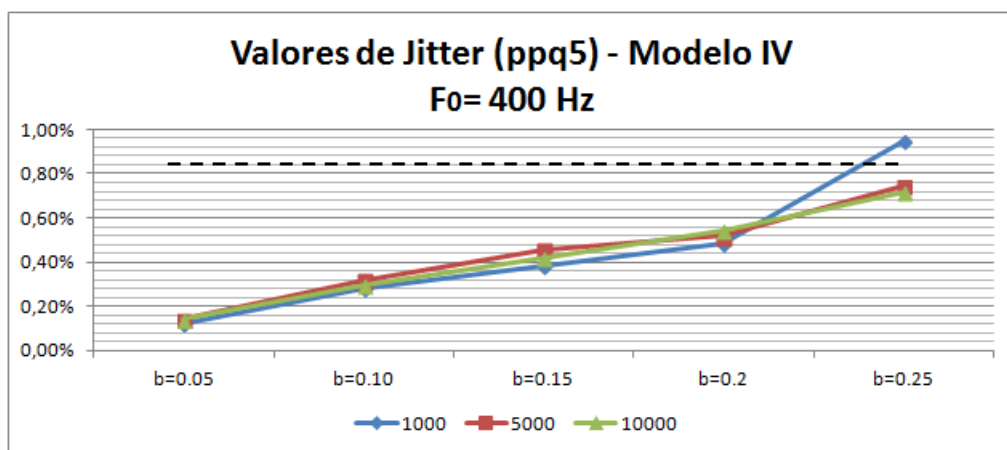


Figura 9.48: Gráfico com os valores de *ppq5* para $f_0 = 400Hz$ do Modelo IV, com intervalos de tempo 1000, 5000 e 10000.

Nota-se que assumindo maior frequência fundamental mais valores de *jitt*, *jitta* e *rap* encontram-se dentro da faixa considerada normal para uma maior variação de valores de *b*.

Tabela 9.14: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de $jitt$ dentro da faixa considerada normal no Modelo IV.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	6			9			12		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	-	-	-	x	x	x	x	x	x
b=0.2	-	-	-	-	-	-	x	x	x

Tabela 9.15: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *jitta* dentro da faixa considerada normal no Modelo IV.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	4			11			15		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	-	-	x	x	x	x	x	x
b=0.15	-	-	-	x	x	x	x	x	x
b=0.2	-	-	-	-	x	-	x	x	x
b=0.25	-	-	-	x	-	-	x	x	x

Tabela 9.16: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *rap* dentro da faixa considerada normal no Modelo IV.

[illegible]

Tabela 9.17: Distribuição da quantidade de valores de b para cada valor de frequência fundamental dos resultados obtidos de *ppq5* dentro da faixa considerada normal no Modelo IV.

	$f_0 = 110Hz$			$f_0 = 200Hz$			$f_0 = 400Hz$		
	1000	5000	10000	1000	5000	10000	1000	5000	10000
	10			11			14		
b=0.05	x	x	x	x	x	x	x	x	x
b=0.10	x	x	x	x	x	x	x	x	x
b=0.15	x	x	-	x	x	x	x	x	x
b=0.2	x	-	-	-	x	-	x	x	x
b=0.25	x	-	-	x	-	-	-	x	x

Considerações

Nas simulações foram encontradas certas discrepâncias nos resultados obtidos de uma mesma simulação. De quarenta e cinco simulações, onze delas apresentaram contrariedades nas medidas de Jitter, seis em simulações com $f_0 = 110Hz$, dois com $f_0 = 200Hz$ e três simulações com $f_0 = 400Hz$

(i) $f_0 = 110Hz$:

- Para os intervalos de tempo 5000 e 10000 com $b = 0.10$.
- Para os intervalos de tempo 1000 e 5000 com $b = 0.15$.
- Para o intervalo 1000 com $b = 0.2$ e $b = 0.25$.

(ii) $f_0 = 200Hz$:

- Para o intervalo de tempo 1000 com $b = 0.2$ e $b = 0.25$.

(iii) $f_0 = 400Hz$:

- Para todos os intervalos de tempo $b = 0.25$.

Modelo V

Analisando os resultados do modelo V para todos os intervalos de tempo (1000, 5000 e 10000) e todos os valores de frequência fundamental:

- Todos os valores de $jitt$ para $f_0 = 110Hz$ foram alcançados dentro do limite para detectar patologias (1.04%). Para $f_0 = 200Hz$ e $f_0 = 400Hz$ todos os valores obtidos foram normais, como podem-se observar na Fig.(9.49), valores acima da linha tracejada correspondem a patologia.

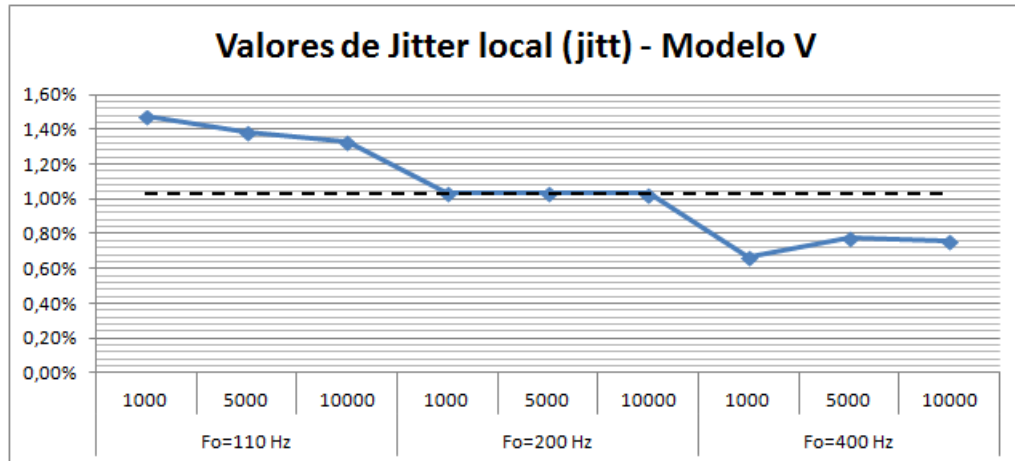


Figura 9.49: Gráfico com os valores de $jitt$ para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000).

- Todos os valores de $jitta$ para $f_0 = 110Hz$ foram atingidos dentro do limite para detectar patologias ($83.200\mu s$). Para $f_0 = 200Hz$ e $f_0 = 400Hz$ todos os valores obtidos foram normais, como podem-se observar na Fig.(9.50), valores acima da linha tracejada correspondem a patologia.

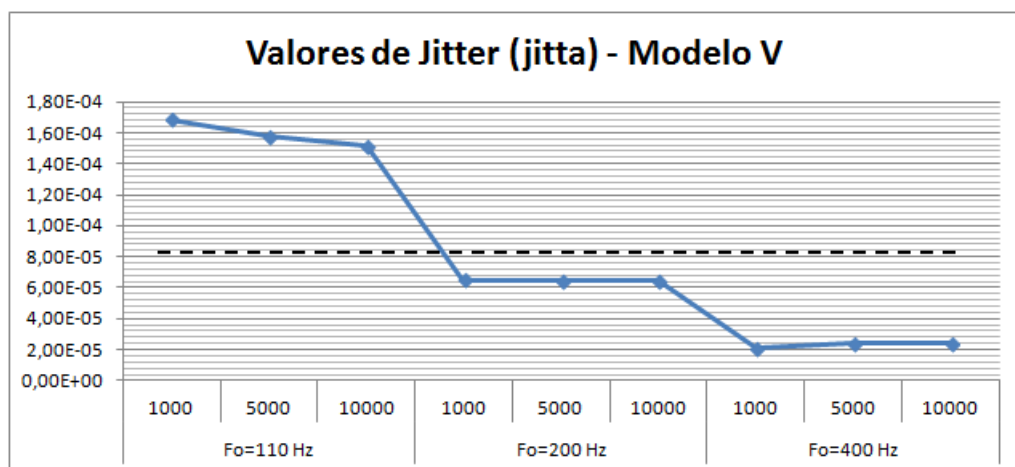


Figura 9.50: Gráfico com os valores de $jitt$ para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000).

- Todos valores de rap para $f_0 = 110Hz$ foram alcançados dentro do limite para detectar patologias (0.68%). Para $f_0 = 200Hz$ e $f_0 = 400Hz$ todos os valores obtidos foram normais, como podem-se observar na Fig.(9.51), valores acima da linha tracejada correspondem a patologia.

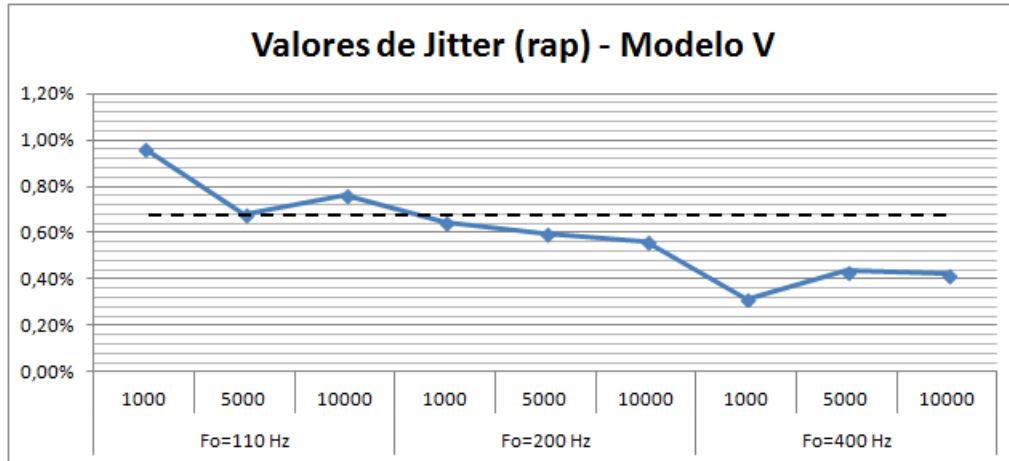


Figura 9.51: Gráfico com os valores de rap para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000).

- Todos valores de $ppq5$ para $f_0 = 110Hz$ foram alcançados dentro do limite para detectar patologias (0.68%), exceto para o intervalo 1000 que corresponde um valor normal. Para $f_0 = 200Hz$ e $f_0 = 400Hz$ todos os valores obtidos foram normais, como podem-se observar na Fig.(9.52), valores acima da linha tracejada correspondem a patologia.

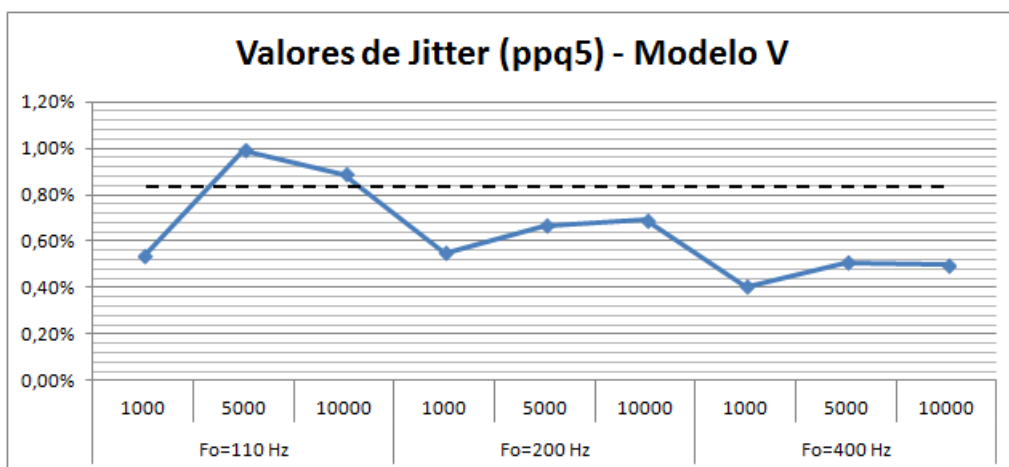


Figura 9.52: Gráfico com os valores de $ppq5$ para as $f_0 = 110Hz$, $f_0 = 200Hz$ e $f_0 = 400Hz$ do Modelo V, com os intervalos de tempo (1000, 5000 e 10000).

Considerações

Nas simulações com $f_0 = 110Hz$, foi encontrada uma discrepância entre as medidas de uma mesma simulação, para o intervalo de tempo 1000 foi obtido um valor de *ppq5* normal, caso contrário aos valores obtidos pelas outras medidas *jitt*, *jitta* e *rap* correspondentes a patologia.

No modelo geral de Jitter, apenas com o valor de $f_0 = 110Hz$ que representa o sinal da glote masculino, foram obtidos valores que detectam patologias para quase todas as diferentes formas de Jitter, entretanto os atingidos para mulheres e bebês encontram-se na faixa normal.

Capítulo 10

Conclusões

A dissertação propõe simular modelos estocásticos de Jitter conhecidos da literatura, fazendo uma análise de sensibilidade dos parâmetros e discutindo diferentes medidas de Jitter.

Basicamente, em todos os modelos simulados, para pequenos valores da constante b , a qual está relacionada com o valor máximo do distúrbio, as diferentes medidas não detectaram patologias.

Porém ao aumentar o valor do distúrbio, nos comprimentos do ciclo glotal, a indicação de patologia era alterada, dependendo da medida usada e da frequência fundamental considerada.

Principalmente, os valores de Jitter que indicavam patologia foram maiores na $f_0 = 110\text{Hz}$, constando-se que com menor frequência fundamental produz-se maior alteração nos ciclos glotais.

Para o menor valor assumido de b , 0.05, não foram observadas discrepâncias entre os resultados obtidos de uma mesma simulação.

As medidas que apresentaram maiores variações foram *jitta* e *jitt*, principalmente nas faixa de frequência fundamental de 110Hz e 400 Hz.

As variações entre os resultados das simulações considerando os diferentes intervalos de tempo foram maiores com o valor de b , 0.25, sendo maiores para o intervalo de tempo 1000.

As diferentes formas de Jitter permitiu avaliar em conjunto os resultados obtidos pelo programa *Praat*, observando similaridades na classificação do sinal, com características de patologia ou normais.

Os resultados do Modelo I apresentaram valores similares aos obtidos no Modelo II apenas com o menor valor de a_1 assumido em 0.25 para os valores de b , apesar que a frequência instantânea é perturbada de forma diferente em cada modelo. Mais valores de Jitter que detectam patologia foram encontrados no modelo II.

As diferentes formas de Jitter permitiu avaliar em conjunto os resultados obtidos pelo programa *Praat*, observando similaridades na classificação do sinal, com características de patologia ou normais.

No modelo geral de Jitter (Modelo V), na faixa de $f_0 = 110Hz$, apenas a medida *ppq5* não atingiu patologia para um intervalo de frequência, caso contrário as outras medidas de Jitter.

Quanto às discrepâncias encontradas em uma mesma simulação, a maioria das vezes, uma medida encontravam-se entre valores muito perto do limite da faixa que caracteriza à voz, gerando uma contrariedade na similitude entre elas.

Pode-se intuir que essas discrepâncias entre as medidas de uma mesma simulação, deve-se a que os modelos I,II,III e IV não levam em conta o comportamento real dos movimentos das cordas vocais, são modelos simples que simulam certas características para fornecer informações sobre o comportamento do modelo mais geral.

Referências

- [1] Abberton, E., and Fourcin, A. (1997). "Electrolaryngography," in Clinical Phonetics, edited by M. J. Ball and C. Code Whurr, London!, pp. 119-148.
- [2] Andrade L.M.(2003), "Determinação dos Limiares de Normalidade dos Parâmetros Acústicos da Voz", [dissertação], Universidade de São Paulo.
- [3] Aronson, A.E. (1980), Clinical Voice Disorders, New York, Decker.
- [4] Aronson, A.E. (1990), Clinical Voice Disorders, 3rd ed. New York: Thieme.
- [5] Atkinson, J. E. (1976). "Inter- and intraspeaker variability in fundamental voice frequency," J. Acoust. Soc. Am. 60, 440-445.
- [6] Baken, R., Orlikoff, F. (2000), Clinical measurement of speech and voice - second edition. USA: Singular
- [7] Bangayan, P., Long, C., Alwan, A., Kreiman, J., and Gerrat, B. (1997). "Analysis by synthesis of pathological voices using the Klatt synthesiser," Speech Commun. 22, 343-368.
- [8] Batista A.P. (2010). "Fadiga Vocal nos Professores do Primeiro Ciclo do Ensino Básico", [dissertação], Universidade de Aveiro. Portugal
- [9] Behlau M. (2001), "O livro do Especialista," Rio de Janeiro: Revinter
- [10] Behlau M., Pontes P. (1995), Avaliação e tratamento das disfonias, São Paulo: Lovise.
- [11] Brandão, Alexandre de Souza ; Cataldo, Edson ; Leta, Fabiana . Um Novo Método Usando Autocorrelação para Extração da Frequência Fundamental em Sinais de Voz. TEMA. Tendências em Matemática Aplicada e Computacional, v. 8, p. 10-20, 2007.
- [12] Brandi, E. (1990), Voz falada: Estudo, Avaliação e Tratamento. Rio de Janeiro: Atheneu.
- [13] Breiman, L. (1986). Probability and Stochastic Processes (Scientific, Palo Alto, CA).
- [14] Boersma P.; Weenink D. Praat: doing phonetics by computer [Computer program] (2013). Version 5.3.55, retrieved 2 September 2013 from <http://www.praat.org/>
- [15] Cataldo E., "Introdução aos processos estocásticos", San Carlos, SP: SBMAC 2012.
- [16] Cataldo, E. ; Soize, C. ; Sampaio, R. . Uncertainty quantification of voice signal production mechanical model and experimental updating. Mechanical Systems and Signal Processing, v. 40, p. 718-726, 2013.

- [17] Cataldo, Edson ; Sampaio, Rubens ; Lucero, Jorge ; Soize, C. . Modeling random uncertainties in voice production using a parametric approach. *Mechanics Research Communications*, v. 35, p. 429-490, 2008.
- [18] Colton, R., Casper, J. (1996). *Compreendendo os Problemas de Voz*. Porto Alegre: Artes Médicas.
- [19] Costa B. (2009). *Características Vocais Acústicas de Homens com Voz e Laringe Normais* [dissertação]. Universidade Federal de Santa Maria. Brasil
- [20] Childers, D. G. (2000). *Speech Processing and Synthesis Toolboxes* (Wiley, New York), pp. 390-393.
- [21] Childers, D. G., Hicks, D. M., Moore, G. P., and Alsaka, Y. A. (1986). "A model for vocal fold vibratory motion, contact area, and the electroglottogram," *J. Acoust. Soc. Am.* 80, 1309-1320.
- [22] Childers, D. G., and Wu, K. (1990). "Quality of speech produced by analysis-synthesis," *Speech Commun.* 9, 97-117.
- [23] Davis, St. B. (1979). "Acoustic characteristics of normal and pathologic voices," in *Speech and Language, Advances in Basic Research and Practice*, Vol. 1, edited by N. J. Lass (Academic, New York), pp. 273-331.
- [24] De Felipe A.C.; Grillo M.H.; Grechi T.H., *Normatização de Medidas Acústicas para Vozes Normais. Rev Bras Otorrinolaringol.*, 72(5) (2006) 659-64.
- [25] De Guchteneere, R. (1996). "Mesure et analyse chronologique des fluctuations de la durée des cycles glottiques," unpublished Doctoral Thesis, Université Libre de Bruxelles, Brussels, Belgium.
- [26] Felipe, A.; Grillo, M; Grechi, T. (2006). *Normatização de medidas acústicas para vozes normais. Revista Brasileira de Otorrinolaringologia*, 72(5): 659-664.
- [27] Forero, L. ; Cataldo, E. ; Vellasco, M. ; Botelho, M. ; Apolinário, J. Classification of Vocal Aging Using Parameters Extracted From the Glottal Signal. *Journal of Voice*, v. 1, p. online-1, 2014.
- [28] Forero, L. ; Vellasco, M. ; Cataldo, E. Classification of voice aging using ANN and glottal signal parameters. In: *ANDESCOM, 2010, Bogota. ANDESCON, 2010 IEEE*, 2010.
- [29] Garcia R., Cobeta I. (1996), *Classificação de las disfonias en los transtornos de la voz*, Madrid: Garsi
- [30] Gardiner, C. W. (1983). *Handbook of Stochastic Methods* (Springer Verlag, Berlin).
- [31] Guimarães, I.; Abberton, E. (2005a) Health and voice quality in smokers: An exploratory investigation. *Logopedics Phoniatric Vocology*, 30: 185-191
- [32] Guimarães, I. (2007), *A Ciência e a arte da voz humana*. Alcoitão: Escola Superior de Saúde do Alcoitão. Santa Casa da Misericórdia de Lisboa.

- [33] González J., Cervera T.; Miralles J. (2002). Análisis acústico de la voz: fiabilidade en un conjunto de parâmetros multidimensionales. *Acta Otorrinolaringol Esp*, 53: 256-268
- [34] Hamilton, J. D. (1994). *Time Series Analysis* (Princeton U. P., Princeton,NJ), pp. 180-200.
- [35] Heiberger, V. L., and Horii, Y. (1982). "Jitter and shimmer in sustained phonation,"in *Speech and Language, Advances in Basic Research and Practice*, Vol. 7, edited by N. J. Lass (Academic, New York), pp. 299-332.
- [36] Hillenbrand, J. (1987). "A methodological study of perturbation and additive noise in synthetically generated voice signals,"*J. Speech Hear. Res.* 30, 448-461.
- [37] Hillenbrand, J. (1988). "Perception of aperiodicities in synthetically generated voices,"*J. Acoust. Soc. Am.* 83, 2361-2371.
- [38] Hollien, H., Michel, J., and Doherty, E. T. (1973). "A method for analysing vocal jitter in sustained phonation,"*J. Phonetics* 1, 85-91.
- [39] Horii, Y. (1979). "Fundamental frequency perturbation observed in sustained phonation,"*J. Speech Hear. Res.* 22, 5-19.
- [40] Imaizumi, S. (1986). "Acoustic measures of roughness in pathological voice,"*J. Phonetics* 14, 457-462.
- [41] Imaizumi, S., and Gauffin, J. (1992). "Acoustic and perceptual modelling of the voice quality caused by fundamental frequency perturbation,"in *Proceedings International Conference on Spoken Language Processing*, Banff, pp. 133-136.
- [42] Kasuya, H., Masubichi, K., Ebihara, S., and Yoshida, H. (1986). "Preliminary results on voice screening,"*J. Phonetics* 3-4, 445-456.
- [43] Kélem Gomes Lourenço (2013), *Modelagem, análise matemática e validação experimental in vitro da fonação*. Instituto de Ciências Exatas Departamento de Matemática, Universidade de Brasília.
- [44] Kelley, W. G., and Peterson, A. C. (2001). *Difference Equations* (Academic, San Diego), pp. 125-150 and 164.
- [45] Kleijn W. B.; Paliwal K. K. eds., Talkin, "A Robust Algorithm for Pitch Tracking (RAPT) in Speech Coding and Synthesis", Elsevier ISBN 0444821694, 1995.
- [46] Kloeden, P. E., and Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin), pp. 305-338.
- [47] Koike, Y., Takahashi, H., and Calcaterra, T. C. (1977). "Acoustic measures for detecting laryngeal pathology,"*Acta Oto-Laryngol.* 84, 105-117.
- [48] Lalwani, A. L., and Childers, D. G. (1991). "Modeling vocal disorders via formant synthesis,"*Proc. IEEE International Conference Acoustics, Speech and Signal Processing* Vol. 1, pp. 505-508.

- [49] Laver, J. (1980). *The Phonetic Description of Voice Quality* (Cambridge U. P., Cambridge).
- [50] Laver, J., Hiller, S., Mackenzie, J., and Rooney, E. (1986). "An acoustic screening system for the detection of laryngeal pathology," *J. Phonetics* 14, 517-524.
- [51] Mauprivez, J. ; Cataldo, Edson ; Sampaio, Rubens . Artificial neural networks applied to the estimation of random variables associated to a two-mass model for the vocal folds. *Inverse Problems in Science & Engineering* (Print), v. 20, p. 209-225, 2012.
- [52] Milenkovic, P. (1987). "Least mean square measures of voice perturbation," *J. Speech Hear. Res.* 30, 529-538.
- [53] Muta, H., Baer, T., Wagatsuma, K., Muraoka, T., and Fukuda, H. (1988). "A pitch-synchronous analysis of hoarseness in running speech," *J. Acoust. Soc. Am.* 84, 1292-1301.
- [54] Nieto A.; Del Palacio A.J; Lorenzo A.J; Vegas A; Cobeta I. (1995), *Los ordenadores en el análisis de la voz: Aplicaciones clínicas*. *Acta Otorrinolaringologica Española*, V.46,n.4, 241-245
- [55] Qi, Y., and Shipp, T. (1992). "An adaptive method for tracking voicing irregularities," *J. Acoust. Soc. Am.* 91, 3471-3477.
- [56] Reetz, H. and Jongman, A. (2009), *Phonetics: Transcription, production, acoustics and perception*. USA: Wiley-Blackwell
- [57] Rozsypal, A., and Millar, B. (1979). "Perception of jitter and shimmer in synthetic vowels," *J. Phonetics* 7, 343-355.
- [58] Pinho S. (1998), *Avaliação e tratamento da voz. Fundamentos em fonoaudiologia. Tratando os distúrbios da voz*. Rio de Janeiro :Guanabara-Koogan, pp 2-48
- [59] Pinto, N. B., and Titze, I. R. (1990). "Unication of perturbation measures in speech signals," *J. Acoust. Soc. Am.* 87, 1278-1289.
- [60] Scalassara, P. R. et al.(2009), Relative entropy measures applied to healthy and pathological voice characterization. *Applied Mathematics and Computation*, Elsevier Science Inc, v. 270, n. 1, p. 95-108.
- [61] Schoentgen J. (2000). *Stochastic models of jitter*. Laboratoire de Phonétique Experimentale, Université Libre de Bruxelles, Brussels, Belgium
- [62] Schoentgen, J., and De Guchteneere, R. (1991). "An algorithm for the measurement of jitter," *Speech Commun.* 10, 533-538.
- [63] Schoentgen, J., and De Guchteneere, R. (1995). "Time series analysis of jitter," *J. Phonetics* 23, 189-201.
- [64] Schoentgen, J., and De Guchteneere, R. (1996). "Searching for nonlinear relations in whitened jitter time series," *Proceedings International Conference on Spoken Language Processing*, Philadelphia, pp. 753-756.

- [65] Schoentgen, J., and De Guchteneere, R. (1997). "Predictable and random components of jitter," *Speech Commun.* 21, 255-272.
- [66] Steiglitz, K. (1996). *A Digital Signal Processing Primer* (Addison-Wesley, Menlo Park, CA), pp. 86-90.
- [67] Tabith, A.JR. (1995), *Foniatria: Disfonia, Fissuras, Paralisia Cerebral*, 7ed. São Paulo, Cortez
- [68] Ternstrom, S. (2005). Does the acoustic waveform mirror the voice. *Logopedics Phoniatrics Vocology*, 30: 100-105
- [69] Teixeira J.P.; Ferreira D.; Carneiro S., *Análise Acústica Vocal-Determinação do Jitter e Shimmer para Diagnóstico de Patologias da Fala*. CLME'2011_0706A.
- [70] Titze, I. R. (1984). "Parametrization of the glottal area, glottal flow, and vocal fold contact area," *J. Acoust. Soc. Am.* 75, 570-580.
- [71] Titze, I. R. (1989). "A four-parameter model of the glottis and vocal fold contact area," *Speech Commun.* 8, 191-201.
- [72] Titze, I. R. (1991). "A model for neurological sources of aperiodicity in vocal fold vibration," *J. Speech Hear. Res.* 34, 460-472.
- [73] Titze, I. R. (2000). "On the simulation of neurologic jitter and tremor with a bio-mechanical model of laryngeal muscle activation," in *Proceedings 4th International Workshop on Advances in Quantitative Laryngoscopy, Voice, Speech Research*, Jena, pp. 132-138.
- [74] Titze, I. R., Horii, Y., and Scherer, R. C. (1987). "Some technical considerations in voice perturbation measurements," *J. Speech Hear. Res.* 30, 252-260.
- [75] Ventsel, H. (1973). *Théorie des probabilités des* (Editions MIR, Moscow!).
- [76] Winholtz, W. S., and Ramig, L. O. (1992). "Vocal tremor analysis with the vocal demodulator," *J. Speech Hear. Res.* 35, 562-573.
- [77] S.A. Zampieri, M. Behlau, O.C. do Brasil, *Análise de cantores de baile em estilo de canto popular e lírico: perceptivo-auditiva, acústica e da configuração laríngea*. *Rev Bras Otorrinolaringol.*, V.68, (2002) 378-86.

APÊNDICE A - Apêndice

APÊNDICE A - Rotina para o cálculo da f_0 pelo Método Cepstral

```
[y,fs,nbit]=wavread('...'); %carrega o arquivo de audio
a1=1;%posição inicial do vector do sinal y a ser analisado
N=length(y)
N1=512/2;
C1=floor(0.002*fs);%arredonda os valores dos harmonicos(indices)
y=detrend(y);%remoção da eventual tendência linear do vector y
c=0;
while a1<N-2*N1
    x=y(a1:a1+511); %segmento do sinal y a ser analisado
    z=rceps(x); %Função do Matlab que determina o cepstro
    [z0,imax]=max(z(C1:512/2)); %Z0 corresponde ao "pico"
    imax=imax+C1; %índice máximo, onde se encontra o "pico"
    t0=imax/fs*1000; %passar para (ms)
    f0=1/t0*1000;
    a1=a1+N1;
    c=c+1;
    F0(c)=f0; %F0(c) é a frequencia fundamental
end
```

APÊNDICE B - Rotina para o cálculo da f_0 pelo Método da autocorrelação

```

clear all
clc
[y,fs,nbit]=wavread('...');
N=floor(0.02*fs); %Define janelas de 20 ms
C1=floor(0.002*fs); %Define uma frequência mínima de 50 Hz
t1=(1:N)/fs*1000;
p=1;
d=0;
while p<length(y)-N
    y1=y(p:p+320).*hanning(320+1);
    p=p+N;
    X=fft(y1,640); %Transformada directa de Fourier
    rx=ifft(abs(X).^2); %Transformada inversa de Fourier
    [x0,imax]=max(rx(C1:N));
    imax=imax+C1;
    t0=imax/fs*1000;
    f0=1/t0*1000;
    d=d+1;
    F0(d)=f0;
end

```

APÊNDICE C - Rotina para o cálculo do Jitter

```

clear all
clc

[x,fs,nbit]=wavread(' '); %carrega o arquivo de audio
y=x(1*fs:length(x)); %Tamanho do Sinal
[fx,tt]=fxrapt(y,fs,'g'); %Funcao que calcula a frequencia fundamental
Fx=fx';
N=length(Fx); %Numero de Periodos
Tx = 1./Fx; %Periodo Fundamental
B = sum(Tx)/N; %Media de Tx
Tx2 = Tx(2:(N-1));
Tx3 = Tx(3:(N-2));

%Diferenca entre dois periodos consecutivos
for i=1:(N-1)
    A(i)=abs(Tx(i+1)-Tx(i));
end

%Somatorio do Tx e seus dois vizinhos
for i=2:(N-1)
    media3(i-1)=(Tx(i-1)+Tx(i)+Tx(i+1))/3;
end

% Somatorio do Tx e seus quatro vizinhos
for i=3:(N-2)
    media5(i-2)=(Tx(i-2)+Tx(i-1)+Tx(i)+Tx(i+1)+Tx(i+2))/5;
end

```

```
jitt a = sum(A)/(N-1)  %Jitter(Jitt a)
```

```
jitt =(jitt a/B)*100  %Jitter(Jitt)
```

```
rap = 100*(sum(abs(Tx2-media3))/((N-2)*B))  %Jitter(rap)
```

```
ppq5 = 100*(sum(abs(Tx3-media5))/((N-4)* B))  %Jitter(ppq5)
```

APÊNDICE D - Rotinas da modelagem do Jitter

D.1 Modelo I

```

clear all;
clc;

delta=10e-5;%time step
h0=0.1;%semi-abertura inicial glotal
b=0.25;%constante que varia o valor de Jitter
f0=400; %freq. fundamental
m=1000;%numero de iterações
n1=0:(m-1);

x=randn(1,m);
y=x./abs(x);
y1=sqrt(delta)*y; %ruído branco

for n=1:m
    soma=0;

    for i=1:n %somatório do ruído branco
        soma=y1(i)+soma;
    end

    teta_n(n)=2*pi*f0*n*delta+2*pi*b*soma;
    h(n)=2*h0+2*sin(teta_n(n)); %deslocamento da corda vocal
end

wavwrite(h,'modelo1')
```

D.2 Modelo II

```

clear all;
clc;

delta=10e-5;%time step
h0=0.1;%semi-abertura inicial glotal
b=0.25;
f0=400; %freq. fundamental
m=10000;%Intervalo de tempo
n1=0:(m-1);
teta0=0;
a1=0.25;

x=randn(1,m);
e=x./abs(x);
e1=sqrt(delta)*e; %ruído branco
Y0=0;%condição inicial
for n=1:
    if n==1
        Y(n)=b*e1(n);
    else
        Y(n)=a1*Y(n-1)+b*e1(n);
    end
end
soma=0;
for n=1:m
    for i=1:n %somatório do
        if (i==1)
            soma=Y0;
        else
            soma=Y(i-1)+soma;
        end
    end
    teta(n)=teta0+2*pi*f0*n*delta+2*pi*soma;
    h(n)=2*h0+2*sin(teta(n));%deslocamento da corda vocal

```

```
end
wavwrite(h, 'modelo2')
```

D.3 Modelo III

```
clear all;
clc;

delta=10e-5;%time step
K=0.1;
b=0.2;
h0=0.1;%semi-abertura inicial glotal
f0=110;%freq. fundamental
m=1000;%Intervalo de tempo
n1=0:(m-1);

xr=randn(1,m);
e1=xr./abs(xr);
er=sqrt(delta)*e1;%ruído branco

xl=randn(1,m);
e2=xl./abs(xl);
e1=sqrt(delta)*e2;%ruído branco

teta_l0=0;
teta_r0=0;
el0=0;
er0=0;

for n=1:m

    if n==1
        teta_l(n)=teta_l0+2*pi*f0*delta-K*sin(teta_l0-teta_r0)+2*pi*b*e1(n);
        teta_r(n)=teta_r0+2*pi*f0*delta-K*sin(teta_r0-teta_l0)+2*pi*b*er(n);
    else
        teta_l(n)=teta_l(n-1)+2*pi*f0*delta-K*sin(teta_l(n-1)-teta_r(n-1));
```

```

+2*pi*b*el(n-1);
    teta_r(n)=teta_r(n-1)+2*pi*f0*delta-K*sin(teta_r(n-1)-teta_l(n-1)
+2*pi*b*er(n-1);
end
end

for n=1:m
    h(n)=2*h0+2*cos((teta_r(n)-teta_l(n))/2)*sin((teta_r(n)
+teta_l(n))/2);%deslocamento da corda vocal
    teta(n)=(teta_r(n)+teta_l(n))/2;
end

wavwrite(h,'modelo3')
```

D.4 Modelo IV

```

clear all;
clc;

delta=10e-5;%time step
h0=0.1;%semi-abertura inicial glotal
b=0.25;%constante que varia o valor de Jitter
f0=110;%freq. fundamental
L0=1.0;%males(L0=1.6)/ female(L0=1.0)
T0=0.35;%males(T0=0.45) female(T0=0.35)
L=0.038*(L0^1.6)*(f0^0.5);%comprimento na vibracao das cordas vocais
T=(T0*L0)/L;
c=200;
fs=10*2*pi;%freq de amostragem
phi=2*pi*f0*T/c;%atraso de fase entre os movimentos das margens i e s
D=phi*fs;%atraso de fase em amostras
teta_i0=0;
m=1000;%Intervalo de tempo
n1=0:(m-1);

x=randn(1,m);
```

```

e=x./ abs(x);
e1=sqrt(delta)*e; %ruído branco

e0=0;
somaj=0;
somak=0;
somal=0;

for n=1:m

for j=1:n %somaj do
    if (j==1)
        somaj=e0;
    else
        somaj=e1(j-1)+somaj;
    end
end

for k=1:(n-(n-1)) %somak do
    if (k==1)
        somak=e0;
    else
        somak=e1(k-1)+somak;
    end
end

for l=1:n %somal do
    if (l==1)
        somal=e0;
    else
        somal=e1(l-1)+somal;
    end
end

if n-D<1

```

```

    if n==1
        teta_i(n)=2*pi*f0*n*delta+2*pi*b*somaj;
        teta_s(n)=2*pi*f0*delta+2*pi*b*e0;
    else
        teta_i(n)=2*pi*f0*n*delta+2*pi*b*somaj;
        teta_s(n)=2*pi*f0*delta+2*pi*b*el(n-1);
    end

else
    teta_i(n)=2*pi*f0*n*delta+2*pi*b*somaj;
    teta_s(n)=2*pi*f0*(n-D)*delta+2*pi*b*somak+2*pi*b*somal;

end

hi(n)=2*h0+2*sin(teta_i(n));%deslocamento da corda vocal
hs(n)=2*h0+2*sin(teta_s(n));%deslocamento da corda vocal

end

wavwrite(hs,'modelo4s')
```

D.5 Modelo V

```

clear all;
clc;

delta=10e-5;%time step
Kl=0.1;
Kr=0.1;
b2=0.05;
b1=0.25;
h0=0.1;%semi-abertura inicial glotal

B=4;
Psi=12*pi;%6Hz
R=1/(B/2);
```

```

varphi=acos((2*R/(1+R^2))*cos(Psi));
a0=(1-R^2)*sin(varphi);
a1=(4*R^2/(1+R^2))*cos(Psi);
a2=-R^2;

f0=200; %freq. fundamental
L0=1.0;%males(L0=1.6)/ female(L0=1.0)
T0=0.35;%males(T0=0.45) female(T0=0.35)
L=0.038*(L0^1.6)*(f0^0.5);%comprimento na vibracao das cordas vocais
T=(T0*L0)/L;
c=200;
fs=10*2*pi; %freq de amostragem
phi=2*pi*f0*T/c; %atraso de fase entre os movimentos das margens sup e in
D1=phi*fs;%atraso de fase em amostras
D=round(D1);
m=1000;%Intervalo de tempo
n1=0:(m-1);

x1r=randn(1,m);
e1=x1r./abs(x1r);
e1r=sqrt(delta)*e1; %ruído branco

x2r=randn(1,m);
e2=x2r./abs(x2r);
e2r=sqrt(delta)*e2; %ruído branco

x1l=randn(1,m);
e3=x1l./abs(x1l);
e1l=sqrt(delta)*e3; %ruído branco

x2l=randn(1,m);
e4=x2l./abs(x2l);
e2l=sqrt(delta)*e4; %ruído branco

Yr0=0;%condição inicial

```

```
Yl0=0;%condição inicial
```

```
Yr1=0;%condição inicial
```

```
Yl1=0;%condição inicial
```

```
for n=1:m
```

```
    if n==1
```

```
        Yr(n)=a1*Yr0+a2*Yr1+2*pi*a0*b1*e1r(n);
```

```
        Yl(n)=a1*Yl0+a2*Yl1+2*pi*a0*b1*e1l(n);
```

```
    elseif n==2
```

```
        Yr(n)=a1*Yr(n-1)+a2*Yr0+2*pi*a0*b1*e1r(n);
```

```
        Yl(n)=a1*Yl(n-1)+a2*Yl0+2*pi*a0*b1*e1l(n);
```

```
    else
```

```
        Yr(n)=a1*Yr(n-1)+a2*Yr(n-2)+2*pi*a0*b1*e1r(n);
```

```
        Yl(n)=a1*Yl(n-1)+a2*Yl(n-2)+2*pi*a0*b1*e1l(n);
```

```
    end
```

```
end
```

```
e2l0=0;
```

```
e2r0=0;
```

```
teta_il0=0;
```

```
teta_ir0=0;
```

```
teta_sl0=0;
```

```
teta_sr0=0;
```

```
for n=1:m
```

```
    if n-D<1
```

```
        if n==1
```

```
            teta_il(n)=teta_il0+2*pi*f0*delta-Kl*sin(teta_il0-teta_ir0)  
            +Yl0+2*pi*b2*e2l0;
```

$$\begin{aligned} \text{teta_ir}(n) &= \text{teta_ir0} + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kr} * \mathbf{sin}(\text{teta_ir0} - \text{teta_il0}) \\ &+ \text{Yr0} + 2 * \mathbf{pi} * b2 * e2r0; \end{aligned}$$

$$\begin{aligned} \text{teta_sl}(n) &= \text{teta_il0} + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kl} * \mathbf{sin}(\text{teta_sl0} - \text{teta_sr0}) \\ &+ \text{Yl0} + 2 * \mathbf{pi} * b2 * e2l0; \end{aligned}$$

$$\begin{aligned} \text{teta_sr}(n) &= \text{teta_ir0} + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kr} * \mathbf{sin}(\text{teta_sr0} - \text{teta_sl0}) \\ &+ \text{Yr0} + 2 * \mathbf{pi} * b2 * e2r0; \end{aligned}$$

else

$$\begin{aligned} \text{teta_il}(n) &= \text{teta_il}(n-1) + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kl} * \mathbf{sin}(\text{teta_il}(n-1) \\ &- \text{teta_ir}(n-1)) + \text{Yl}(n-1) + 2 * \mathbf{pi} * b2 * e2l(n-1); \end{aligned}$$

$$\begin{aligned} \text{teta_ir}(n) &= \text{teta_ir}(n-1) + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kr} * \mathbf{sin}(\text{teta_ir}(n-1) \\ &- \text{teta_il}(n-1)) + \text{Yr}(n-1) + 2 * \mathbf{pi} * b2 * e2r(n-1); \end{aligned}$$

$$\begin{aligned} \text{teta_sl}(n) &= 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kl} * \mathbf{sin}(\text{teta_sl}(n-1) - \text{teta_sr}(n-1)) \\ &+ \text{Yl}(n-1) + 2 * \mathbf{pi} * b2 * e2l(n-1); \end{aligned}$$

$$\begin{aligned} \text{teta_sr}(n) &= 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kr} * \mathbf{sin}(\text{teta_sr}(n-1) - \text{teta_sl}(n-1)) \\ &+ \text{Yr}(n-1) + 2 * \mathbf{pi} * b2 * e2r(n-1); \end{aligned}$$

end

else

$$\begin{aligned} \text{teta_il}(n) &= \text{teta_il}(n-1) + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kl} * \mathbf{sin}(\text{teta_il}(n-1) \\ &- \text{teta_ir}(n-1)) + \text{Yl}(n-1) + 2 * \mathbf{pi} * b2 * e2l(n-1); \end{aligned}$$

$$\begin{aligned} \text{teta_ir}(n) &= \text{teta_ir}(n-1) + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kr} * \mathbf{sin}(\text{teta_ir}(n-1) \\ &- \text{teta_il}(n-1)) + \text{Yr}(n-1) + 2 * \mathbf{pi} * b2 * e2r(n-1); \end{aligned}$$

$$\begin{aligned} \text{teta_sl}(n) &= \text{teta_il}(n-D) + 2 * \mathbf{pi} * f0 * \text{delta} - \text{Kl} * \mathbf{sin}(\text{teta_sl}(n-1) \\ &- \text{teta_sr}(n-1)) + \text{Yl}(n-1) + 2 * \mathbf{pi} * b2 * e2l(n-1); \end{aligned}$$

```
teta_sr(n)=teta_ir(n-D)+2*pi*f0*delta-Kr*sin(teta_sr(n-1)
-teta_sl(n-1))+Yr(n-1)+2*pi*b2*e2r(n-1);

end

end

for n=1:m

    hi(n)=2*h0+sin(teta_ir(n))-sin(teta_il(n)+pi);
    hs(n)=2*h0+sin(teta_sr(n))-sin(teta_sl(n)+pi);

end

wavwrite(hs,'modelo5s')
```