



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Simulação computacional de uma fenda glótica

Elias Amadeu de Souza Gomes

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientador
Prof. Dr. Jorge C. Lucero

Brasília
2015

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Coordenador: Prof.^a Dr.^a Maristela Holanda

Banca examinadora composta por:

Prof. Dr. Jorge C. Lucero (Orientador) — CIC/UnB

Prof. Dr. Pedro Berger — CIC/UnB

Prof.^a Dr.^a Aveliny Lima-Gregio — FCE/UnB

CIP — Catalogação Internacional na Publicação

Gomes, Elias Amadeu de Souza.

Simulação computacional de uma fenda glótica / Elias Amadeu de Souza Gomes. Brasília : UnB, 2015.

129 p. : il. ; 29,5 cm.

Monografia (Graduação) — Universidade de Brasília, Brasília, 2015.

1. Síntese de Voz, 2. Fenda Glótica, 3. Simulação Computacional,
4. Patologias da Voz

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil

Dedicatória

Dedico a todos os interessados no estudo da síntese da voz em geral, a todos os profissionais que buscam melhorar a qualidade de vida das pessoas através de pesquisas, desenvolvimento e aplicação do conhecimento e a todos que me ajudaram a progredir durante todos esses anos de estudo.

Agradecimentos

Agradeço primeiramente a Deus pelas oportunidades de aprendizado e desenvolvimento dos meus estudos, agradeço a minha família pelo apoio, amor e carinho, ao meu orientador, o Dr. Jorge C. Lucero, pelo suporte e por me guiar durante esse trabalho e no decorrer do curso e, a todos aqueles que me apoiaram, suportaram e ensinaram durante todos esses anos.

Abstract

O presente trabalho busca simular computacionalmente uma fenda glótica a fim de auxiliar os profissionais da área de saúde da voz no estudo e diagnóstico dessa patologia vocal e, aos interessados no campo de estudo da síntese da voz humana, este trabalho fornece uma visão geral acerca de alguns modelos e métodos que permitem o desenvolvimento de um sintetizador de voz computacional.

Palavras-chave: Síntese de Voz, Fenda Glótica, Simulação Computacional, Patologias da Voz

Abstract

This study aims to simulate computationally a glottal gap to assist professionals on the fields of voice healthcare in the study and diagnosis of this vocal pathology, and for those interested in the field of vocal synthesis, this paper provides an overview of some models and methods that allow the development of a computational voice synthesizer.

Keywords: Vocal Synthesis, Glottal Gap, Computer Simulation, Vocal Pathologies

Sumário

1	Introdução	1
1.1	Aparelho Fonador	1
1.1.1	A Anatomia	1
1.1.2	Músculos e Cartilagens	3
1.1.3	Pregas Vocais	4
1.1.4	Overview: Geração de Som e Ressonadores	5
1.2	Patologias da Voz	8
2	Fundamentos Biofísicos para a Produção da Voz	9
2.1	A Biomecânica da Laringe	9
2.1.1	Fatores Biológicos que Afetam a Produção de Som na Laringe	9
2.2	Fluxo de Ar nas Vias Respiratórias	12
2.2.1	O Sistema Pulmonar	13
2.2.2	Leis de Conservação para Fluxos em Dutos	15
2.3	Oscilação das Pregas Vocais	16
2.3.1	Critérios para Oscilação	17
2.3.2	Tipos de Oscilação	17
2.3.3	Oscilador Massa-Mola	17
2.4	Geração e Propagação do Som	18
2.4.1	Reflexão do Som	19
3	Computação e Síntese de Voz	21
3.1	Modelos Computacionais para Representação das Pregas Vocais	21
3.1.1	Modelo Mono-Massa	21
3.2	Modelo Computacional para Representação do Trato Vocal	23
3.2.1	Modelo Reflexão de Onda	24
4	SynPath	26
4.1	Descrição do Sistema	26
4.2	Concepção Matemática e Modelagem do Sistema	27
4.2.1	Trato Vocal	27
4.2.2	Pregas Vocais	27
4.2.3	Fluxo de Ar na Glote	28
4.2.4	<i>Jitter</i> Vocal	29
4.3	Requisitos do Sistema	30
4.3.1	Requisitos Funcionais	30
4.3.2	Requisitos Não-Funcionais	30

4.3.3	Requisitos de Dependabilidade	31
4.4	Diagramas do Sistema	31
4.4.1	Atores do Sistema	31
4.4.2	Diagrama de Casos de Uso	32
4.5	Arquitetura do Sistema	32
4.5.1	Arquitetura Conceitual	32
4.5.2	Arquitetura de Implementação	33
4.5.3	Diagrama de Atividades	34
5	Implementação da Fenda Glótica	36
5.1	Modelagem Matemática	36
5.2	Implementação	37
5.2.1	Classe main.py	37
5.2.2	Classe config.py	37
5.2.3	Classe soundOutput.py	38
5.2.4	Classe vocalTractModels.py	38
5.2.5	Classe vocalFoldModels.py	38
5.2.6	Classe synthesis.py	39
5.3	Resultados e Analise	40
5.3.1	Resultados S/N - Voz Masculina	41
5.3.2	Resultados S/N - Voz Feminina	42
5.3.3	Gráficos - Voz Masculina	43
5.3.4	Gráficos - Voz Feminina	47
5.3.5	Análise dos Resultados	52
5.4	Conclusão	53
5.5	Trabalhos Futuros	53
	Referências	54

Lista de Figuras

1.1	Aparelho Fonador [12]	2
1.2	Esquema da laringe e parte superior da traquéia. [7]	3
1.3	Músculos Intrínsecos da Laringe [6]	4
1.4	Músculos Extrínsecos da Laringe [6]	4
1.5	Cordas Vocais e Componentes [6]	5
1.6	ciclo Vibratório das Pregas Vocais [12]	6
1.7	Trato Vocal [12]	7
1.8	(a) Espectro inicial do som. (b) A função de transferência do ressonador. (c) O espectro e voz gerado	7
2.1	Curva Hipotética Tensão-Deformação das Cordas vocais Humanas [23]	11
2.2	Pressões Atuantes no Torso [27]	13
2.3	Sistema Pulmonar [7]	14
2.4	Dinâmica do Fluido em um Duto com Mudança de Área Transversal [7]	15
2.5	Um Oscilador Massa-Mola [7]	18
3.1	Modelo Mono-Massa [23]	22
3.2	Modelagem do trato vocal para diferentes vogais [7]	25
4.1	Esquema das Pregas Vocais [11]	27
4.2	Diagrama de Casos de Uso do sistema.	32
4.3	Diagrama de Pacotes do sistema.	33
4.4	Diagrama de Componentes do sistema.	34
4.5	Diagrama de Atividades do sistema.	35
5.1	Parte da classe config.py	38
5.2	Parte da classe vocalFoldsModel.py	39
5.3	Fluxo Glotal	43
5.4	Espectro das Frequências de Resposta	43
5.5	Onda Sonora	44
5.6	Fluxo Glotal	44
5.7	Espectro das Frequências de Resposta	45
5.8	Onda Sonora	45
5.9	Fluxo Glotal	46
5.10	Espectro das Frequências de Resposta	46
5.11	Onda Sonora	47
5.12	Fluxo Glotal	47
5.13	Espectro das Frequências de Resposta	48

5.14 Onda Sonora	48
5.15 Fluxo Glotal	49
5.16 Espectro das Frequências de Resposta	49
5.17 Onda Sonora	50
5.18 Fluxo Glotal	50
5.19 Espectro das Frequências de Resposta	51
5.20 Onda Sonora	51

Lista de Tabelas

5.1	Resultados da síntese linear	41
5.2	Resultados da síntese quadrática	41
5.3	Resultados da síntese linear	42
5.4	Resultados da síntese quadrática	42

Capítulo 1

Introdução

Diversos trabalhos sobre modelagem e representação da voz foram produzidos nas últimas décadas, tendo como marca inicial, os trabalhos pioneiros de Flanagan e Landgraf [5], e Ishizaka e Flanagan [8].

A partir daí, vários modelos matemáticos foram propostos para reproduzir computacionalmente o comportamento da voz humana sobre diferentes aspectos, desde vozes produzidas dentro da normalidade funcional, até vozes com complexos padrões de vibração e instabilidades.

Entretanto, poucos estudos abordam a real simulação ou síntese de vozes com distúrbios. Lucero [16] propõe em seu trabalho a modelagem e o desenvolvimento de um sintetizador para este fim. É com base nesse software, nomeado de SynPath, que o trabalho aqui será desenvolvido.

Portanto, o objetivo deste trabalho é primeiramente abordar a modelagem e síntese da voz em um contexto computacional, explicando os diferentes modelos matemáticos propostos para estudá-la, e então desenvolver um novo módulo para o sintetizador SynPath cuja funcionalidade seja sintetizar vozes afetadas pela patologia de fenda glótica, que o atual sistema ainda não o faz.

Com isso, este software então será capaz de auxiliar fonoaudiólogos no diagnóstico desta patologia.

1.1 Aparelho Fonador

Para estudar a produção e a síntese da voz, é necessário ter um conhecimento acerca da anatomia e do funcionamento físico da voz [12]. Sendo assim, as subseções seguintes descreverão brevemente detalhes da anatomia do sistema fonador humano e como o som é produzido, moldado e influenciado por este sistema.

1.1.1 A Anatomia

A Figura 1.1 [12], mostra os órgãos associados com a produção da voz.

Dentro das condições normais, a voz é produzida quando um fluxo de ar vindo dos pulmões é convertido em energia acústica através da vibração das pregas vocais, localizadas na laringe. Os padrões de vibrações resultantes são moldados acusticamente quando o som passa pelo trato vocal acima da laringe. O sistema respiratório serve como uma

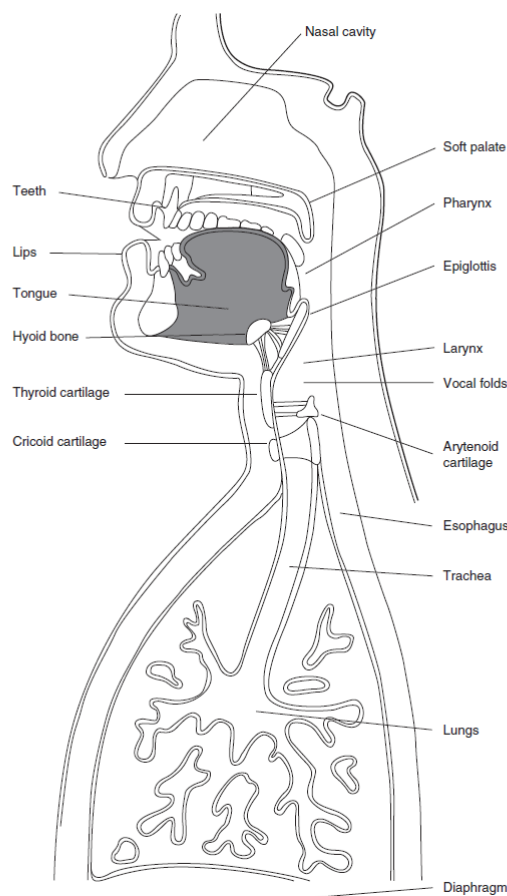


Figura 1.1: Aparelho Fonador [12]

fonte de potência para a produção do som, sendo responsável por movimentar o ar através do trato vocal.

A laringe atua como um oscilador convertendo a potência aerodinâmica produzida em energia sonora, sendo frequentemente retratada como a fonte da voz. No entanto, a mais importante função da laringe não é a produção de som, e sim, vedar as vias aéreas aos pulmões completamente, protegendo-as de objetos estranhos ou líquidos, principalmente durante a deglutição. De maneira análoga, a laringe serve como uma válvula de acesso às vias respiratórias e por essa característica, atua também no controle do fluxo de ar que por elas passam.

Sendo assim, é fácil notar que há uma necessidade de mobilidade para toda estrutura da laringe, logo é de se esperar que sua estrutura seja formada em sua maioria por cartilagens. De fato o é, com exceção de um osso chamado de Hioide, a laringe é basicamente formada por cartilagens e músculos.

A seguir, analisaremos brevemente a dinâmica dos músculos e cartilagens da laringe.

1.1.2 Músculos e Cartilagens

Os músculos e cartilagens atuam na sustentação e nos ajustes internos (tensão e relaxamento) da laringe. As pregas vocais estão localizadas dentro da laringe e devido à dinâmica das cartilagens e dos músculos, podem executar os movimentos citados de forma a produzir som.

Cartilagens da Laringe

A Figura 1.2, mostra um esquema da laringe, detalhando as cartilagens presentes.

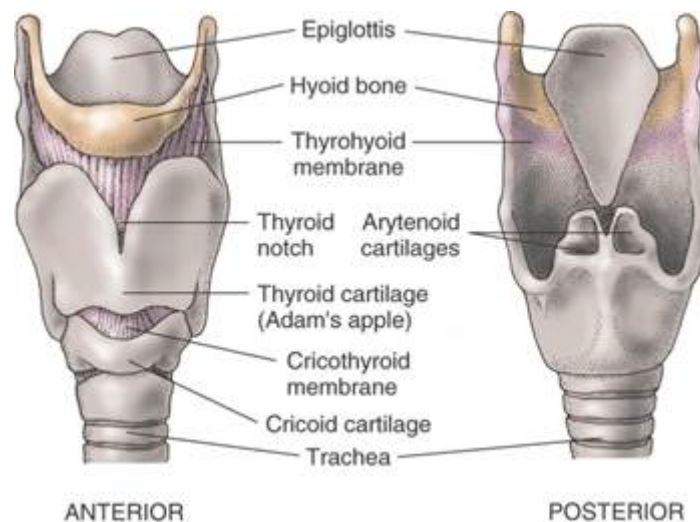


Figura 1.2: Esquema da laringe e parte superior da traquéia. [7]

De maneira sucinta, estas cartilagens servem como base de interconexão para os músculos intrínsecos ao redor da laringe. Dentre as cartilagens acima, a *epiglote* é responsável por vedar as vias respiratórias movimentando-se sobre a entrada das mesmas. Em geral, as cartilagens garantem a flexibilidade da laringe em conjunto com as articulações e ligamentos.

Músculos da Laringe

Os músculos na laringe podem ser divididos em dois grupos, os intrínsecos e os extrínsecos [1]. Os músculos intrínsecos interconectam as cartilagens da laringe, ao passo que, os extrínsecos conectam a laringe à outras estruturas externas, como o osso hióide. A Figura 1.3 detalha alguns dos músculos intrínsecos da laringe.

Alguns desses músculos têm influência direta em algumas características da voz. Por exemplo, o músculo tensor *cricotiroideo* é o músculo primário utilizado no controle do tom da voz. Por sua vez, o músculo cricoaritenóideo posterior atua na abdução das pregas vocais, ao passo que o músculo cricoaritenóideo anterior atua como adutor das pregas vocais. Um músculo muito relevante na produção de voz é o tireoaritenóideo, conhecido como TA, esse músculo é responsável pelo relaxamento das pregas vocais durante a fonação auxiliando no controle da tensão e no ajuste da laringe.

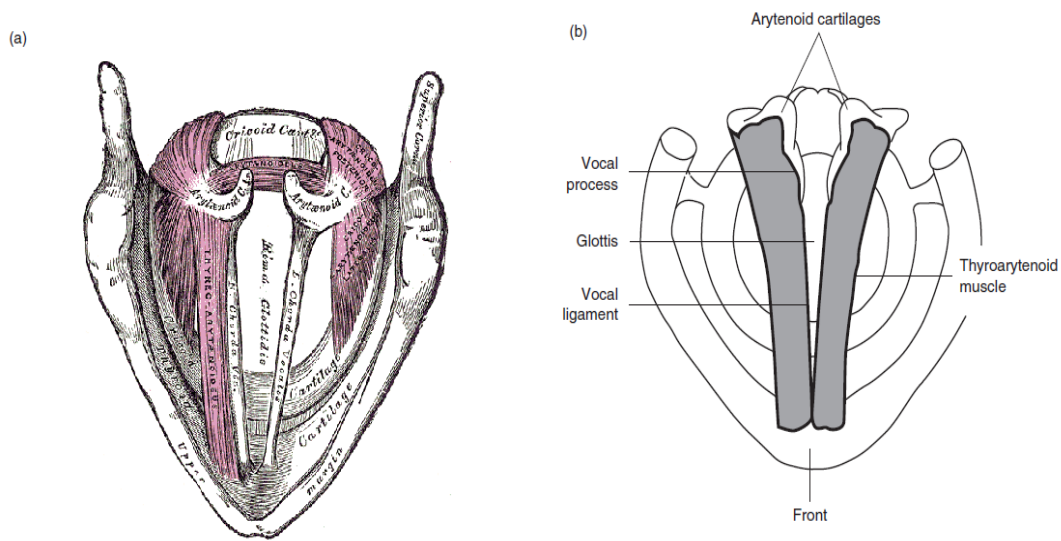


Figura 1.3: Músculos Intrínsecos da Laringe [6]

Os músculos extrínsecos, Figura 1.4, atuam basicamente na sustentação e no posicionamento da laringe, agindo como depressor e elevador da estrutura laríngea. Além disso também conectam estruturas do trato vocal à estrutura laríngea, como por exemplo a língua ao osso hioide, garantindo estabilidade e o movimento de elevação.

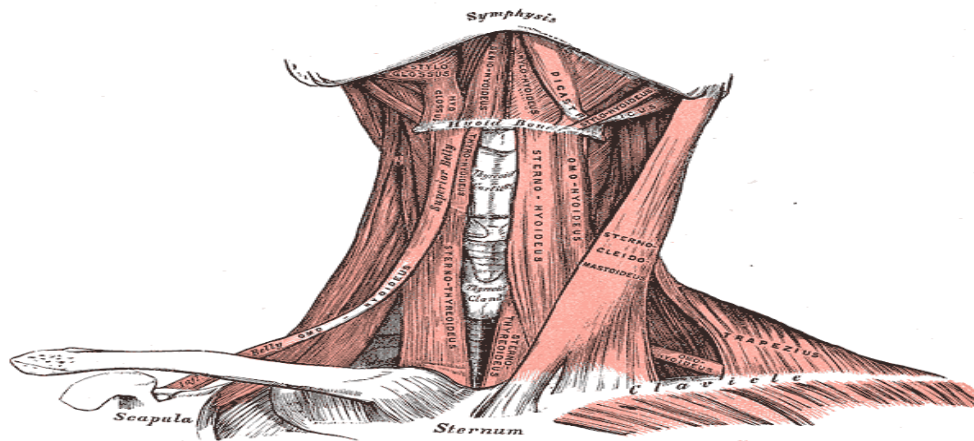


Figura 1.4: Músculos Extrínsecos da Laringe [6]

1.1.3 Pregas Vocais

As pregas vocais, como dito anteriormente, estão localizadas dentro da laringe, mais especificamente na parte superior da traqueia. Elas estão posteriormente ligadas às cartilagens aritenoides, e anteriormente ligadas à cartilagem tireoide. As suas bordas exteriores

estão ligadas ao músculo tireoaritenoideo citado anteriormente, enquanto as suas bordas interiores são livres.

As bordas das pregas vocais são construídas de epitélio, sendo compostas também de algumas fibras musculares. As pregas vocais são bandas triangulares planas de cor branca e acima de ambos os lados destas, se encontram as pregas vestibulares ou falsas pregas vocais.

O espaço entre as pregas vocais é chamado de *glote*, sendo que o que está acima da glote é denominado *supraglotal* e o que está abaixo é denominado *subglotal*. A Figura 1.5 mostra em mais detalhes a anatomia das pregas vocais, os componentes musculares e as cartilagens atuantes.

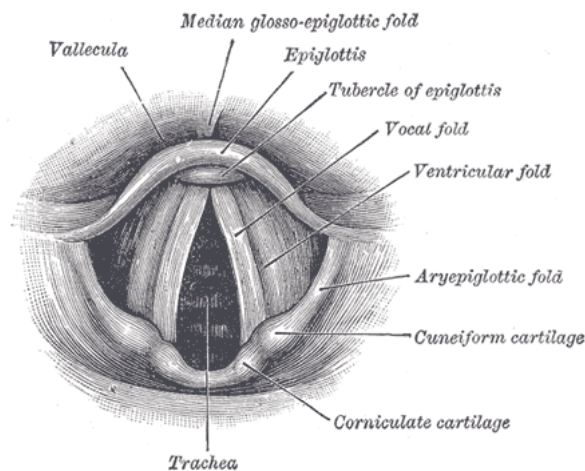


Figura 1.5: Cordas Vocais e Componentes [6]

1.1.4 Overview: Geração de Som e Ressonadores

A produção do som da voz é composta de duas etapas importantes que ocorrem na região subglotal/glotal e supraglotal. A primeira é a transformação da energia aerodinâmica em energia sonora, pelo movimento e vibração das pregas vocais. O segundo é a transformação do som primitivo gerado em voz, através da atuação dos ressonadores e formantes na região supraglotal.

A vibração das pregas vocais é extremamente complexa e, diversos músculos em união com a pressão exercida pelo ar atuam para tornar esse movimento possível. De maneira sintetizada, as pregas vocais vibram do topo ao fundo de maneira que não vibram como se fosse um bloco, mas sim de forma ondulatória, conforme mostrada na Figura 1.6.

Essa vibração é responsável por mudanças de pressão entre a região subglotal e a região glotal, causando uma diferença de pressão (condensação e rarefação) nessa área durante o movimento das pregas vocais, o que ocasiona a geração de som.

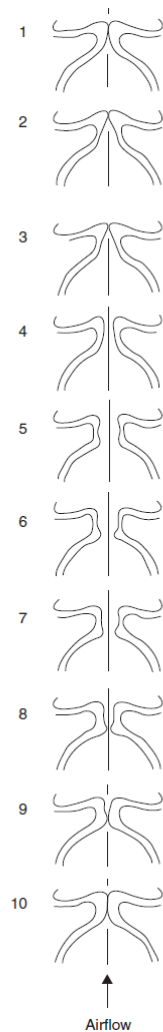


Figura 1.6: ciclo Vibratório das Pregas Vocais [12]

Entretanto, o som produzido é um som primitivo, conforme dito anteriormente, e para se transformar na voz característica humana, ele deve ser filtrado e moldado pelos ressonadores(formantes) no trato vocal(Figura 1.7).

Todas as cavidades mostradas no trato vocal atuam como ressonadores para a onda sonora produzida pelas pregas vocais. Um ressonador pode entrar em estado de vibração através de uma força aplicada ao mesmo em inércia ou por interação com algo que já esteja em estado de vibração. Neste segundo caso, as vibrações produzidas pelo ressonador serão ampliações ou atenuações dependendo de quão próximas ou distantes, em termos de frequência, são as vibrações da onda sonora em contato com o ressonador.

Caso a onda sonora possua vibrações cujas frequências se assemelhem às frequências do ressonador, estas então serão amplificadas pelo ressonador. Entretanto, caso as frequências do som gerado vibrem em uma frequência distante da frequência natural do ressonador, então estas serão abafadas.

A voz passa por esse processo ao ser formada. Um som primitivo advindo das pregas vocais entra em contato com os ressonadores no trato vocal, estes por sua vez em conjunto

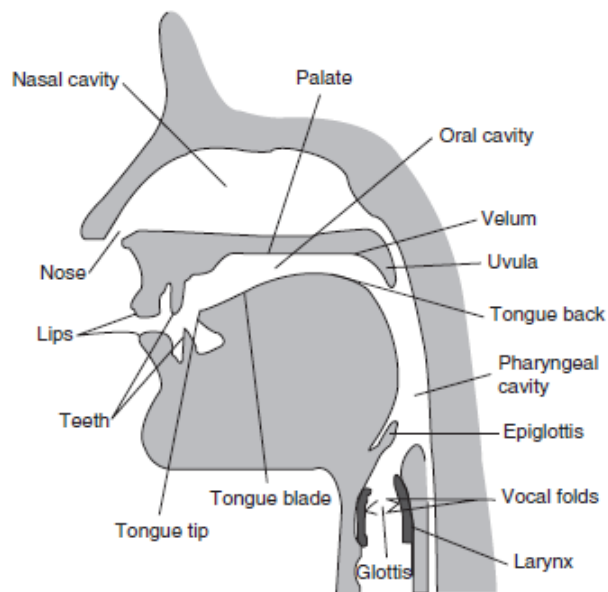


Figura 1.7: Trato Vocal [12]

com a forma do trato vocal, filtram as frequências de acordo com a frequência fundamental do som primitivo e geram o timbre característico da voz humana.

A Figura 1.8, ilustra a interação entre um ressonador e uma fonte sonora. A primeira figura mostra o som em sua forma primitiva quando gerado pelas pregas vocais, em seguida temos o ressonador com suas frequências naturais(formantes) e por fim, o espectro de voz gerado.

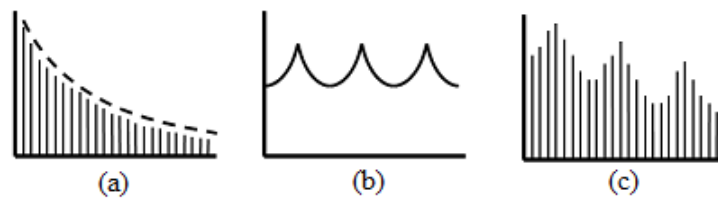


Figura 1.8: (a) Espectro inicial do som. (b) A função de transferência do ressonador. (c) O espectro e voz gerado

1.2 Patologias da Voz

Classificar sistematicamente patologias da voz é algo extremamente delicado. Tradicionalmente, médicos da voz classificam as desordens vocais em dois nichos: desordens orgânicas e desordens funcionais. Patologias orgânicas são aquelas relacionadas a uma lesão identificável em algum órgão, por sua vez, patologias funcionais são aquelas em que não há uma lesão identificável. A fenda glótica é uma patologia de desordem funcional proveniente de uma desordem orgânica, podendo esta ser um nódulo vocal, um cisto, uma alteração estrutural mínima da cobertura das pregas vocais, entre outros. Usualmente, profissionais da área de saúde que trabalham com voz se referem a fenda glótica como uma patologia por estresse mecânico.

As patologias por estresse mecânico geralmente ocorrem por colisões prolongadas e repetidas dos tecidos das pregas vocais. Acrescidos de tensão produzidos no pescoço e em todo o corpo, muitas vezes pessoas que se utilizam da voz como profissão acabam que por impôr condições de pressão, volume e desgaste da voz bem acima do normalmente suportado. Sem um devido cuidado com a voz, desde aquecimento e treino vocal até uma alimentação saudável, isso acaba gerando lesões orgânicas como nódulos vocais nas pregas. Curiosamente, na maioria dos casos, os nódulos vocais ocorrem no centro da membrana das pregas vocais, o que ocorre pois esta é a região em que há o maior impacto durante fonação. Outro fator curioso é que normalmente esses nódulos vocais são bilaterais, o que demonstra uma aplicação da terceira lei de Newton na geração de estresses simétricos nas pregas vocais, para toda ação há uma reação equivalente.

Após a incidência de uma lesão de desordem orgânica, esta acaba que por gerar uma pequena abertura nas pregas vocais durante o período de vibração, o que caracteriza uma voz soprosa ao detentor de tal lesão. Mesmo após a cura dessa lesão, deve-se tratar a voz com exercícios de fonoterapia a fim de que a fenda não torne a abrir ou piorar, aumentando o espaço entre as pregas vocais durante a vibração e prejudicando a fonação. É importante ressaltar que uma fenda pode existir sem que haja uma lesão orgânica nas pregas vocais. Sendo assim, trataremos aqui nesse trabalho a simulação de uma fenda uniforme sem a existência de uma lesão orgânica, o que significa que a simetria das pregas vocais será mantida e o seu corpo não terá alterações orgânicas.

Capítulo 2

Fundamentos Biofísicos para a Produção da Voz

Neste capítulo serão introduzidos princípios físicos e biológicos da produção da voz. Além dos conceitos introduzidos aqui, serão analisados os processos mecânicos da geração da voz por parte da laringe em conjunto com o trato vocal pois são de extrema importância para a concepção dos modelos computacionais e matemáticos utilizados para se desenvolver o sintetizador.

2.1 A Biomecânica da Laringe

Primeiramente, devemos ter em mente que o principal papel da laringe não é a produção de voz e sim a proteção das vias respiratórias. Dito isto, podemos fazer uma simples análise, visto que apesar de seu papel principal, a laringe também atua como um instrumento da fala humana. Se analisarmos os instrumentos criados pelo homem podemos notar que estes dependem basicamente de sua geometria, do material que o compõe e da interação de suas partes acústicas. Do mesmo modo, a laringe possui uma determinada geometria, é composta por tecido humano e em conjunto com o trato vocal compõe a parte acústica do nosso corpo. Entretanto, nada é tão simples, a sua geometria e as propriedades do material humano envolvidos na produção do som são bastante irregulares [23].

Outra analogia interessante é sobre o instrumento e quem o utiliza. Um bom pianista por exemplo, sua música é boa por que ele é habilidoso com o instrumento? Ou sua música é boa por que o instrumento é bem feito e o som gerado por este é agradável? Ou os dois? Essas perguntas também podem ser feitas com respeito a voz. Para entendermos o que influencia na qualidade da síntese da voz é necessário analisar a biomecânica da voz, que nada mais é analisar o movimento do material vivo e as forças atuantes sobre ele [23].

2.1.1 Fatores Biológicos que Afetam a Produção de Som na Laringe

A parte da (bio)mecânica que se relaciona diretamente com a atuação da laringe na produção do som é a mecânica dos meios contínuos, que é a parte da mecânica que lida com a matéria distribuída sobre uma determinada região no espaço, e conseqüentemente, se contrapõe à mecânica de partículas. Dentro da mecânica de meios contínuos, mais

especificamente, a parte que irá nos auxiliar no estudo do comportamento da laringe se chama mecânica de sólidos e fluidos.

Dito isto, analisaremos a seguir alguns conceitos físicos que tem forte ligação com os processos que ocorrem na região da laringe durante a produção do som:

- **Tensão e Deformação:** São características de forças atuantes em superfícies, como por exemplo a resistência do ar.

Tensão é quantidade de força por unidade de área [23], podemos escrever na forma da equação 2.1:

$$\sigma = f/A \quad (2.1)$$

Onde σ é a tensão, f é a força aplicada e A a área de aplicação desta força.

Deformação é a medida de deformação de um meio após a aplicação de uma tensão [23] e pode ser escrito na forma da equação 2.2:

$$\epsilon = (L - L_0/L_0) \quad (2.2)$$

Onde ϵ é a medida de deformação, L é o comprimento após a tensão e L_0 é o comprimento antes da tensão.

Normalmente uma dada deformação em uma dimensão resulta em uma deformação oposta em outra dimensão em um dado meio. Se uma deformação é uniforme por todo o corpo de um objeto, então chamamos de *compressão*, se o volume diminui por conta desta deformação, e *expansão*, se o volume aumenta.

- **Viscosidade:** É a velocidade de deformação(consequentemente, de restauração) de um determinado fluido quando atuam forças de tensão no mesmo. Matematicamente pode ser expresso conforme a equação 2.3:

$$\sigma = \eta \frac{d\epsilon}{dt} \quad (2.3)$$

Para η viscosidade e t tempo.

Quanto maior a viscosidade, mais devagar será a deformação de um meio.

- **Elasticidade:** É uma propriedade do meio que determina quão completa será a restauração do meio após uma dada deformação.

Os conceitos e propriedades descritos acima são extremamente importantes para se entender a manutenção da produção do som. Como dito anteriormente no capítulo 1, as pregas vocais são músculos e músculos são compostos por fibras, logo, as pregas vocais consistem de uma grande concentração de fibras. Além disso, entre as fibras que compõe as pregas vocais existem também fluidos atuantes, o que caracteriza as pregas vocais como um material viscoelástico. Para se entender a capacidade de absorção e regeneração das pregas vocais, em detrimento das vibrações de alta frequência e as pressões do ar, deve-se primeiramente estudar as propriedades absorcivas do material que as compõe. Ou seja, em outras palavras, deve-se estudar as propriedades mecânicas do tecido viscoelástico, e

uma ferramenta que facilita o entendimento é o estudo da *curva força-alongamento* de um material.

Entretanto, construir uma curva de *força-alongamento* depende essencialmente da geometria da amostra do material [23] e, por se tratar de uma material biológico, é difícil obter uma geometria precisa pois as fibras estão constantemente se reorientando em detrimento de lesões e cortes fibrosos. Para viabilizar este estudo, Titze [23] sugere normalizar as forças atuantes e as deformações resultantes para que não haja a dependência direta da geometria. Essa normalização se dá através da substituição da curva *força-alongamento* por um curva *tensão-deformação*.

A Figura 2.1 retirada do estudo feito por Titze [23] demonstra uma curva hipotética de *tensão-deformação* para os tecidos que compõe as pregas vocais humanas. Esta Figura ilustra o comportamento das fibras das pregas vocais através da relação entre uma força atuante e a deformação gerada por esta. A importância desta análise se deve ao fato de que é possível estabelecer uma relação direta entre nódulos vocais e uma fonação prolongada, alta(em termos de frequência) e intensa.

A partir da análise desta curva é então possível estabelecer um precedente para a formação de nódulos vocais: a frequência e amplitude da vibração estão diretamente ligadas ao surgimento de um nódulo vocal e conseqüentemente o de uma fenda pois a força de impacto entre as pregas vocais é proporcional à altura tonal quando acima do tom natural e à intensidade durante a fonação.

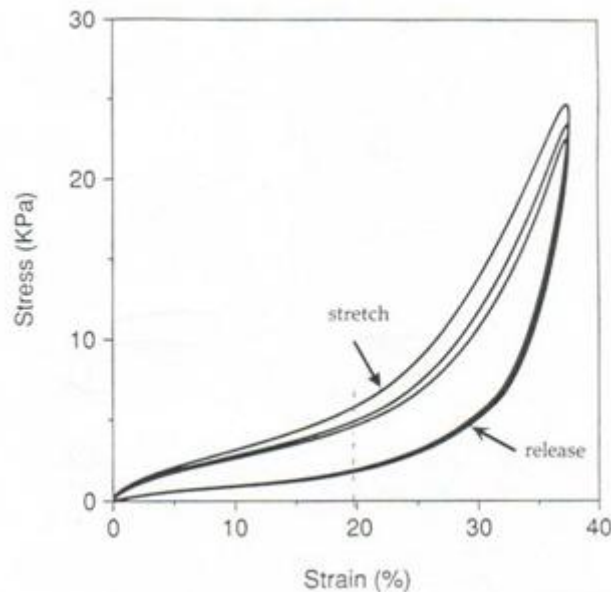


Figura 2.1: Curva Hipotética Tensão-Deformação das Cordas vocais Humanas [23]

Para confirmar o precedente supracitado, tomemos a segunda lei de Newton:

$$f = \frac{\textit{impeto}}{\textit{periodo}} \quad (2.4)$$

Isto pode ser traduzido em termos de proporção:

$$f \propto \frac{mv_m}{T} \quad (2.5)$$

Onde m é a massa do tecido das pregas vocais, T é o período de vibração. A velocidade máxima (v_m) é proporcional à máxima amplitude e inversamente proporcional ao tempo de fechamento das pregas vocais. Ou seja:

$$v_m \propto \frac{A}{T} \quad (2.6)$$

Em conjunto com a equação 2.5, e sabendo que a frequência é o inverso do período, temos que:

$$f \propto mAF_0^2 \quad (2.7)$$

O que biomecanicamente corrobora para a afirmação de que a força de impacto é proporcional à amplitude e ao quadrado da frequência.

2.2 Fluxo de Ar nas Vias Respiratórias

Em seções posteriores serão apresentados modelos para representar a vibração das pregas vocais, entretanto para entender por completo esses modelos, devemos antes entender o papel do fluxo de ar nesse processo e para isso devemos analisar primeiro como a pressão de ar é gerada e exercida.

À primeira vista, imagine-se que o fluxo de ar é gerado somente pelos pulmões, entretanto, se formos analisar como inspiramos e expiramos percebemos que existem outros fatores contribuintes para a criação e manutenção de um fluxo de ar nas vias respiratórias. Os pulmões, como é de se imaginar, é responsável pelo armazenamento do ar e pela troca de oxigênio entre os tecidos que o compõe. Aqui, porém, somente nos interessa esse papel de armazenamento que estes exercem. Ao inspirar o ar, há um aumento da pressão interna gerando uma tensão aplicada ao tecido elástico que compõe os pulmões. A expansão desse tecido gera uma mudança de pressão devido ao aumento do volume pulmonar e à deformação do tecido pulmonar. Em conjunto com essa mudança de pressão, o *diafragma*, músculo localizado abaixo dos pulmões, é contraído levando a uma outra mudança de pressão. O diafragma, por sua vez, é um músculo que pode ser controlado, e em virtude disso, é possível controlar a pressão exercida nos pulmões, algo que cantores utilizam para sustentar a técnica vocal. Assim como o diafragma, a caixa torácica também sofre mudanças, expandindo, o que gera um aumento no volume dos pulmões e consequentemente uma maior pressão torácica sobre os pulmões. Devido ao fato de ser uma estrutura óssea e portanto ter uma maior rigidez e menor controle se comparado à parede abdominal, a pressão exercida nos pulmões pela caixa torácica é significativamente menor do que a pressão exercida pela aplicação de uma força à parede abdominal.

A Figura 2.2 ilustra essa transmissão de pressão por todo o torso citada acima:

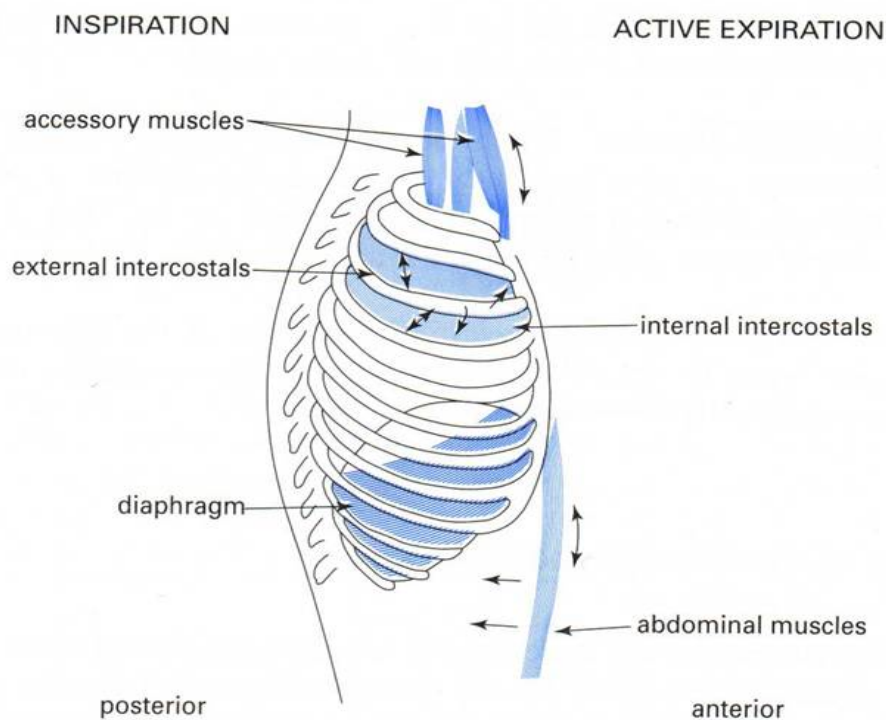


Figura 2.2: Pressões Atuantes no Torso [27]

2.2.1 O Sistema Pulmonar

O sistema pulmonar consiste dos pulmões e vias respiratórias, compostas pela traqueia, glote e trato vocal. Conforme descrito anteriormente, o tórax e o abdômen em conjunto com o pulmão atuam na geração do fluxo de ar. A glote, por sua vez, atua como reguladora do fluxo de ar através de variações em seu fechamento que permitem mudanças na pressão envolvida com o movimento do fluxo do ar pelas vias respiratórias. Isso significa que é possível dar constância ao fluxo de ar (Figura 2.3).

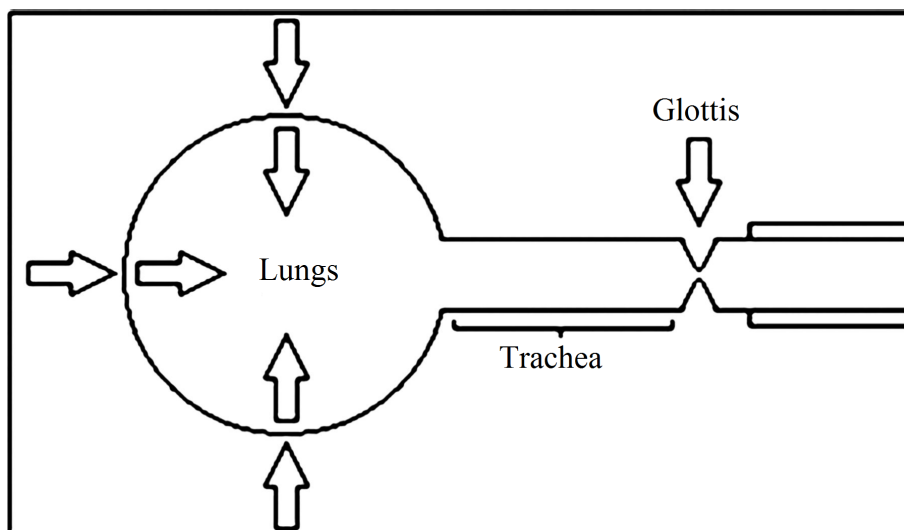


Figura 2.3: Sistema Pulmonar [7]

O Processo Físico da Respiração

Uma importante lei que nos auxilia no entendimento da relação entre o volume pulmonar e a pressão pulmonar é a Lei de Boyle. A lei nos diz que em um ambiente cujas paredes não são rígidas e em uma temperatura constante, pressão e volume são inversamente proporcionais.

Se aplicarmos a lógica dessa lei à respiração humana, fica fácil visualizar o processo físico que envolve a respiração. Quando inspiramos, o diafragma é contraído o que aumenta o volume do pulmão e diminui a pressão interna. Conseqüentemente, o pulmão se enche de ar. Relacionada à pressão atmosférica, a pressão pulmonar se torna menor, o que gera uma busca pelo equilíbrio das pressões no meio. Esse equilíbrio é alcançado ao se expelir o ar, diminuindo o volume pulmonar e então aumentando a pressão interna novamente. Na verdade é um pouco mais complexo do que isso, entretanto, esse conceito de equilíbrio de pressões é fundamental para o entendimento da abdução e adução das pregas vocais, o que permite a transformação da energia aerodinâmica em energia acústica.

2.2.2 Leis de Conservação para Fluxos em Dutos

Porém, antes de abordar a transformação de energia aerodinâmica em acústica, é importante se familiarizar com alguns conceitos que regem o comportamento de fluidos em dutos. Isso porque trataremos o fluxo de ar na síntese de voz como um fluido e a modelagem da glote e do trato vocal como a concatenação de pequenos dutos.

Lei da Continuidade para um Fluxo Incompressível

Quando dizemos que um fluido é incompressível significa dizer que sua densidade não se altera quando este é forçado a passar por uma restrição. Imagine agora um fluido confinado a um duto que possui uma mudança, ao longo de seu comprimento, em sua área transversal formando uma restrição (Figura 2.4). Se não for permitido ao fluido vazar pelas paredes do duto, então todas as partículas do fluido devem ser mantidas mesmo durante a mudança de área transversal ao longo do duto. Para que isso seja possível, as partículas devem acelerar durante a restrição, mantendo-se constante o número de partículas em movimento por unidade de área. Essa relação pode ser expressa pela equação 2.8:

$$v_1 A_1 = v_2 A_2 = \text{constant} = U \quad (2.8)$$

Onde v e A são respectivamente a velocidade da partícula e a área em que se encontra. U é o fluxo. Matematicamente então, a Lei da Continuidade expressa que um fluxo incompressível em um duto é constante, independente do que acontece com a área transversal ao longo do mesmo [23].

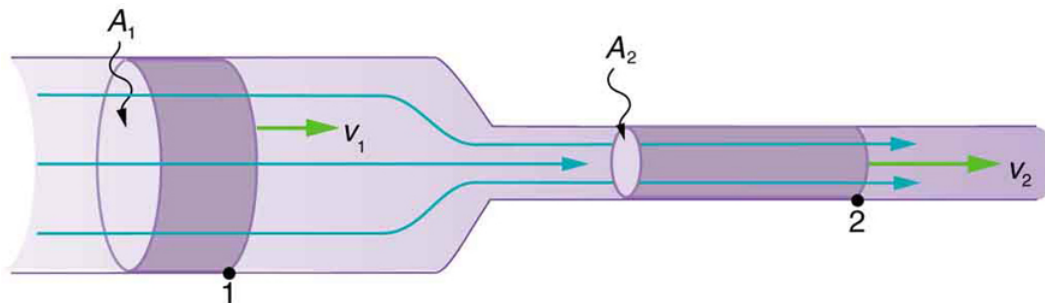


Figura 2.4: Dinâmica do Fluido em um Duto com Mudança de Área Transversal [7]

Lei de Bernoulli para Conservação de Energia

Outra importante lei de aplicação para o entendimento da síntese de voz é a Lei de Bernoulli para Conservação de Energia. Essa lei foi desenvolvida a partir do reconhecimento de que a energia total atuante no fluido em qualquer ponto durante o trajeto no duto tem duas componentes, a saber, uma energia potencial e uma energia cinética. A energia potencial está relacionada diretamente com a pressão no duto e a energia cinética

é proporcional ao quadrado da velocidade da partícula. Portanto, podemos expressá-la conforme a equação 2.9:

$$P + \frac{1}{2}\rho v^2 = \text{constante} \quad (2.9)$$

Onde ρ é a densidade do fluido, v a velocidade da partícula e P a pressão no duto. Segue disso o *Princípio de Bernoulli*: Se a energia em uma corrente de fluido é constante, um aumento na velocidade da partícula deve ser acompanhado de uma queda na pressão [23].

Analisando a Figura 2.4, a pressão na área de restrição deve ser menor que a pressão na área maior do duto, considerando que não hajam perdas de energia no processo. Esse conceito é importante para se compreender o caráter oscilatório das pregas vocais.

Resistência Glotal

A resistência a um fluxo é uma característica do sistema de transporte, ou seja, do meio, e pode ser descrita como a razão entre pressão e fluxo (equação 2.10)

$$R = \frac{P}{U} \quad (2.10)$$

Em sistemas de transporte, as restrições atuam como ponto de resistência a um fluxo. No caso das vias respiratórias humanas, isso normalmente ocorre na glote ou em alguma parte do trato vocal com pouco espaço. Podemos então definir que a *resistência glotal* será a pressão na glote dividida pelo fluxo que a atravessa. A resistência glotal está diretamente ligada a qualidade vocal pois é responsável por auxiliar no controle do fluxo de ar.

2.3 Oscilação das Pregas Vocais

Nesta seção analisaremos os processos e princípios físicos que ocorrem na laringe para que as pregas vocais oscilem, tendo como objetivo destrinchar tais processos a fim de entendermos a reprodução dos mesmos por modelos computacionais utilizados para simulação da voz humana.

Os primeiros estudos sobre vibração das pregas vocais especulavam que as pregas vocais se ajuntavam por um efeito de pressão negativa (Princípio de Bernoulli) na glote. A esta descrição de vibração nas pregas vocais se deu o nome de *Teoria Mioelástica-Aerodinâmica da Vibração das Pregas Vocais*, conhecida também pelo nome de seu criador como *Teoria de van der Berg* [2]. A teoria de van der Berg serviu como marco para o desenvolvimento teórico do campo de estudos da voz e conseqüentemente o aparecimento dos primeiros modelos matemáticos de representação do funcionamento das pregas vocais [5].

Entretanto, apesar disso, sua teoria é inadequada para explicar a vibração auto-sustentável das pregas vocais. Isso porque as forças de Bernoulli por si só não são capazes de distinguir entre os movimentos interiores e posteriores das pregas vocais. Assim sendo, são necessários mecanismos para prover um aumento ou decréscimo às forças de Bernoulli

durante a abertura e o fechamento das pregas vocais respectivamente, caso contrário, as oscilações serão amortecidas.

2.3.1 Critérios para Oscilação

Alguns critérios devem ser atendidos para que um determinado padrão de movimento seja considerado como uma oscilação mecânica, a saber:

- No sistema onde ocorre o movimento deve haver uma posição de equilíbrio estável, que é caracterizada por uma força restaurativa que sempre acelera o corpo em movimento de volta para a sua posição de repouso.
- Deve haver inércia (no caso do sistema mecânico, a massa atua como propriedade de inércia) no sistema para superar esta posição de equilíbrio.
- A perda, em excesso, de energia por ciclo de oscilação deve ser zero.

2.3.2 Tipos de Oscilação

Ingo R. Titze, descreve uma terminologia auxiliar para a caracterização de movimentos oscilatórios [23]. A saber:

- *Oscilação Natural*: Quando um sistema que se encaixa nos critérios anteriores se move sem interferência após um distúrbio inicial.
- *Oscilação Forçada*: Requer uma fonte externa de condução que por si só é um oscilador. Dita grande parte do padrão de vibração do sistema.
- *Oscilação Auto-Sustentável*: Requer uma fonte de energia estável e uma interação não-linear entre os componentes internos ao sistema. As perdas de energia são compensadas, mantendo o padrão oscilatório.

Algumas observações importantes referentes aos tipos de oscilação seguem:

- No caso de uma oscilação auto-sustentável, quando uma força média aplicada ao ciclo tem a mesma direção da velocidade, energia é transmitida ao oscilador. A oscilação pode ser mantida se esse acréscimo de energia for pelo menos tão grande quanto às perdas por fricção durante o ciclo do sistema.
- Quando há sincronia entre a oscilação forçada e a oscilação natural ocorre o que conhecemos como ressonância. A ressonância é utilizada para se obter a máxima resposta com o menor estímulo possível em um sistema oscilatório.

2.3.3 Oscilador Massa-Mola

É um oscilador montado sobre um sistema elástico com uma massa acoplada a uma mola. Como veremos mais adiante, este sistema pode ser mono-massa ou multi-massa. A força de restauração deste sistema resulta do alongamento ou da compressão da mola e essa força é sempre oposta ao deslocamento x , sempre buscando retornar ao ponto de equilíbrio. Aqui a inércia da massa causa uma superação da posição de equilíbrio, produzindo a oscilação (Figura 2.5).

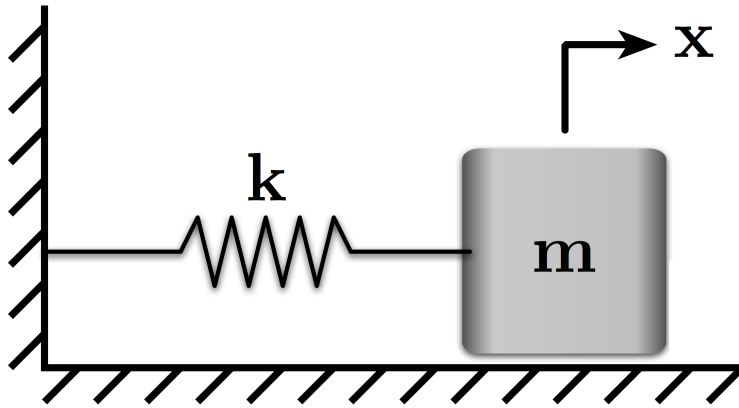


Figura 2.5: Um Oscilador Massa-Mola [7]

A frequência de oscilação F_0 é definida como o número de movimentos ida e volta que são feitos em um segundo. Naturalmente, essa frequência pode ser calculada pela equação 2.11,

$$F_0 = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (2.11)$$

Onde k é a rigidez da mola e m sua massa. Pela equação podemos notar que quanto maior a rigidez, maior a frequência, ao passo que, quanto maior a massa menor será a frequência de oscilação. Ou seja, rigidez e massa são quantidades chaves na determinação de F_0 [23]. A curva produzida por um oscilador é de característica senoidal e portanto o seu movimento é governado por três métricas, a saber: a amplitude, a fase e a frequência. A amplitude escala a forma de onda verticalmente, a fase desloca a forma de onda para a esquerda ou para a direita e a frequência escala-a horizontalmente.

A seguir no próximo capítulo, veremos como representar e sintetizar a voz computacionalmente. Para isso analisaremos os conceitos entorno dos modelos massa-mola-amortecedor [5, 8]. Veremos como esses modelos possibilitam a representação do mecanismo de vibração auto-sustentável das pregas vocais. Entretanto, primeiramente, deve-se entender como o som é gerado e como este se propaga em um meio.

2.4 Geração e Propagação do Som

Fisicamente, o som é gerado quando há uma perturbação do equilíbrio da densidade (ou pressão) de um gás, líquido ou sólido [23]. Se a perturbação na pressão é positiva, então haverá uma condensação no meio e portanto um aumento na densidade do ar. Se a perturbação for negativa, haverá uma rarefação no meio o que ocasionará um decréscimo na densidade do ar. Considere um pistão em um tubo cilíndrico preenchido de ar. Pelo movimento oscilatório do pistão, o ar tende a condensar ou rarefazer. Sendo assim, as partículas em contato com o pistão tendem a seguir o movimento oscilatório do mesmo, porém, por não estarem rigidamente conectadas nem à superfície do pistão nem entre si mesmas, elas estabelecem um padrão próprio de movimento em detrimento das colisões. Dessa forma, quantidade de movimento é transmitida às partículas adjacentes, criando

uma nova perturbação em um ponto mais longe ao pistão. A este fenômeno se dá o nome de propagação. Da mesma maneira, acontece com a propagação do som na glote.

Podemos pensar a respeito da **geração** do som na glote com a seguinte analogia: o movimento das pregas vocais pode ser comparado ao movimento de bater palmas, a cada choque das mãos, uma perturbação é gerada e conseqüentemente, som é gerado. Se pensarmos no papel da glote na geração do som, podemos pensar que a glote é como se fosse uma sirene, o fluxo de ar que passa por ela é espremido no canal o que leva a uma perturbação e conseqüentemente a geração do som.

A partir do conceito de propagação de ondas sonoras é possível deduzir o conceito de reflexão sonora, que está ligado à propagação da onda em um meio com propriedades de impedância acústica. Devido a estas propriedades de impedância, a propagação da onda ao entrar em contato com a superfície do meio, é alterada.

2.4.1 Reflexão do Som

A reflexão está diretamente ligada com as propriedades de amortecimento e rigidez do meio. Um meio muito rígido ou denso provoca reflexões mais dramáticas quando em contato com a onda sonora. Para efeitos de análise considere que este meio muito rígido(meio 2) possui um coeficiente de impedância de onda aproximando o infinito. Neste caso a pressão na superfície do meio 2 pela coluna de ar(meio 1) se aproxima do dobro de seu valor inicial simplesmente porque as partículas do ar não conseguem atravessar o meio 2 e portanto se aglomeram na superfície. Em decorrência disso, as partículas do ar em transição se ajuntam mais, dobrando também a densidade de partículas no ar. Essa nova perturbação no meio 1 criará uma nova onda de pressão superimposta à onda inicial. De acordo com Titze [23], a propagação é sempre em direção oposta à fonte, no caso é na direção contrária à coluna de ar(meio 1). Portanto, haverão duas ondas no mesmo meio com movimentos cujas direções são opostas.

Podemos então descrever a pressão espacial e temporal total como a soma das pressões de incidência e reflexão, logo temos a eq 2.12:

$$p = p_i + p_r \quad (2.12)$$

Consideraremos a polarização da pressão que causa condensação como positiva, ou seja, excesso de pressão. Portanto a polarização de rarefação será negativa, ou seja, redução de pressão. Analogamente ao meio muito rígido, podemos considerar um meio cuja densidade se aproxima à zero(meio 2), para critérios de análise. O mesmo que ocorreu para o meio rígido ocorrerá neste, entretanto, ao invés de causar condensação devido a uma impenetrabilidade do meio, neste caso ocorrerá uma rarefação. Tendo em vista que o coeficiente deste meio é próximo a zero, as partículas de ar não encontram resistência nenhuma neste meio e rapidamente o atravessam. O excesso de pressão contido no meio 1(coluna de ar) é rapidamente exaurido em decorrência disso. Isso faz com que uma perturbação seja gerada neste meio, uma pressão negativa, ou seja, uma rarefação. Essa pressão negativa(de reflexão) se superimpõe à onda de pressão de incidência, com sua direção oposta.

Estes dois exemplos servem para ressaltar que a reflexão de onda ocorre tanto para meios com alta rigidez quanto para baixa rigidez ou densidade.

Coefficiente de Reflexão

Diferentemente do mencionado anteriormente, na maioria das situações físicas o meio de reflexão não é nem infinitamente denso e nem infinitamente esparso. A impedância de um meio geralmente é finita e maior do que zero, o que significa que somente uma fração da pressão de incidência é refletida. Isso também significa que parte da energia acústica é transmitida ao meio de reflexão. Um coeficiente de reflexão na superfície então pode ser definido como a fração refletida da onda incidente, conforme eq 2.13:

$$r = \frac{p_r}{p_i} \quad (2.13)$$

A pressão e o fluxo das partículas deve ser contínuo [23], essa condição permite que o coeficiente de reflexão seja calculado de acordo com dois parâmetros de impedância ρ e c , sendo assim temos a eq 2.14 :

$$r = \frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_1 c_1} \quad (2.14)$$

Note que caso o meio 2 seja extremamente denso, o coeficiente tende a +1 e, caso contrário, tende a -1. Para qualquer outro caso o coeficiente será um número entre -1 e +1. Quando não há reflexão, ou seja, quando os dois meios possuírem o mesmo coeficiente de impedância, o coeficiente de reflexão será 0.

Dessa forma, pode-se notar que o coeficiente de reflexão é útil para se obter tanto a magnitude como a polaridade da onda refletida. A direção desta nós já sabemos, é sempre oposta à de incidência. Além disso, caso haja uma reflexão parcial, como é o que ocorre na maioria dos casos, o coeficiente de reflexão também define a onda transmitida, cuja pressão é equivalente à soma das pressões das ondas de incidência e reflexão.

O resultado disto será explicado quando mencionarmos o modelo reflexão de onda para representação do trato vocal, no Capítulo 3. Será visto que o coeficiente de reflexão é útil em descrever a ressonância na cavidade oral do trato vocal.

Capítulo 3

Computação e Síntese de Voz

Neste capítulo serão tratados os modelos computacionais e os conceitos matemáticos que possibilitaram o desenvolvimento do SynPath.

3.1 Modelos Computacionais para Representação das Pregas Vocais

No trabalho pioneiro de Flanagan e Landgraf [5], foi proposta uma fonte de auto-oscilação para simular computacionalmente o comportamento das pregas vocais humanas. O modelo utilizado para representar as pregas vocais era um modelo mecânico simples de segunda-ordem, logo, podendo ser representado por um simples sistema mecânico, um oscilador com uma massa, uma mola e um amortecedor, chamado sistema mono massa-mola-amortecedor. Posteriormente, Ishizaka e Flanagan propuseram em seu trabalho [8] uma extensão do modelo mono massa-mola-amortecedor para acoplar duas massas ao modelo a fim de se obter maior fidelidade ao comportamento das pregas vocais humanas. Como o Synpath foi desenvolvido sobre o modelo mono-massa, analisar-se-á este a seguir. Caso tenha interesse a respeito do modelo multi-massa, é interessante buscar os trabalhos de Titze [23], Ishizaka *et al* [8, 9], Erath [4] e etc.

3.1.1 Modelo Mono-Massa

Os modelos de uma massa são os mais simples para se representar as pregas vocais, com essa única massa representando toda a estrutura das pregas vocais. Como é representado por uma massa somente, o grau de liberdade neste sistema é 1, ou seja, o deslocamento é dado somente em uma direção [4]. Conforme foi citado no Capítulo 2, para um modelo viscoelástico das pregas vocais apresentar auto-oscilação, energia deve ser periodicamente adicionada ao sistema de modo que supere as fricções internas do tecido laríngeo. Tal energia é acrescida ao sistema por uma carga externa quando a mesma está em fase com a velocidade de vibração das pregas vocais. A transferência de energia entre o fluxo e as vibrações estruturais é dada por uma carga acústica (pressões acústicas) em conjunto com o fluxo e a resistência glotal variante no tempo [17, 26]. Mais especificamente, a auto-oscilação somente ocorrerá na presença de uma carga acústica ou algum outro mecanismo de imposição de cargas, como mostrado por McGowan e Howe [18].

A modificação na pressão aerodinâmica devido às mudanças de orientação convergente-divergente da glote é o mecanismo primário pelo qual o supracitado ocorre fisiologicamente. Por definição, os sistemas mono-massa não podem desenvolver por si só a variação (no tempo) glotal necessária para que as pressões supraglotal e glotal sejam carregadas assimetricamente, permitindo que o efeito de Bernoulli aja em reconhecimento à abertura e fechamento da glote. Para que isso ocorra neste sistema, deve haver uma interação da glote com um tubo acústico supraglotal ou subglotal [23].

Considere agora o sistema representando cada prega vocal, cujos componentes são uma massa m , uma constante de rigidez k e uma constante de amortecimento b . A constante de amortecimento representa a viscosidade do tecido, ou seja, atua como um absorvedor de energia. A constante k representa as propriedades elásticas do tecido ao passo que m é a massa do tecido em movimento, Figura 3.1 [23].

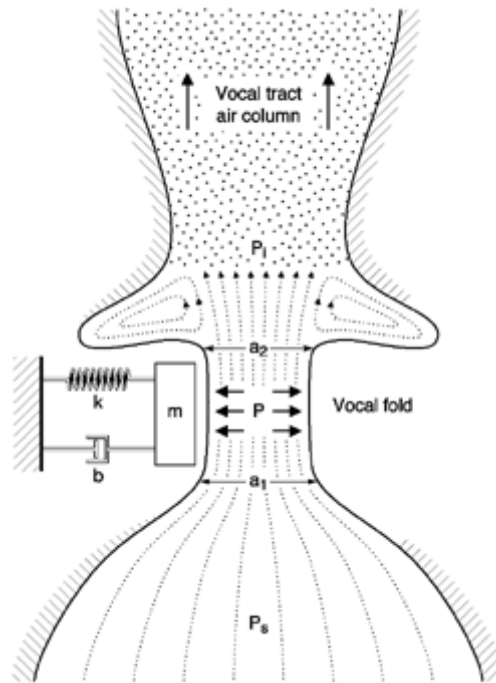


Figura 3.1: Modelo Mono-Massa [23]

A pressão P na glote atua perpendicular à superfície do tecido. Se esta pressão se alterar de acordo com a direção da velocidade, então energia será transmitida ao tecido pelo fluido, e esta pode superar a energia perdida com a viscosidade do tecido. Entretanto, como essa pressão pode ser diferente ou assimétrica no caminho de volta do que foi ao sair? Ou seja, como a pressão pode ser alterada e conseqüentemente acompanhar e influenciar a abertura e o fechamento da glote?

Titze derivou uma expressão para a pressão intraglotal na superfície das pregas vocais [26] e forneceu uma versão simplificada dessa expressão para o modelo em questão [23], conforme mostra a equação 3.1:

$$P = \left(1 - \frac{a_2}{a_1}\right)(P_s - P_i) + P_i \quad (3.1)$$

Onde a_1 e a_2 são as áreas de entrada e saída da glote respectivamente, P_s é a pressão subglótica e P_i é a pressão exercida sobre o trato vocal, chamada de pressão de *input*. O primeiro termo da equação é um fator geométrico que descreve o formato da glote. O segundo termo, a diferença de pressões, descreve a pressão transglotal, ou seja, aquela que percorre a glote. Para o modelo mono-massa, temos que $a_1 = a_2$ o que faz com que a pressão intraglotal seja simplesmente equivalente à pressão supraglotal.

A relação resultante de que a pressão condutora P é igual à pressão supraglotal sugere que algo deve acontecer **acima da glote** para que essa pressão seja alterada durante o ciclo glotal. O elemento chave aqui é a inércia do ar no trato vocal. A lentidão em resposta da coluna de ar acima das pregas vocais causa uma outra condição de superação que auxilia na vibração, isto porque quando a glote está abrindo e o fluxo de ar aumentando, a coluna de ar está sendo acelerada pelo fluxo glotal. Isto cria uma pressão positiva P_i na entrada do trato vocal, fazendo com que as pregas vocais se separem. Essa pressão positiva gera também um aumento no ímpeto(ou na quantidade de movimento) da coluna de ar. Quando a glote fecha, o ímpeto da coluna de ar continua e o fluxo na glote não se sustenta com o fluxo da coluna de ar, gerando uma pressão negativa(sucção) acima das pregas vocais auxiliando no fechamento das pregas vocais. Portanto, P_i conduz as pregas vocais em sincronia com seu movimento natural.

Essa explicação é muito similar ao efeito de Bernoulli em ação, entretanto, um termo adicional muito importante está envolvido: o atraso na resposta à pressão da coluna de ar no trato vocal.

Dessa maneira, fica claro que em decorrência do grau de liberdade do modelo mono-massa, não há como este modelo reproduzir uma movimentação ondular das pregas vocais, o que auxiliaria na criação e manutenção da auto-oscilação. Em decorrência disso, o trato vocal atua suprindo esta falta de modo que o mesmo possibilita a alteração nas pressões durante as fases de movimento das pregas vocais o que induz uma auto-oscilação no sistema mono-massa, conforme explicado.

3.2 Modelo Computacional para Representação do Trato Vocal

Para a representação do trato vocal, Fraj [19], Titze [24], entre outros [21, 22] descrevem um modelo físico, onde o trato vocal é representado como uma concatenação de pequenos tubos cilíndricos de tamanhos variados e conseqüentemente de áreas diferentes. As condições de continuidade da pressão acústica e da velocidade do fluido em cada junção de tubos permite o mapeamento das ondas acústicas de pressão e de velocidade de entrada com as de saída. Aplicando isso iterativamente, obtêm-se a simulação da propagação da onda no trato vocal. Levando em consideração que o tempo de propagação em um duto equivale à uma amostra no sistema, é possível aproximar o tamanho necessário dos tubos cilíndricos de acordo com a frequência da amostra e a velocidade do som.

Existem dois modelos para a representação do trato vocal como um concatenação de tubos cilíndricos. O primeiro, foco deste trabalho, é chamado de modelo de reflexão de onda [3, 19]. O segundo, caso haja interesse em pesquisa, é chamado de modelo linha de transmissão e pode ser também encontrado no livro de Titze [24].

3.2.1 Modelo Reflexão de Onda

Este modelo tem seus fundamentos na reflexão das ondas sonoras no trato vocal, conforme citado no Capítulo 2. Enquanto a glote produz uma onda acústica com muitas frequências, o trato vocal amplifica um subconjunto dessas frequências para radiar pela boca. O filtro desses subconjuntos está diretamente ligado às características de ressonância e reflexão do trato vocal, e é isto que este modelo busca implementar.

O modelo tem seu fundamento nos conceitos de reflexão e impedância acústica, ambos descritos no Capítulo 2. Entretanto, precisaremos adequar aqueles conceitos ao uso de tubos cilíndricos para a representação do trato vocal.

Impedância Acústica em um Tubo

Se a velocidade média de uma partícula for multiplicada pela área transversal do tubo, um fluxo médio pode ser obtido. Esse fluxo médio é conservado quando um tubo se expande ou contrai. Dessa maneira, podemos definir a impedância acústica como a razão entre a pressão acústica p e o fluxo acústico u . Logo temos a eq 3.2:

$$Z = \frac{p}{u} \quad (3.2)$$

Se não há reflexões no tubo, a razão entre pressão e fluxo adota a seguinte forma:

$$Z = \frac{\rho c}{A} \quad (3.3)$$

Onde ρ é a densidade do ar, c é a velocidade do som e A é a área transversal do tubo. No Capítulo 2, na seção de reflexão do som, foi definido uma fórmula para a impedância acústica livre como ρc . A partir disso, é fácil notar que o coeficiente de impedância acústica no tubo nada mais é do que o coeficiente de impedância acústica livre dividido pela área transversal do tubo.

Dessa maneira, temos três fatores que podem alterar a propagação das ondas acústicas:

- Uma mudança na velocidade do som
- Uma mudança na densidade do ar
- Uma mudança na área transversal

Qualquer uma dessas mudanças causarão uma reflexão.

Coefficiente de Reflexão nos Tubos

Como dito no início dessa seção, computacionalmente o trato vocal é representado como uma concatenação de tubos cilíndricos em série. Isso permite a modelagem de qualquer forma do trato vocal, vide Figura 3.2, já que a complexidade da forma é subdividida em figuras relativamente simples. Com base no conceito definido para impedância acústica em tubos, podemos definir um coeficiente de reflexão para a interface de transição entre um tubo e o outro, como mostra a eq 3.4:

$$r = \frac{\frac{\rho_2 c_2}{A_2} - \frac{\rho_1 c_1}{A_1}}{\frac{\rho_2 c_2}{A_2} + \frac{\rho_1 c_1}{A_1}} \quad (3.4)$$

Onde os subscritos 1 e 2 representam dois tubos adjacentes.

Se considerarmos a densidade do ar e a velocidade do som como constantes durante o trajeto nas vias respiratórias, temos que $\rho_2 = \rho_1$ e $c_2 = c_1$. Logo a fórmula acima pode ser simplificada algebricamente para a eq 3.5:

$$r = \frac{A_1 - A_2}{A_1 + A_2} \quad (3.5)$$

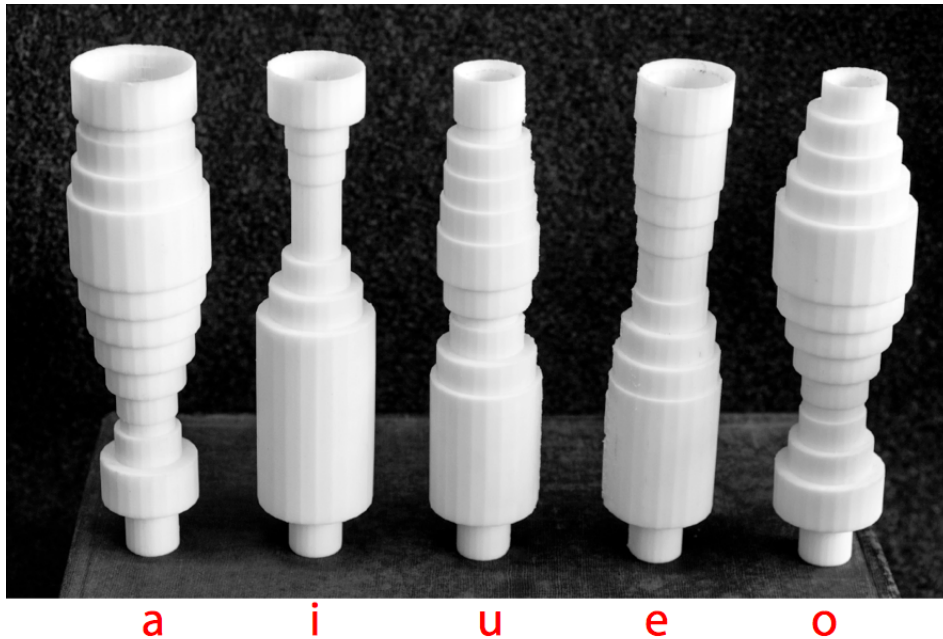


Figura 3.2: Modelagem do trato vocal para diferentes vogais [7]

O resultado expresso pela eq 3.5 é um resultado muito importante para a síntese, pois indica que o coeficiente de reflexão entre dois tubos adjacentes é a razão entre a diferença das áreas pela soma das áreas transversais destes tubos. A partir deste resultado, Titze [23] mostra que diferentemente da pressão aerodinâmica (Bernoulli), que reduz em uma região constrita, a pressão acústica aumenta em uma região constrita.

A partir disto, podemos modelar o trato vocal independente da forma que este venha a assumir e saberemos qual o comportamento da onda sonora no decorrer de seu trajeto no trato vocal.

Capítulo 4

SynPath

4.1 Descrição do Sistema

O SynPath é um sintetizador computacional, desenvolvido em linguagem Python [13, 20, 28], criado por Lucero [10]. Este software é uma extensão do sintetizador concebido por Fraj[19], incorporando um modelo de vibração para as pregas vocais. O seu propósito é aumentar a fidelidade fisiológica do sintetizador de Fraj e permitir o controle direto dos sons sintetizados em termos de parametrização da laringe.

Para se obter um simples controle sobre o sintetizador e facilitar o seu uso para aplicações práticas, é necessário que a representação das pregas vocais seja simples. Além disso, o modelo das pregas vocais deve garantir variações suaves no fluxo gerado na glote. A falta de suavidade gera timbres não naturais e conseqüentemente a perda da fidelidade fisiológica buscada. Sendo assim, o modelo multi-massa para representação das pregas vocais não pode ser utilizado visto que produz variações não suaves e é um modelo matematicamente muito complexo levando a instabilidades numéricas, afetando o uso para aplicações práticas.

O SynPath tomou como base para a representação das pregas vocais o modelo de onda mucosa desenvolvido por Titze [26]. Basicamente, o modelo é um oscilador mono-massa, conforme descrito no Capítulo 3, incorporando a transferência de energia do fluxo de ar para as pregas vocais. Entretanto, o modelo de Titze possui duas restrições, uma que foi solucionada e aplicada no desenvolvimento do SynPath e a outra que ainda não foi solucionada e portanto é também uma restrição do modelo computacional do SynPath.

A primeira restrição é que o modelo de Titze foi concebido para o estudo em pequenas oscilações, sendo que para oscilações de maior amplitude, este não é apropriado. Entretanto, esse modelo foi estendido por Lucero [11] para abranger maiores oscilações utilizando um mecanismo limitador de amplitude durante as oscilações [14].

Mesmo com essa extensão para maiores amplitudes, o modelo ainda apresentava uma outra restrição, um atraso pequeno para o deslocamento da onda no canal glotal. A consequência desse atraso é que a pressão limite para que ocorra a vibração se torna independente da frequência de vibração [15]. Porém, sabe-se que um esforço maior é necessário para que tons mais agudos sejam vocalizados, ou seja, a pressão para que se ocorra vibração em frequências maiores(tons maiores) é maior. Essa restrição ainda não foi solucionada tendo em vista a dificuldade de se realizar o supracitado computacionalmente. Mesmo com essa restrição, o modelo de Lucero [11] é o modelo utilizado para

a representação do caráter oscilatório das pregas vocais no software SynPath, sendo que esta restrição não solucionada não tem forte influência no produto final.

4.2 Concepção Matemática e Modelagem do Sistema

4.2.1 Trato Vocal

O trato vocal no SynPath é representado como uma sequência de tubos cilíndricos concatenados pelos quais se propaga uma onda acústica planar com componentes anteriores e posteriores [10]. Tanto o trato vocal subglótico quanto o trato vocal supraglótico estão incluídos na representação, sendo que no caso do trato supraglótico, ainda estão inclusos as cavidades nasais e os sinus paranasais. Neste modelo, as perdas por vibração, viscosidade do ar e condução térmica são consideradas.

A propagação da onda acústica foi resolvida numericamente para um valor de amostra de 88.2 kHz, o que fixa o comprimento dos tubos cilíndricos em 0,4 cm. Sendo assim, a traqueia foi modelada com 36 tubos cilíndricos cujas áreas transversais foram fixadas em $2,5 \text{ cm}^2$.

4.2.2 Pregas Vocais

O modelo de Titze [26] para representação das cordas vocais pode ser visto na Figura 4.1. Este modelo assume completa simetria entre as pregas vocais e somente permite o movimento horizontal das mesmas. As propriedades biomecânicas do tecido são aglomeradas no ponto médio da glote, levando à equação 4.1, de movimento das pregas vocais:

$$M\ddot{x} + B(1 + \eta x^2)\dot{x} + Kx = P_g \quad (4.1)$$

Onde x é o deslocamento do tecido, M, B e K são a massa, o amortecimento e a rigidez respectivamente por unidade de área da superfície das pregas vocais, η é o coeficiente de amortecimento não linear e P_g é a pressão intraglotal. Essa equação representa o movimento de cada prega vocal dentro do modelo estabelecido.

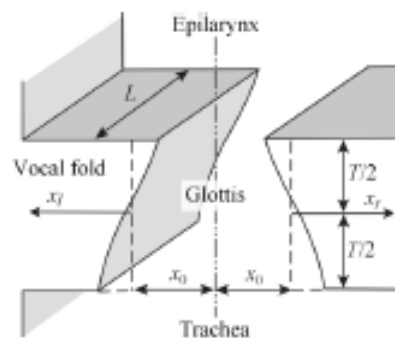


Figura 4.1: Esquema das Pregas Vocais [11]

Na abertura glotal, se nos atermos aos conceitos entre pressão e área do capítulo 2 e 3, a pressão intraglotal pode ser definida da seguinte forma:

$$P_g = P_i + \left(\frac{P_s - P_i}{k_t} \right) \frac{a_1 - a_2}{a_1}, \quad (a_1, a_2 > 0) \quad (4.2)$$

Onde P_s é a pressão subglotal, P_i é a pressão supraglotal, k_t é o coeficiente de pressão transglotal e a_1 e a_2 são as áreas transversais das bordas inferiores e superiores das cordas vocais, respectivamente.

Sendo assim, podemos definir as áreas de acordo com a seguinte equação 4.3:

$$a_{1,2}(t) = L[x_0 + x_r(t \pm \tau_r)] + L[x_0 + x_l(t \pm \tau_l)] \quad (4.3)$$

Onde L é a largura da prega vocal, x_0 é a metade do comprimento glotal quando as pregas vocais estão em estado de repouso e $\tau_{r,l}$ é o atraso no tempo para que a onda atravesse a altura glotal T . Os sinais de positivo e negativo correspondem a a_1 e a_2 respectivamente.

Para simplificar a equação, o fator que contempla a área glotal pode ser aproximado pela linearização

$$\frac{a_1 - a_2}{a_1} \approx \frac{1}{x_0} (\tau_r \dot{x}_r + \tau_l \dot{x}_l) \quad (4.4)$$

Essa linearização elimina o perigo da divisão por um número pequeno, como pode ocorrer na equação 4.2, em detrimento de uma pequena perda na precisão da representação.

Dessa maneira, temos que a área no centro da glote será:

$$a(t) = L[2x_0 + x_r(t) + x_l(t)] \quad (4.5)$$

A equação 4.5 produz uma transição não suave da área glotal nos limiares entre a abertura e o fechamento da glote. Para obter uma variação mais suave, Lucero [10] aplicou à função uma interpolação com um polinômio cúbico de Hermite. Essa transformação modela o comportamento de fechamento e abertura glótica mais próximo da fisiologia real, onde as transições ocorrem de forma menos abrupta.

4.2.3 Fluxo de Ar na Glote

O fluxo de ar na glote foi definido por Titze [25], em seu trabalho, na seguinte equação 4.6:

$$U_g = \pm \left(\frac{a_m c}{k_t} \right) \left\{ \frac{-a_m}{A^*} + \left[\left(\frac{a_m}{A^*} \right)^2 \pm \left(\frac{4k_t}{c^2 \rho} \right) (P_s^+ - P_i^-) \right]^{1/2} \right\} \quad (4.6)$$

Onde A^* é a área glotal efetiva computada pelas áreas A_i e A_s supra e subglotal respectivamente, ρ é densidade do ar, c é a velocidade do som e P_s^+ e P_i^- são as pressões de incidência na entrada e saída da glote respectivamente. Detalhes a cerca dos valores utilizados nos parâmetros das equações citadas aqui podem ser encontrados no trabalho de Lucero [10].

Uma vez descoberto o fluxo, as pressões de reflexão P_s^- e P_i^+ (opostas às de incidência) podem ser definidas pelas seguintes equações:

$$P_s^- = P_s^+ - (\rho c/A_s) U_g \quad (4.7)$$

$$P_i^+ = P_i^- - (\rho c/A_i) U_g \quad (4.8)$$

Quando ocorre o fechamento da glote, então $a(t) = 0$, $P_g = \frac{(P_s - P_i)}{2}$ e $U_g = 0$. As forças de contato, durante a colisão das pregas vocais, são modeladas com a adição de uma mola cujo coeficiente de rigidez é $K_{rc} = 3K_r$.

4.2.4 *Jitter* Vocal

Jitter é o nome dado a pequenas variações de curto prazo na frequência fundamental de uma onda sonora. Toda voz em seu timbre natural, seja ele qual for, possui *jitter*, sendo este uma característica fundamental para representar o realismo da fisiologia da voz. Logo, uma síntese de voz sem aplicação de *jitter* gera vozes com características mais artificiais. No SynPath há a preocupação de se aproximar o máximo possível de timbres naturais do ser humano, sendo necessário a implementação de *jitter* na síntese.

O *jitter* é implementado através da aplicação de uma perturbação estocástica aos coeficientes de rigidez. Isso torna possível a simulação das irregularidades das tensões musculares.

O seguinte termo estocástico é adicionado ao lado esquerdo da eq 4.1:

$$a_r \epsilon_r K_r x_r \quad (4.9)$$

Onde a_r é um coeficiente de escalabilidade e ϵ_r é o fator de ruído.

Com isso, a equação de movimento das pregas vocais se torna a seguinte:

$$M\ddot{x} + B(1 + \eta x^2)\dot{x} + K(1 + a\epsilon)x = P_g \quad (4.10)$$

O fator de ruído é um fator probabilístico e foi definido da seguinte maneira:

$$\epsilon_r = \begin{cases} +1 & p = 0,5 \\ -1 & p = 0,5 \end{cases} \quad (4.11)$$

Onde p é a probabilidade.

4.3 Requisitos do Sistema

4.3.1 Requisitos Funcionais

O SynPath é relativamente novo e de natureza científica, sendo assim, não possui ainda uma interface clara de interação com o usuário. Entretanto, é possível descrever as funcionalidades do sistema com base no que o mesmo oferece quando executado por um usuário. Ressalto que para o seu pleno funcionamento, os parâmetros que devem ser interpretados pelo sistema devem ser passados pelo usuário diretamente ao programa.

Tendo sido dito isto, podemos enumerar os requisitos funcionais deste sistema a seguir:

- **I.** Verificar os parâmetros passados pelo usuário, se condizem com as restrições do programa.
- **II.** Plotar um gráfico inicial do trato vocal de acordo com os parâmetros do usuário.
- **III.** Plotar três gráficos referentes as propriedades da voz simuladas com os parâmetros fornecidos pelo usuário.
 - O primeiro gráfico refere-se às posições adotadas pelas cordas vocais, a área da glótis, ao fluxo de ar nessa área e às características desse fluxo.
 - O segundo gráfico refere-se às características do som gerado pela simulação física do aparato fonador pelo programa.
 - O terceiro gráfico refere-se ao espectro de frequência do som gerado e do fluxo da glótis.
- **IV.** Gerar um arquivo de texto com as características da voz gerada, tais como a frequência, os aspectos físicos das cordas vocais, amplitude do som e do ruído na voz, entre outros.
- **V.** Gerar um arquivo de som contendo o áudio da voz simulada de acordo com as características exigidas.

4.3.2 Requisitos Não-Funcionais

Seguindo a lógica acima, a seguir se encontram destacados os requisitos ligados ao desempenho do sistema e não às suas funcionalidades:

- **I.** O sistema deve exibir os gráficos da simulação de voz dentro do tempo corrido de 1 minuto.
- **II.** O sistema deve gerar um arquivo texto e um arquivo de áudio exatamente após os gráficos terem sido exibidos.
- **III.** Para que o sistema seja executado corretamente, o usuário deve ter instalado os pacotes padrões da linguagem Python, juntamente com os pacotes de plotagem de gráficos matemáticos e calculos numéricos com matrizes a saber:
 - Matplotlib
 - NumPY

- **IV.** O sistema deve ser executado nas plataformas de sistema operacional Windows, Linux ou MacOS em suas mais recentes versões de acordo com a data deste documento.
- **V.** O sistema não necessita de atualizações e nem de conexão com a internet, sendo portanto classificado com um sistema integralmente offline.

4.3.3 Requisitos de Dependabilidade

Após um estudo sobre o sistema, alguns requisitos que não fazem parte nem das funcionalidades nem das propriedades do sistema, chamados requisitos de dependabilidade, foram estabelecidos a fim de que em breve o sistema possa garantir uma maior fidelidade aos seus usuários. A saber os requisitos:

- **I. Confiabilidade:** Para garantir que o diagnóstico e o estudo do comportamento da voz e suas patologias estejam dentro do espectro da correteza exigida dentro desta área de saúde, o sistema deve ser capaz de oferecer um serviço correto contínuo, de modo que o mesmo possa garantir que os gráficos, dados e arquivos de som fornecidos ao final da execução do programa sejam igualmente corretos dado uma segunda execução do programa com os mesmos parâmetros de entrada.
- **II. Segurança:** O sistema deve garantir que não haja consequências catastróficas no resultado esperado por qualquer tipo de interação entre o usuário e o sistema. Ou seja, deve-se garantir que o usuário não seja capaz de causar uma consequência grave por mérito próprio na interação dele com o sistema, em suas funcionalidades plenas.
- **III. Integridade:** O sistema deve agir da mesma maneira para todos os usuários em todo e qualquer tempo de execução. Isto significa que o sistema deve garantir a ausência de alterações declaradas impróprias no comportamento do mesmo para quaisquer que sejam o número de usuários ou plataformas onde o sistema será executado.
- **IV. Manutenibilidade:** O sistema deve ser capaz de ser submetido a reparos e modificações com baixa queda de performance e sem que haja o comprometimento das suas funcionalidades básicas.

4.4 Diagramas do Sistema

4.4.1 Atores do Sistema

O sistema por ser de caráter científico, não prescreve vários atores em seu uso, dito isto podemos destacar os seguintes:

- **Clientes do Sistema:** Utilizam o sistema e suas funcionalidades para atingir os objetivos ao qual este foi desenvolvido.
- **Engenheiros de Software:** Responsáveis pelos testes e implementações necessárias ao desenvolvimento e manutenibilidade do sistema.

4.4.2 Diagrama de Casos de Uso

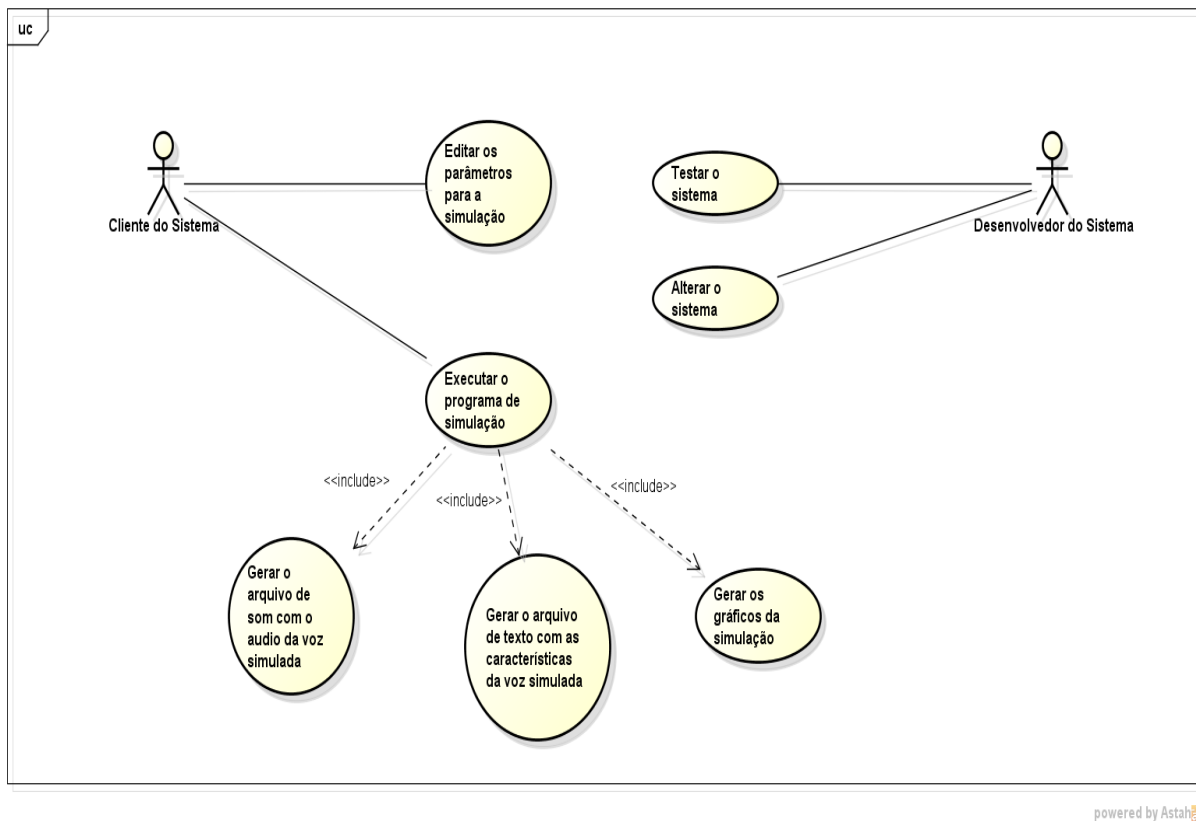


Figura 4.2: Diagrama de Casos de Uso do sistema.

A Figura 4.2 é um diagrama de casos de uso para o SynPath e descreve a interação dos atores supracitados com o sistema.

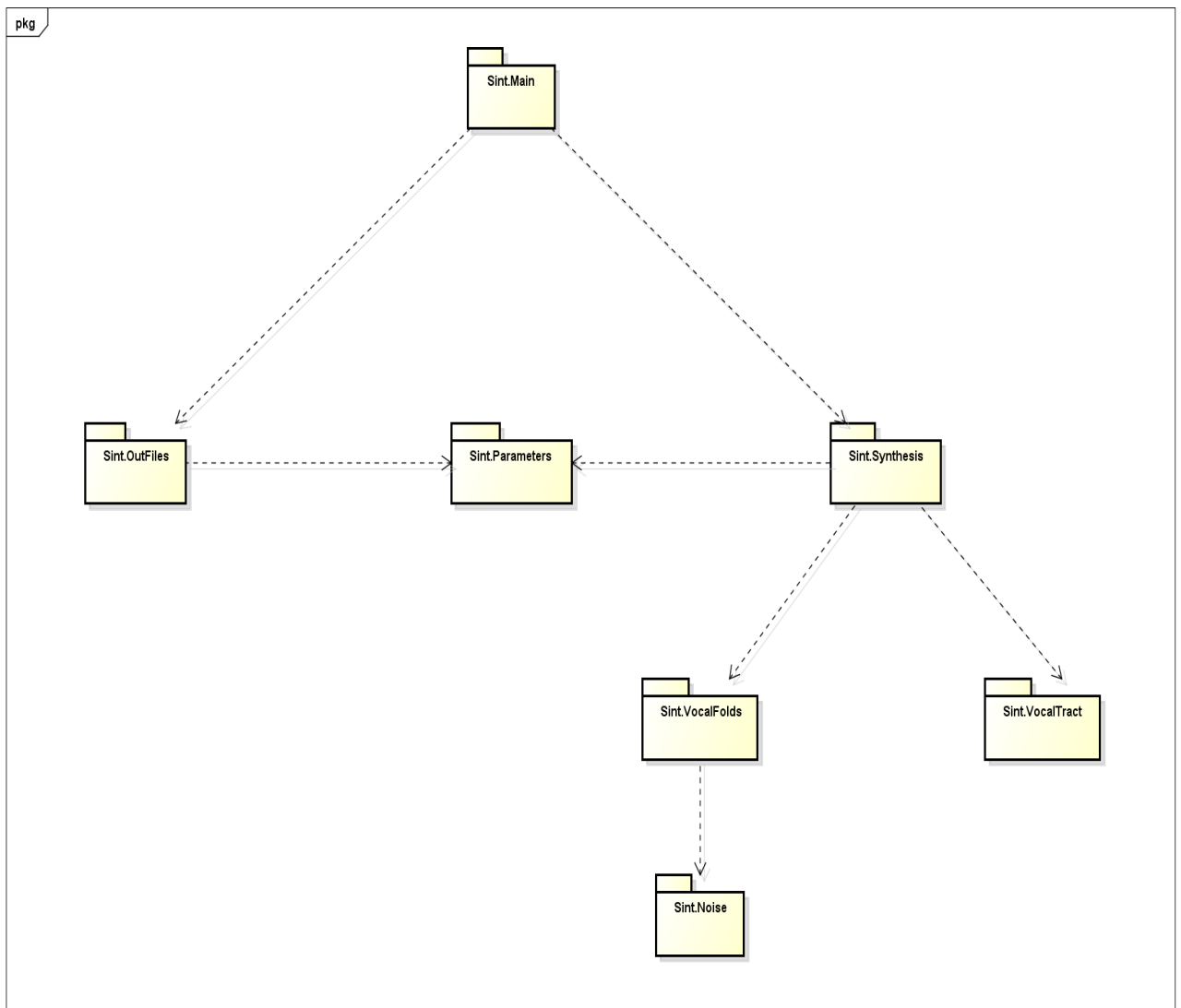
4.5 Arquitetura do Sistema

Nesta seção será descrita a arquitetura do software SynPath e seus componentes principais, sendo que a especificação será subdividida em duas arquiteturas:

- **Fase 1:** Arquitetura Conceitual
- **Fase 2:** Arquitetura de Implementação

4.5.1 Arquitetura Conceitual

Nesta subseção encontra-se a arquitetura do sistema exibida como decomposição do mesmo em subsistemas e nas interfaces que cada subsistema provê. Para a representação gráfica, iremos utilizar o diagrama UML de pacotes. Segue o diagrama:



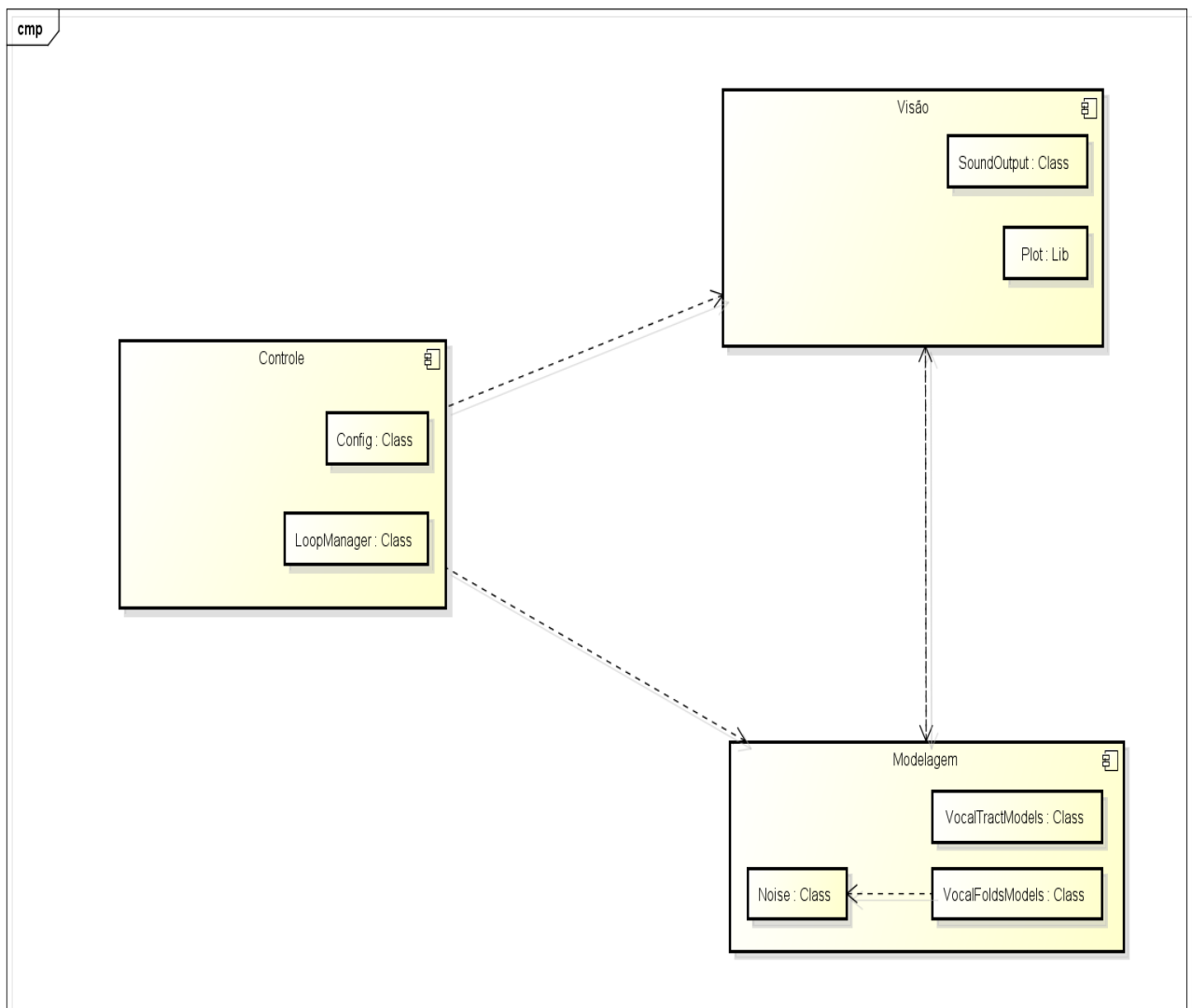
powered by Astah

Figura 4.3: Diagrama de Pacotes do sistema.

4.5.2 Arquitetura de Implementação

O SynPath segue a lógica do padrão MVC(Model-View-Controller), onde a fase de modelagem é literalmente a modelagem das pregas vocais, a fase de visão são os gráficos gerados e os arquivos de som e texto criados, e a fase do controle está sob as classes de parâmetros que serão responsáveis por implementar a lógica do funcionamento e da interação das partes do sistema.

A Figura 4.4 abaixo expõe o que foi citado anteriormente de forma a fundamentar o sistema no conceito:

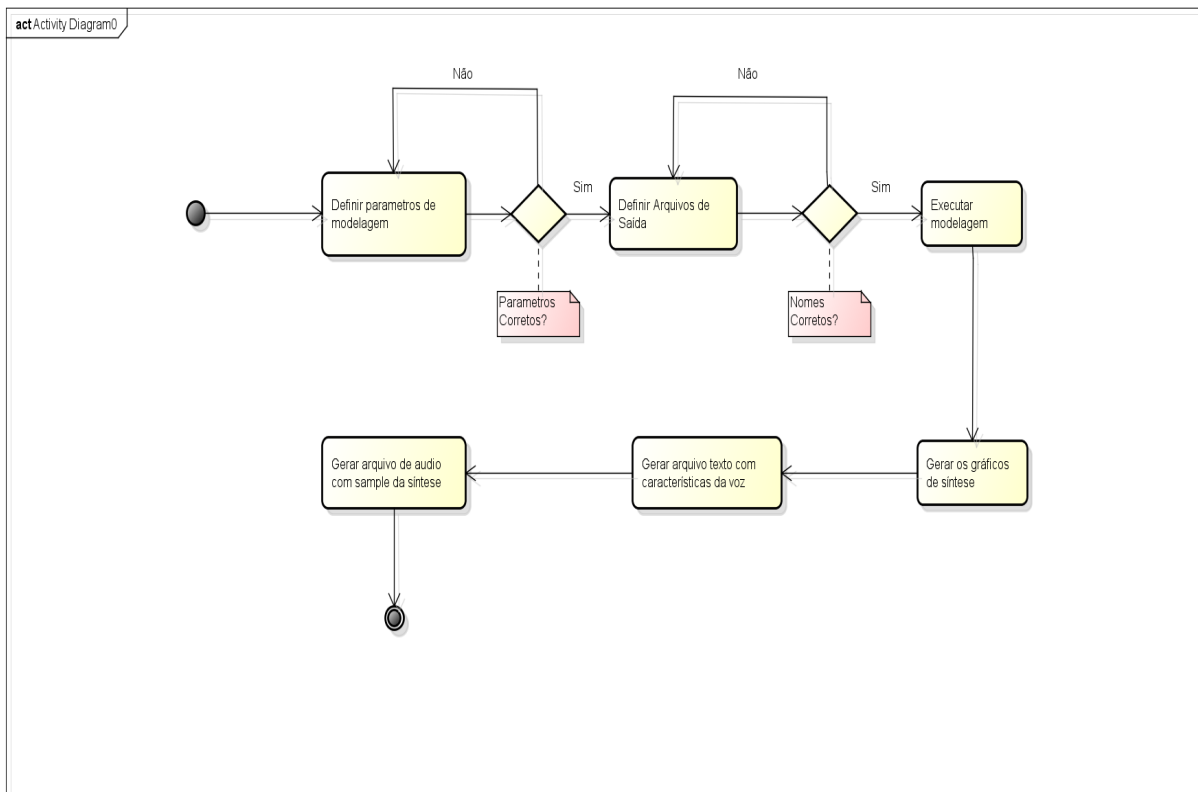


powered by Astah

Figura 4.4: Diagrama de Componentes do sistema.

4.5.3 Diagrama de Atividades

O diagrama a seguir é chamado de diagrama de atividades do sistema, este diagrama tem por intuito fornecer uma visão geral do funcionamento do sistema e da sequência de ações que ocorrem até que seja obtido o resultado esperado ao final de sua execução. A Figura 4.5 detalha este diagrama.



powered by Astah

Figura 4.5: Diagrama de Atividades do sistema.

Capítulo 5

Implementação da Fenda Glótica

Neste último capítulo descreveremos a implementação de uma fenda glótica patológica para o sintetizador SynPath. Será explicado a modelagem da mesma dentro dos moldes matemáticos utilizados para descrever o trato vocal, as pregas vocais e a glote. Em seguida será exposta a transposição do modelo matemático para o programa sintetizador e a partir daí serão colhidos os resultados da síntese e analisados.

5.1 Modelagem Matemática

Para se implementar uma fenda, é necessário que se implemente um mecanismo aditivo de ruído à síntese, sendo que este pode ser subdividido em dois termos, um ruído pulsátil referente às turbulências durante a produção da voz e um ruído de aspiração causado pelo 'vazamento' de ar do fluxo devido a um mal fechamento das cordas vocais durante a fonação em detrimento da fenda glótica.

O ruído pulsátil será analisado no software de duas maneiras, a primeira é como um fator linear proporcional ao fluxo glotal:

$$b\epsilon U_g \quad (5.1)$$

Onde b é o coeficiente de escala e ϵ é o fator de ruído computado na equação 4.11 e filtrado por um filtro passa-baixo de 1000 Hz de banda de frequência.

Em seguida, analisaremos o ruído como um fator proporcional ao quadrado do fluxo glotal a partir de um limiar pré-determinado, chamado de limiar de Reynolds [22]. Dessa maneira, calcular-se-á o número de Reynolds, para comparação com o limiar pré-estabelecido, pela seguinte equação:

$$Re = \frac{u_g \rho}{L \nu} \quad (5.2)$$

Onde u_g o valor do fluxo glotal, L é o comprimento da glote, ρ é a densidade do ar e ν é a viscosidade do ar. Utilizando a eq 5.2 podemos então calcular o ruído do fluxo da seguinte maneira:

$$ruído_p = \begin{cases} N_f(Re^2 - Re_c^2)(6,1 \times 10^{-5}) & , Re > Re_c \\ 0 & , Re \leq Re_c \end{cases} \quad (5.3)$$

Onde N_f é um ruído filtrado por um filtro passa-baixa de 300 Hz, Re é o número de Reynolds calculado na eq 5.2 e Re_c é o limiar de Reynolds estabelecido em 1200, de acordo com o trabalho de Story [22], o que significa que não serão gerados os ruídos pulsáteis abaixo deste limiar.

O ruído de aspiração por sua vez deve ser implementado diretamente na equação 4.6 do fluxo de ar na glote. Isto porque este ruído está diretamente ligado ao problema de vazamento ocorrido pela presença de uma fenda entre as pregas vocais. Dessa maneira, a equação do fluxo deve ser modificada para incorporar um termo que alterará o valor mínimo da área de contato entre as pregas vocais durante a fonação. Sendo assim, a nova equação de fluxo pode ser descrita da seguinte maneira:

$$U_g = \pm \left(\frac{ca_m + ca_f}{k_t} \right) \left\{ - \left(\frac{a_m + a_f}{A^*} \right) + \left[\left(\frac{a_m + a_f}{A^*} \right)^2 \pm \left(\frac{4k_t}{c^2\rho} \right) (P_s^+ - P_i^-) \right]^{1/2} \right\} \quad (5.4)$$

Onde o termo a_f é o termo limitante inferior da área de contato entre as pregas vocais.

5.2 Implementação

Como dito anteriormente, o SynPath foi implementado todo em linguagem Python. Como pode ser visto na Figura 4.3, o programa pode ser dividido em sete pacotes principais. Cada pacote contém as classes necessárias à implementação de uma determinada função dentro da síntese. A seguir serão descritas as classes principais do programa e suas funcionalidades, dando prioridade ao detalhamento de como foi implementada a simulação da fenda glótica.

5.2.1 Classe main.py

Responsável pela execução do programa. É ela quem controla o acoplamento das funcionalidades providas pelas outras classes de modo a resultar no produto final esperado, um arquivo texto com as características do som sintetizado, gráficos referentes ao som produzido e o arquivo contendo o som gerado. Esta classe é a responsável por inicializar a classe *config.py* que contém os parâmetros que serão utilizados no decorrer da execução pelas outras classes do programa e também por criar e definir o arquivo texto com as principais características do som sintetizado.

5.2.2 Classe config.py

Responsável por definir todas as constantes utilizadas no curso de execução do programa por basicamente todas as outras classes. Constantes como por exemplo a área dos cilindros que formam o trato vocal, a área da fenda entre as pregas vocais, velocidade do som, densidade do ar, o caminho onde serão salvos os arquivos finais e etc. Mais detalhadamente, para a implementação da fenda no programa, foram utilizadas constantes que definem a área da mesma, o tipo de ruído pulsátil (linear ou quadrático) e seu valor e o limiar de Reynolds, caso utilize-se a modelagem do ruído quadrático (Figura 5.1).

```

416
417 # ... Vocal Gap
418 ... VG_AREA = 0.1
419 ...
420 # ... Additive and pulsatile noise
421 ...
422 ... NOISE_DETERMINANT = "QUADRATIC" #QUADRATIC OR LINEAR
423 ...
424 ... NOISE_BAND = 1000.
425 ... ASPIRATION = 0.
426 ... PULSATILE = 100 #25. #.25. #1.#10. #100. #20.
427 ... REYNOLDS_LIMIT = 1200.

```

Figura 5.1: Parte da classe config.py

5.2.3 Classe soundOutput.py

Responsável pela criação do arquivo '.wav' contendo o som da síntese. Essa classe tem por função normalizar o som gerado na síntese para uma faixa preestabelecida de valores mínimo e máximo, adicionar pequenas quantidades de tremor ao som, a fim de garantir um aspecto mais realista ao mesmo, quantizar e codificar o som a fim de que este possa ser representado por uma quantidade finita de bits, ou seja, converter o sinal analógico para o seu representante digital e por fim criar e definir o arquivo sonoro final.

5.2.4 Classe vocalTractModels.py

Responsável pela modelagem dos diversos componentes físicos que compõem o trato vocal, assim como também a definição das características do trato vocal. De forma mais clara, essa classe contém diversas subclasses, chamadas também de *inner* classes, responsáveis por definir um componente do trato vocal ou as características do mesmo. Dentre essas podemos citar por exemplo, as subclasses *Lips* e *Glottis* responsáveis pela modelagem física dos lábios e da glote, respectivamente, e as subclasses *ViscousLosses* e *Propagation* responsáveis pela definição das características de perda por viscosidade e propagação do trato vocal.

5.2.5 Classe vocalFoldModels.py

Responsável pela modelagem computacional das pregas vocais. Assim como a classe *vocalTractModels*, esta classe possui diversas subclasses responsáveis por modelar de forma íntegra as pregas vocais de acordo com os modelos propostos nos capítulos anteriores. É nessa classe que se implementa a simulação da fenda glótica, utilizando-se as equações 5.4, 5.1 e 5.3. Essas equações foram implementadas dentro da função que define o fluxo, conforme mostrado na Figura 5.2.

Como pode ser visto, na linha 101, primeiramente adiciona-se a área da fenda à área glotal de modo que a fenda se torna o limiar inferior no fechamento das pregas vocais. Nas linhas 112 e 114 encontra-se a eq 5.4 do fluxo glotal, devidamente modificada pela fenda glótica. O 'if-else' implementado é utilizado para a bifurcação da equação 5.4 para os dois casos que podem ocorrer por conta do sinal \pm .

```

95     def flow(self, ag, w, x0, Ai, ps_in, pi_in):
96         .....
97         c = self.c
98         .....
99         Ai = c.CORRCOEF1*Ai
100        A = c.A_TR*Ai/(c.A_TR+Ai)
101        ag += c.VG_AREA
102        .....
103        if ag > 0. :
104            .....
105            # re = (c.A_TR-ag)/(c.A_TR+ag)
106            # rs = (Ai-ag)/(Ai+ag)
107            .....
108            re=1.
109            rs=1.
110            .....
111            if ps_in >= pi_in:
112                ug_clean = ag*c.SOUND_SPEED/c.VF_KT*(-ag/A+np.sqrt((ag/A)**2.+2.*c.VF_KT/c.SOUND_SPEED**2./c.RHO*((1.+rs)*ps_in-(1.+re)*pi_in)))
113            else:
114                ug_clean = ag*c.SOUND_SPEED/c.VF_KT*(ag/A-np.sqrt((ag/A)**2.-2.*c.VF_KT/c.SOUND_SPEED**2./c.RHO*((1.+rs)*ps_in-(1.+re)*pi_in)))
115            .....
116            #####
117            .....
118            if c.NOISE_DETERMINANT == "QUADRATIC":
119                .....
120                reynolds = (ug_clean * c.RHO)/(c.GLOTTAL_LENGTH*c.NU)
121                .....
122                if reynolds > c.REYNOLDS_LIMIT:
123                    noise_buffer = rd.normalvariate(0., c.SQRT_DELTA_T)
124                    self.filtered_noise = self.lp_filter_pulsatile.get_otp(noise_buffer,self.a3,self.b3,self.c3)
125                    pulsatile_noise = self.filtered_noise * c.PULSATILE * (reynolds*reynolds -
126                                                                c.REYNOLDS_LIMIT*c.REYNOLDS_LIMIT) * (0.000061)
127                    aspiration_noise = self.filtered_noise * c.ASPIRATION
128                else:
129                    pulsatile_noise = 0.
130                    aspiration_noise = 0.
131                .....
132            elif c.NOISE_DETERMINANT == "LINEAR":
133                noise_buffer = rd.normalvariate(0.,c.SQRT_DELTA_T)
134                self.filtered_noise = self.lp_filter_pulsatile.get_otp(noise_buffer,self.a3,self.b3,self.c3)
135                pulsatile_noise = self.filtered_noise * (c.PULSATILE * ug_clean)
136                aspiration_noise = self.filtered_noise * c.ASPIRATION
137            .....
138            #####
139            .....
140            noise = pulsatile_noise + aspiration_noise
141            ug = ug_clean + noise

```

Figura 5.2: Parte da classe vocalFoldsModel.py

A partir daí, a começar na linha 116 até a linha 138, são definidos os dois modos de ruído pulsátil analisados no presente trabalho pelas equações 5.1 e 5.3. A partir da escolha do usuário, através da constante *NOISE_DETERMINANT*, o programa simulará a fenda com base em um desses dois ruídos, seguindo o modelo matemático implementado nas linhas subsequentes.

Por fim, nas linhas 140 e 141, o ruído é somado ao fluxo definido anteriormente para ser então retornado pela função ao programa para utilização na síntese.

5.2.6 Classe *synthesis.py*

Aqui ocorre a síntese propriamente dita, ou seja, aqui são computados a excitação glotal, os coeficientes de *jitter* e abertura glotal, o ruído, a excitação do trato vocal e a propagação no mesmo. Em seguida, essa classe pega o resultado de cada computação feita por intermédio das outras classes que compõem o aparelho fonador simulado, trato vocal e pregas vocais, e retorna ao programa para serem plotados os gráficos e retiradas as informações necessárias.

5.3 Resultados e Analise

Nesta seção serão expostos os resultados obtidos pela simulação de uma fenda glótica no software SynPath. Em seguida, após exibidos os detalhes acerca dos resultados obtidos, será feita uma análise partindo de duas perspectivas específicas, a primeira, teórica, ou seja, analisar as implicações teóricas que levaram aos diferentes resultados e a segunda, prática, analisar se o resultado obtido se aproxima do esperado em um caso clínico real de fenda patológica nas pregas vocais.

Os resultados serão analisados de três maneiras a saber:

A primeira maneira, a mais simples e prática, de analisarmos a existência de uma fenda glótica e suas influências na voz é ouvir se os sons gerados, pelo SynPath, característicos de uma voz com essa patologia, apresentam os sintomas característicos, que são: quanto maior a fenda maior a incidência de sopro e rugosidade exacerbada na voz se comparada à uma voz natural sem a existência de fenda glótica. Entretanto para que essa análise seja feita de maneira a prover resultados confiáveis, é necessário que profissionais que trabalhem na área de fonoaudiologia e que possuem o conhecimento certo sejam as pessoas a analisar o produto final do software.

A segunda maneira, teórica, é analisarmos o comportamento das vozes geradas com relação à razão sinal/ruído. A razão sinal/ruído compara o nível do sinal desejado (nesse caso a voz) com o nível de ruído existente no canal (gerado pela fenda). Caso haja uma fenda, conforme explicado anteriormente, há um vazamento no fluxo de ar durante a produção da voz, o que gera um ruído no canal. Quanto maior for o resultado dessa razão, menor é a influência do ruído na voz, o que significa que para vozes não afetadas pela fenda, a razão sinal/ruído deve ser necessariamente maior do que para aquelas afetadas pela fenda. É possível por essa razão também, analisar a influência do tamanho da fenda na voz, pois quanto maior a fenda, maior o ruído e conseqüentemente, menor o resultado da razão S/N (sinal/ruído).

Uma terceira maneira de analisarmos, também de forma teórica, é verificarmos se os gráficos acerca das características da produção e síntese da voz gerados pelo software SynPath diferem, entre uma voz considerada livre de patologias e uma voz afetada pela fenda glótica, conforme o esperado.

A seguir serão expostas as tabelas contendo os diferentes valores obtidos para a razão sinal/ruído de cada voz patológica sintetizada. Para se obter esses resultados, foram feitas precisamente 10 simulações para cada tipo de implementação de ruído (linear e quadrático) e para cada tipo de voz (masculina e feminina), totalizando 40 simulações. Dentro dessa gama de simulações, os parâmetros relativos ao tamanho da fenda e à quantidade de ruído pulsátil foram alterados para se mensurar o comportamento do programa e o produto final da síntese. Para cada mudança de parâmetro, a síntese foi executada três vezes, sendo que os valores obtidos para a razão S/N foram extraídos dessas execuções por meio de média aritmética.

Após as tabelas serão apresentados os gráficos comparativos para ambos os tipos de voz e para ambas implementações de ruído (linear e quadrático) comparando a voz sintetizada não afetada pela fenda com a voz sintetizada afetada. Mais especificamente, esses gráficos farão a comparação do espectro das frequências de resposta, do fluxo glotal e do espectro do som final.

5.3.1 Resultados S/N - Voz Masculina

Ruído Pulsátil Linear

	Tamanho da Fenda(em cm ²)	Coef. <i>b</i> na Eq. 5.1	Razão S/N(em dB)
Som Normal	0	0	224,549
Fenda 1	0,07	25	41,954
Fenda 2	0,07	50	35,764
Fenda 3	0,07	100	30,174
Fenda 4	0,1	25	41,802
Fenda 5	0,1	50	36,161
Fenda 6	0,1	100	30,203
Fenda 7	0,14	25	41,939
Fenda 8	0,14	50	35,900
Fenda 9	0,14	100	29,808

Tabela 5.1: Resultados da síntese linear

Ruído Pulsátil Quadrado

	Tamanho da Fenda(em cm ²)	Coef. <i>b</i> na Eq. 5.1	Razão S/N(em dB)
Som Normal	0	0	224,549
Fenda 1	0,07	25	43,058
Fenda 2	0,07	50	36,775
Fenda 3	0,07	100	30,290
Fenda 4	0,1	25	42,810
Fenda 5	0,1	50	36,782
Fenda 6	0,1	100	29,967
Fenda 7	0,14	25	42,640
Fenda 8	0,14	50	36,587
Fenda 9	0,14	100	29,472

Tabela 5.2: Resultados da síntese quadrática

5.3.2 Resultados S/N - Voz Feminina

Ruido Linear

	Tamanho da Fenda(em cm ²)	Coef. b na Eq. 5.1	Razão S/N(em dB)
Som Normal	0	0	223,990
Fenda 1	0,07	25	41,858
Fenda 2	0,07	50	35,791
Fenda 3	0,07	100	29,606
Fenda 4	0,1	25	42,303
Fenda 5	0,1	50	35,850
Fenda 6	0,1	100	29,739
Fenda 7	0,14	25	42,244
Fenda 8	0,14	50	35,938
Fenda 9	0,14	100	29,831

Tabela 5.3: Resultados da síntese linear

Ruido Quadrado

	Tamanho da Fenda(em cm ²)	Coef. b na Eq. 5.1	Razão S/N(em dB)
Som Normal	0	0	223,990
Fenda 1	0,07	25	43,453
Fenda 2	0,07	50	37,354
Fenda 3	0,07	100	31,174
Fenda 4	0,1	25	43,251
Fenda 5	0,1	50	37,043
Fenda 6	0,1	100	30,753
Fenda 7	0,14	25	42,767
Fenda 8	0,14	50	36,700
Fenda 9	0,14	100	29,831

Tabela 5.4: Resultados da síntese quadrática

5.3.3 Gráficos - Voz Masculina

Voz Normal

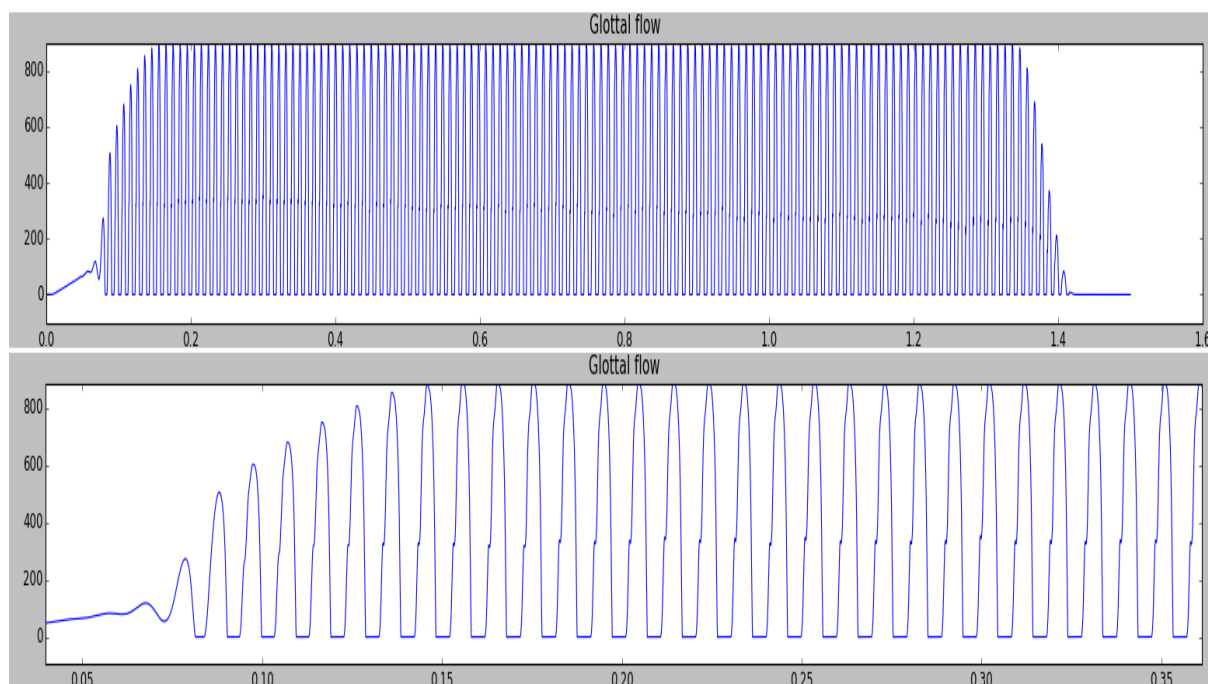


Figura 5.3: Fluxo Glotal

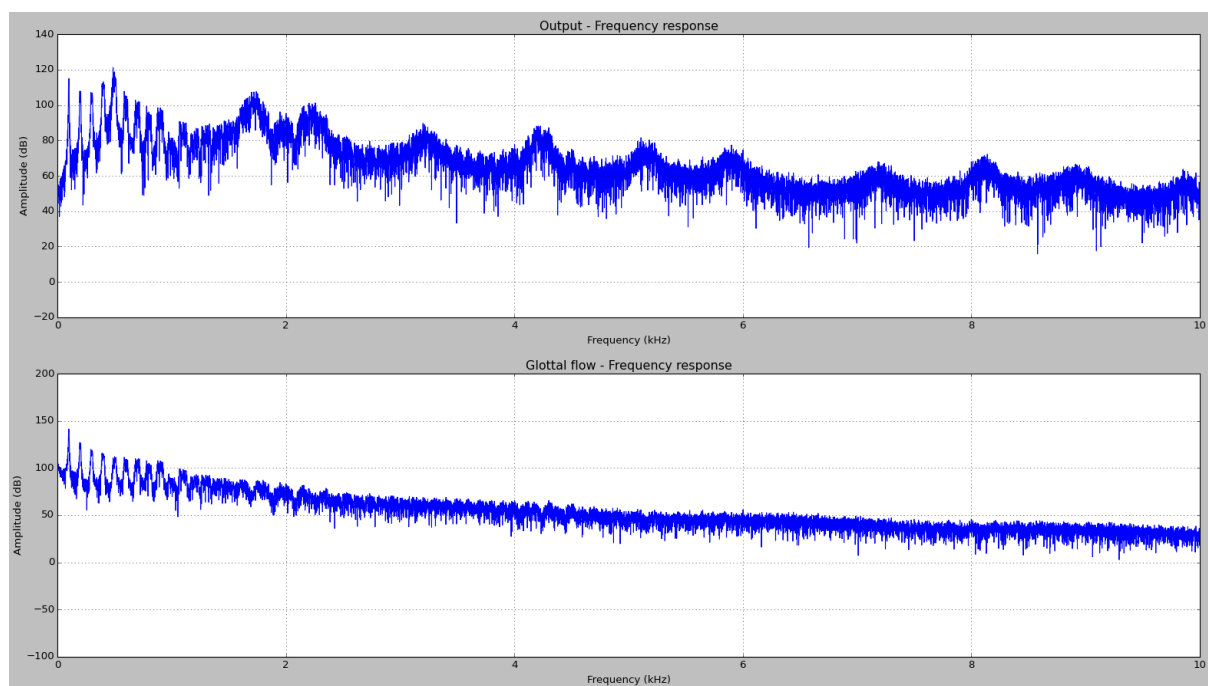


Figura 5.4: Espectro das Frequências de Resposta

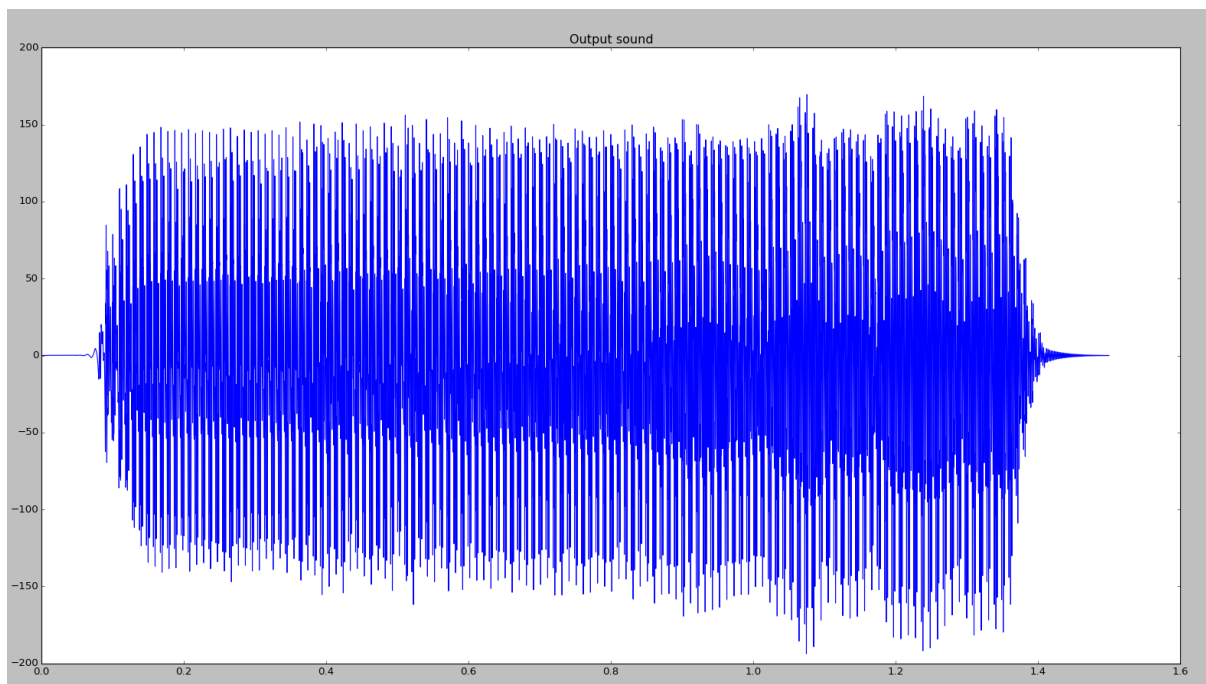


Figura 5.5: Onda Sonora

Voz com Fenda - Ruido Linear

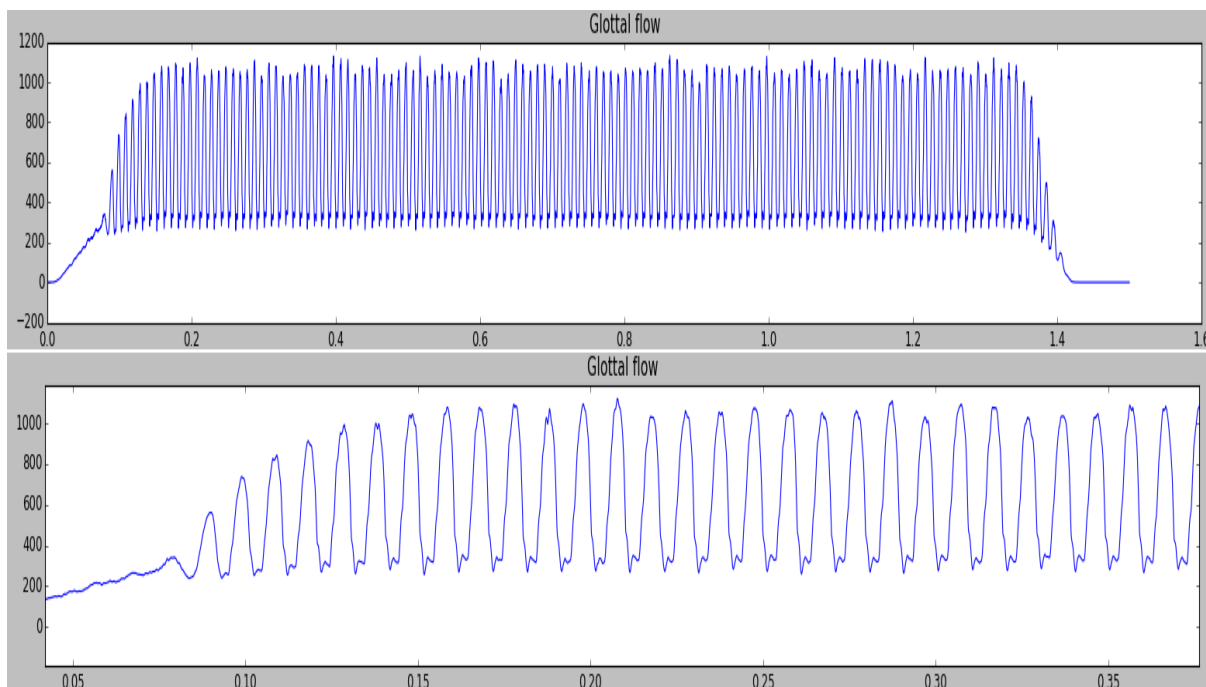


Figura 5.6: Fluxo Glotal

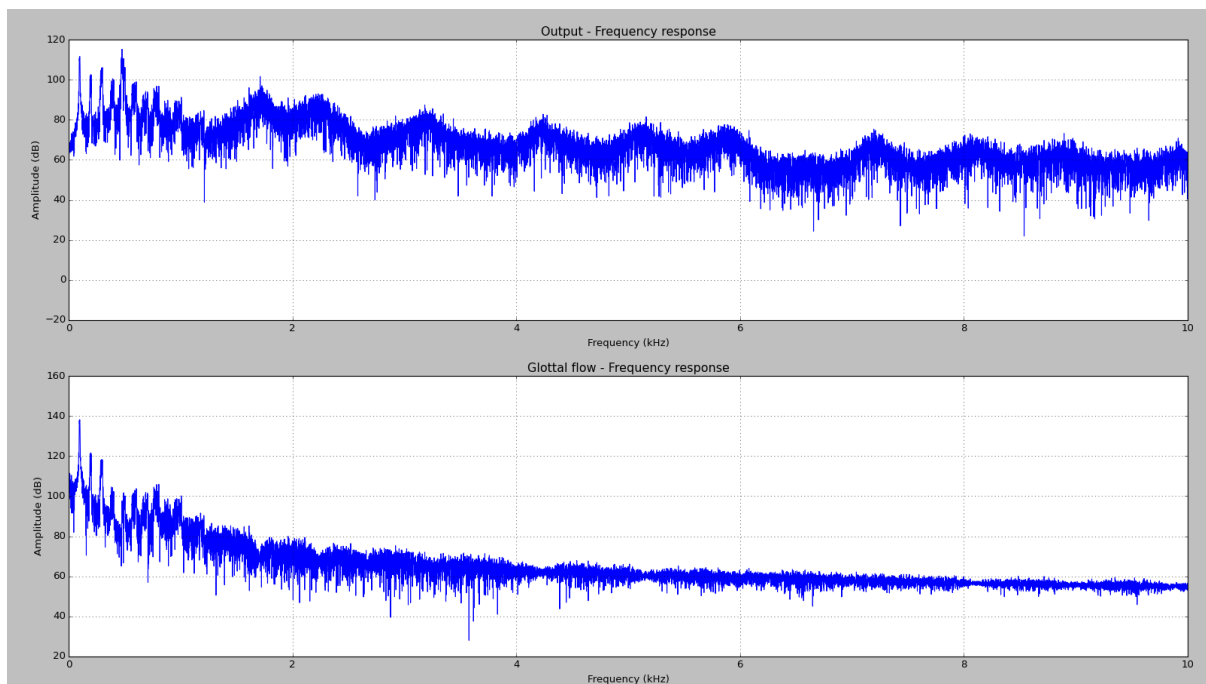


Figura 5.7: Espectro das Frequências de Resposta

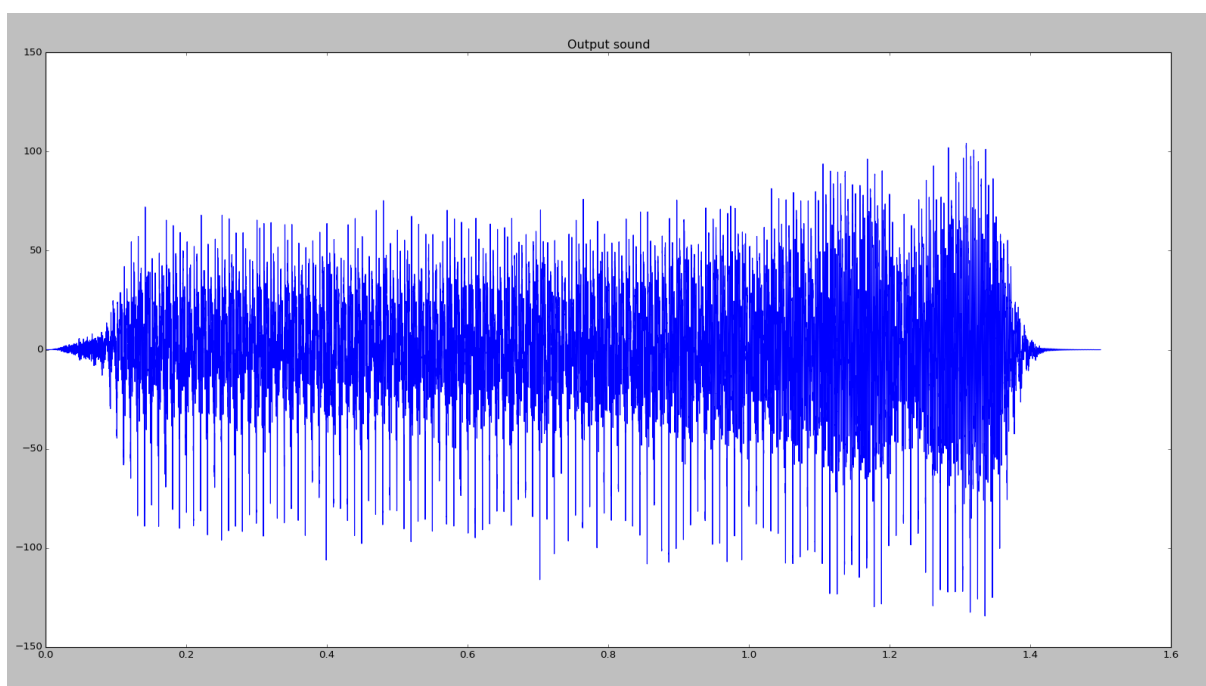


Figura 5.8: Onda Sonora

Voz com Fenda - Ruido Quadrado

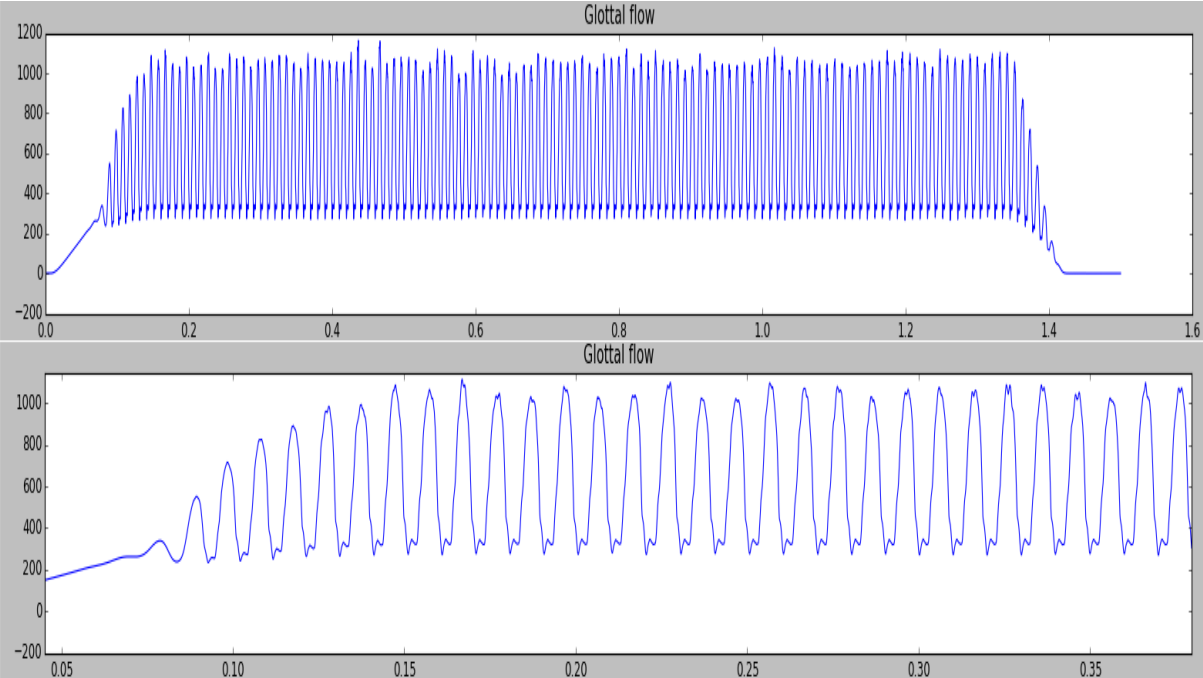


Figura 5.9: Fluxo Glotal

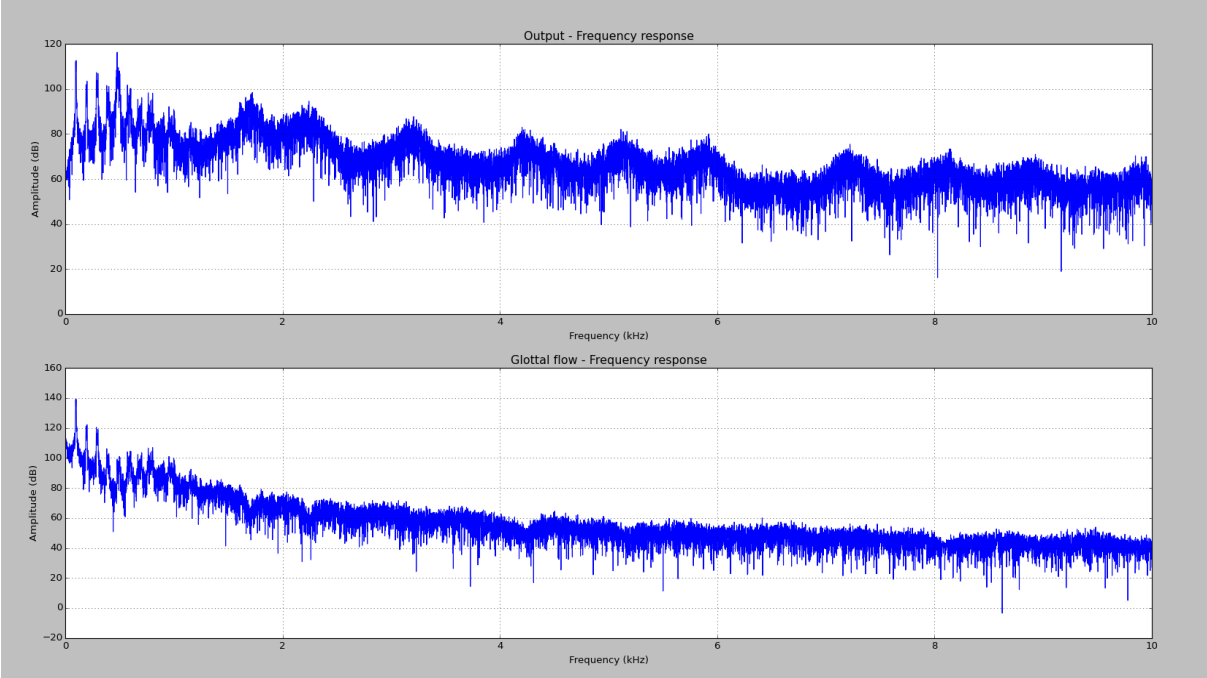


Figura 5.10: Espectro das Frequências de Resposta

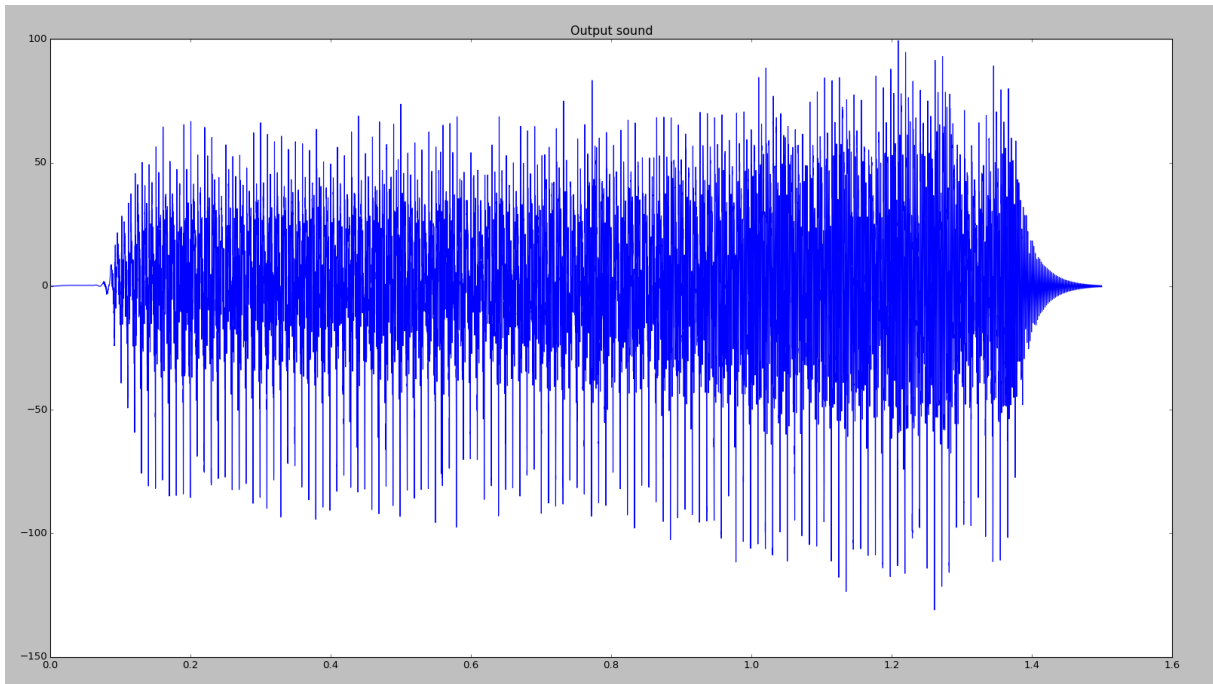


Figura 5.11: Onda Sonora

5.3.4 Gráficos - Voz Feminina

Voz Normal

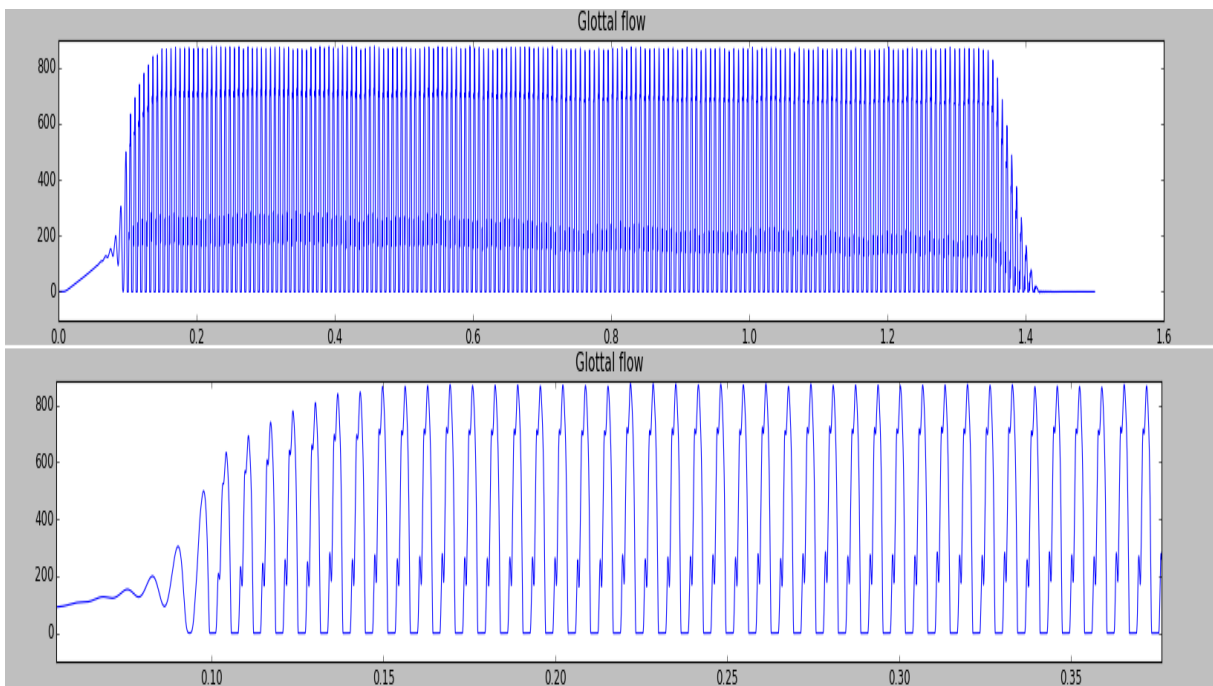


Figura 5.12: Fluxo Glotal

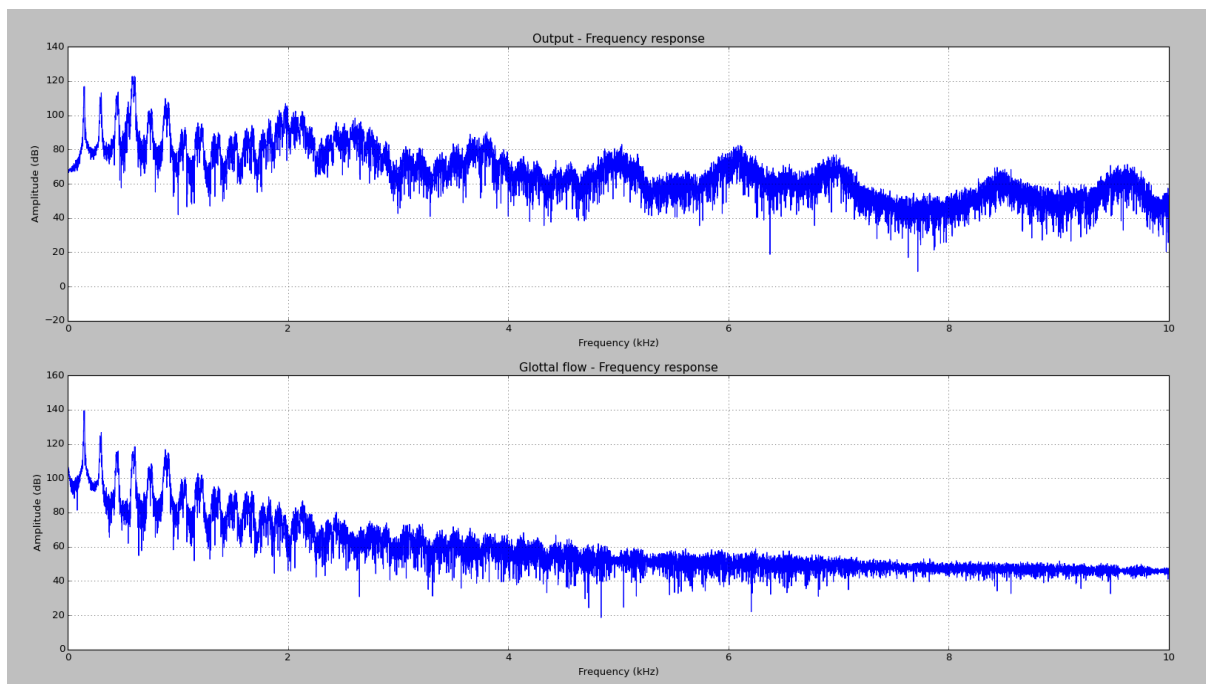


Figura 5.13: Espectro das Frequências de Resposta

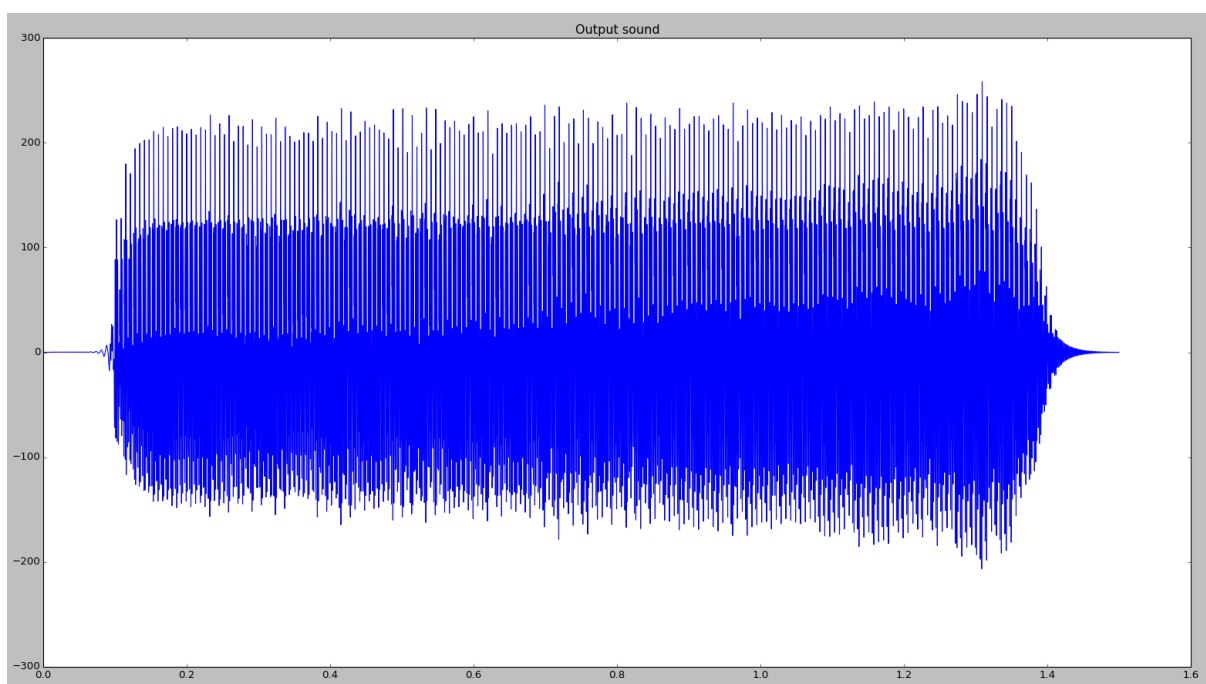


Figura 5.14: Onda Sonora

Voz com Fenda - Ruido Linear

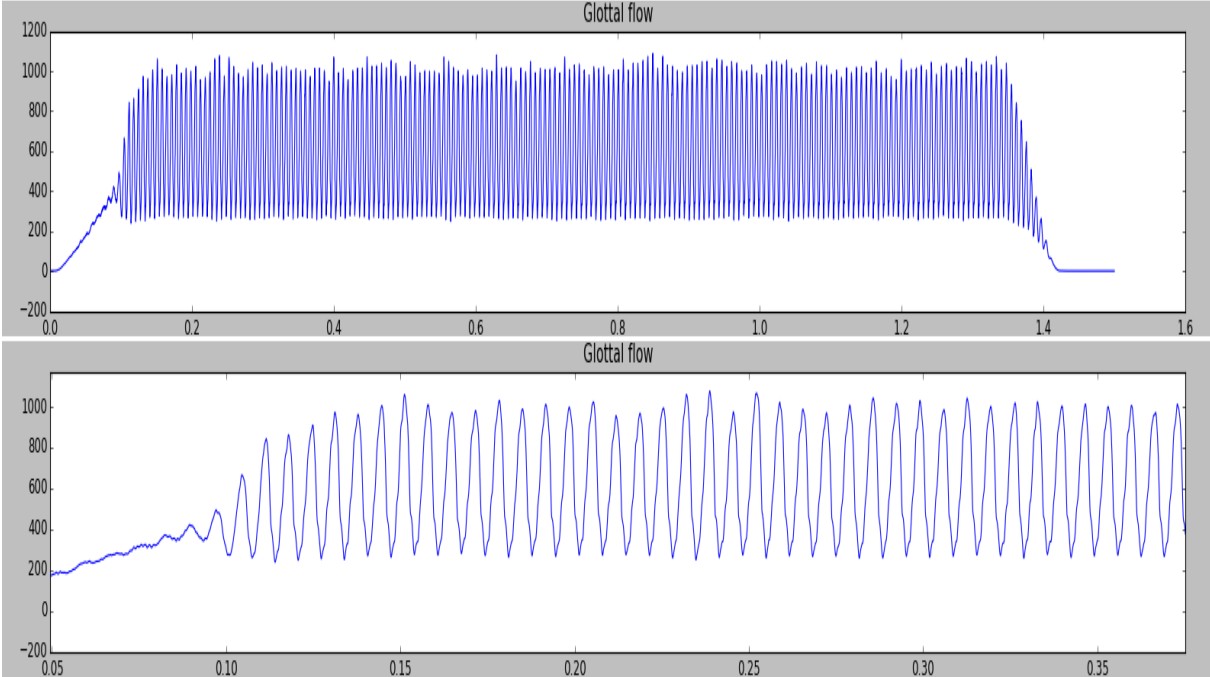


Figura 5.15: Fluxo Glotal

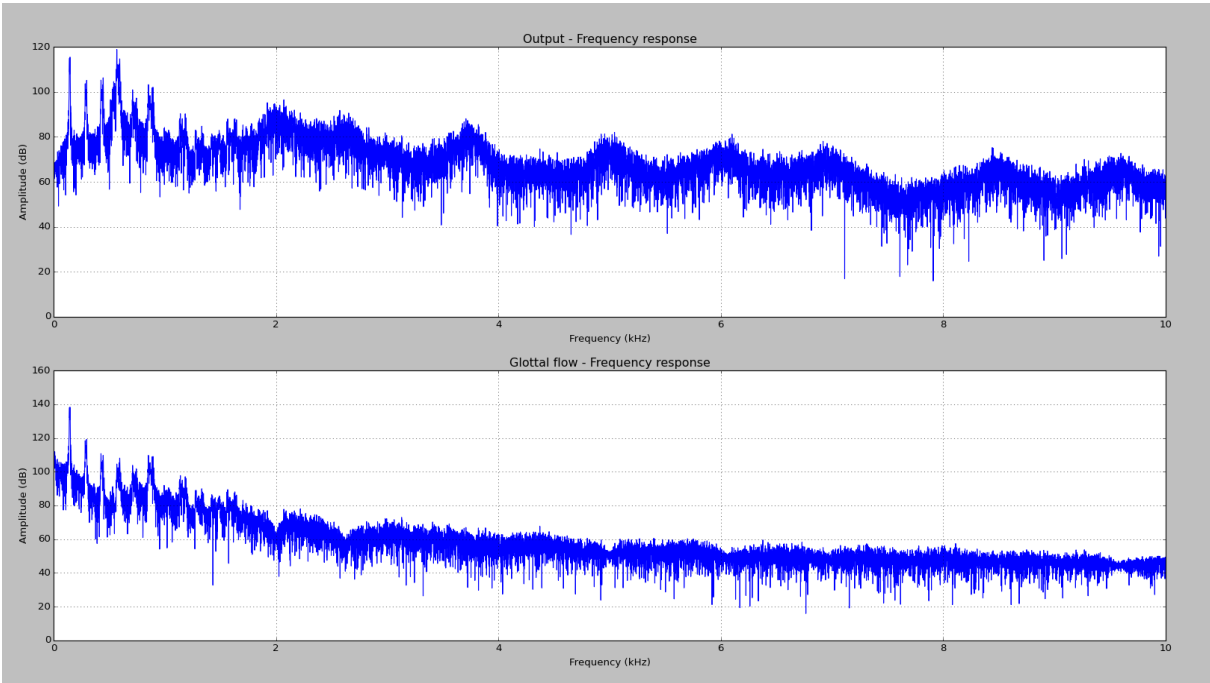


Figura 5.16: Espectro das Frequências de Resposta

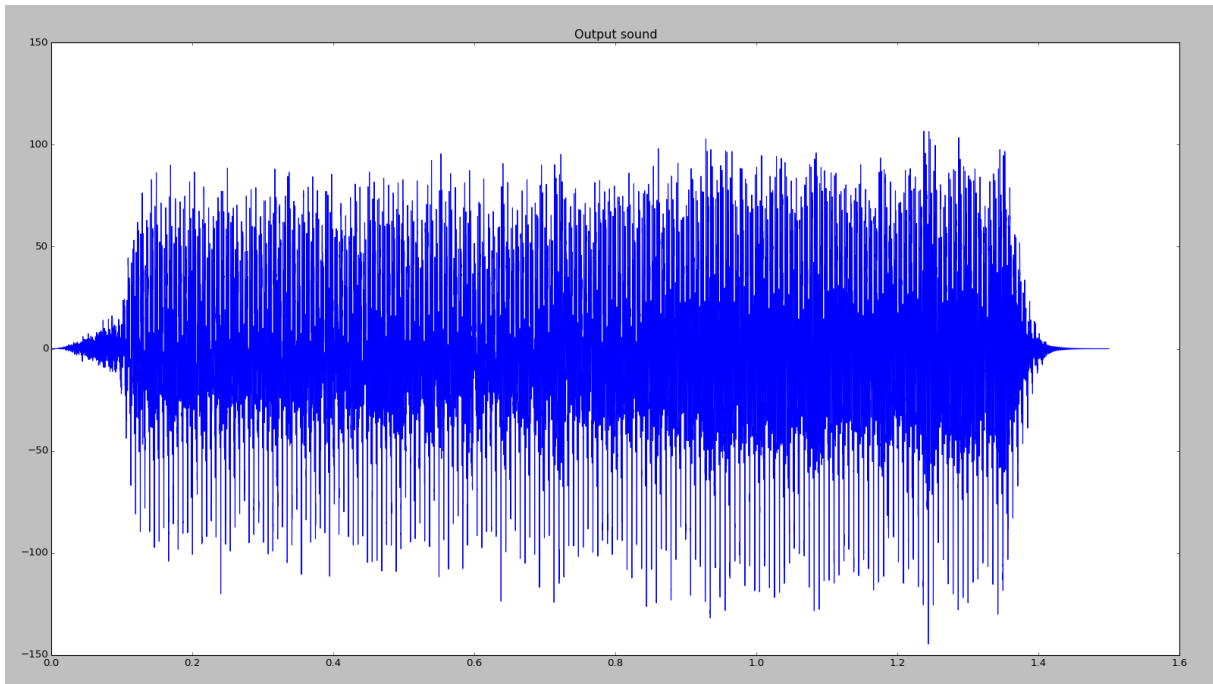


Figura 5.17: Onda Sonora

Voz com Fenda - Ruido Quadrado

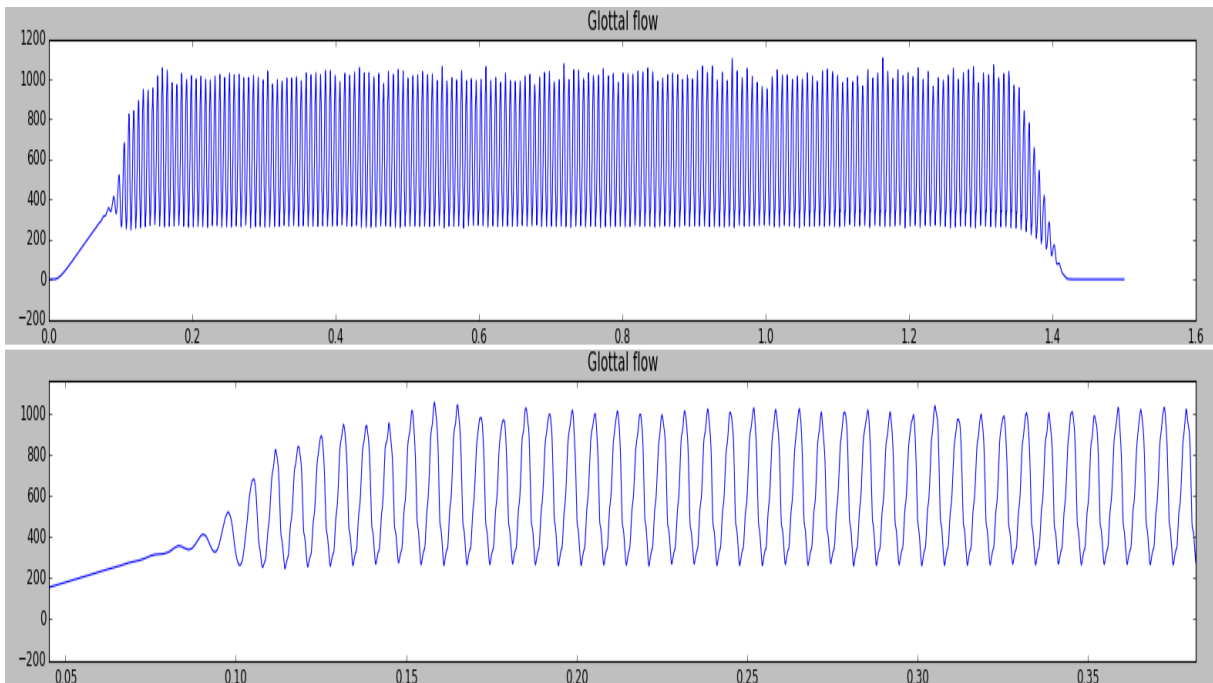


Figura 5.18: Fluxo Glotal

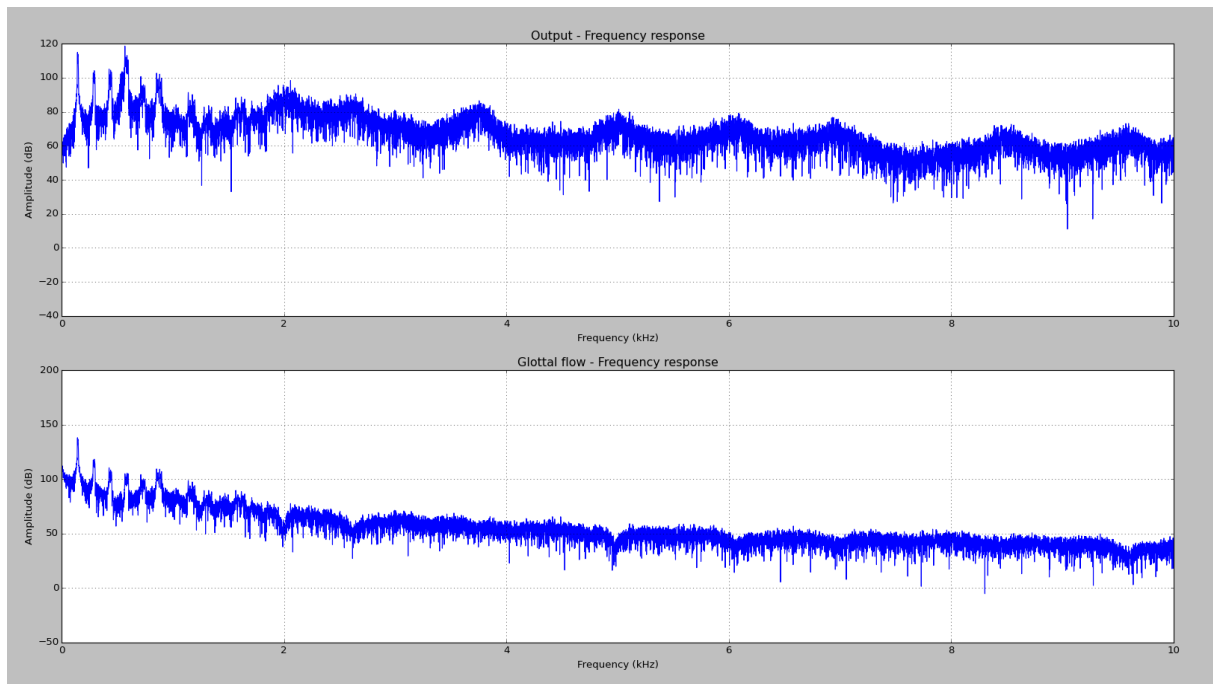


Figura 5.19: Espectro das Frequências de Resposta

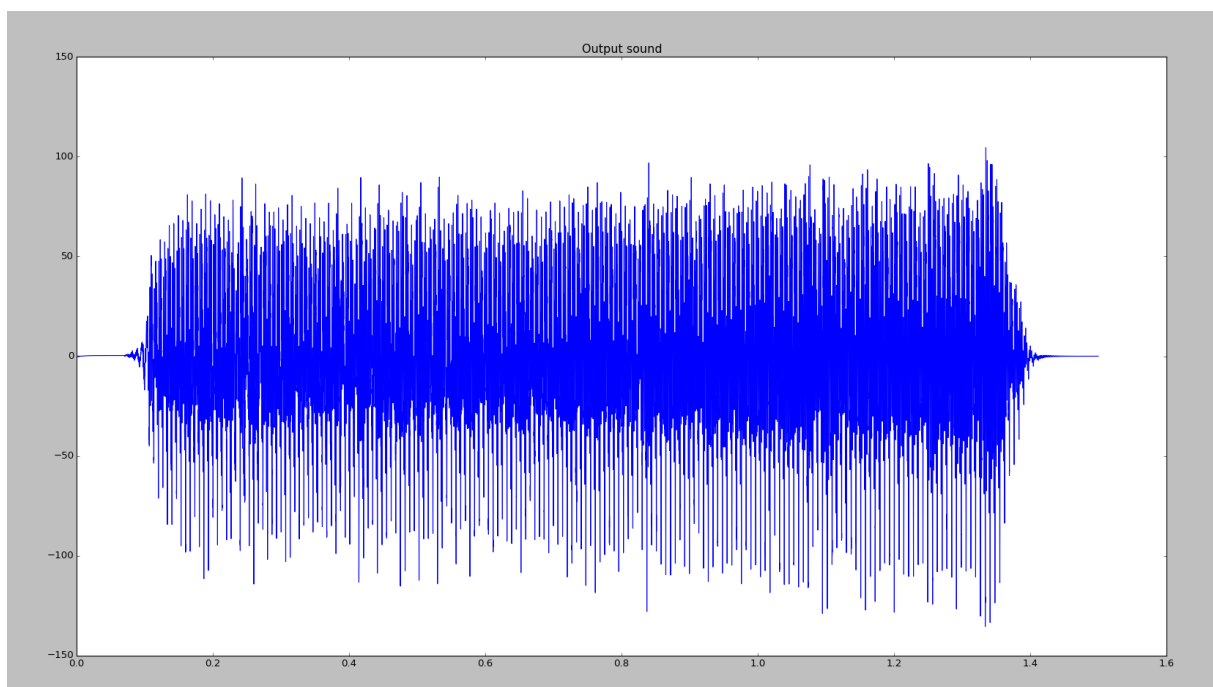


Figura 5.20: Onda Sonora

5.3.5 Análise dos Resultados

Primeira Análise - Avaliação dos Sons Sintetizados

Em parceria com o Centro de Estudos da Voz em São Paulo, foram enviados para análise os arquivos sonoros contendo a simulação das vozes afetadas por fenda glótica. As fonoaudiólogas Glaucya Madazio, Mara Behlau e Marina Englert analisaram do ponto de vista clínico os sons sintetizados e constataram que ambos os modelos (com ruído linear e quadrático) parecem naturais e que é perceptível a variação da intensidade da soproidade em ambos. A diferença constatada entre o modelo de ruído linear e o modelo de ruído quadrático é que no modelo de ruído linear a soproidade é mais audível e no modelo de ruído quadrático a rugosidade é mais intensa, apesar de em ambos estarem presentes as características citadas. Em relação aos casos clínicos reais, as fonoaudiólogas afirmaram que os arquivos sonoros podem servir como base para o aprendizado e estudo das características das vozes afetadas por fenda glótica pois as simulações oferecem uma qualidade sonora suficiente à análise e ao estudo da patologia.

Segunda Análise - Avaliação das Tabelas S/N

As tabelas sinal/ruído nos fornecem uma visão interessante a cerca do efeito da fenda glótica na qualidade do som da voz. Quando não há fenda, a razão sinal ruído atinge o valor em torno de 220 dB ao passo que quando implementada a menor fenda, o valor cai em torno de 40 dB para ambos os tipos de ruído. O que isso significa? Significa que a qualidade do sinal cai em torno de 78%. Traduzindo isso para a síntese de voz, significa que o produto final, a voz simulada, será afetado de forma que a qualidade do timbre e do brilho da voz serão alterados pelo ruído. Nesse caso, o ruído é a passagem do excesso de ar pela fenda, o que não permite o total fechamento das pregas vocais para executar a transformação da energia aerodinâmica em energia acústica. Consequentemente, o portador dessa patologia possui uma maior dificuldade em controlar a intensidade da voz, o que faz com que haja uma necessidade de esforço maior para falar e manter a fonação dentro de condições que outrora não seria necessário esforço algum.

Logo, a razão SNR nos auxilia na quantificação da qualidade da voz em detrimento da existência de uma fenda glótica e, nos permite avaliar, mais precisamente, os efeitos que essa patologia tem sobre a voz no que concerne a sua eficácia em sua utilização.

Terceira Análise - Comparação dos Gráficos

Essa última avaliação nos permitirá analisar as características das vozes simuladas com base no que for apresentado pelos gráficos.

- **Gráfico do Fluxo Glotal:** Comparando os gráficos das vozes afetadas pela fenda glótica com os gráficos da voz não afetada, é notável que o fluxo glotal da voz não afetada por fenda possui para cada pulso, a partir do pico máximo, a mesma amplitude máxima, mostrando uma regularidade amplitudinal entre pulsos. Além disso, os pulsos todos são iniciados no ponto 0 no eixo Y, o que caracteriza o fechamento total das pregas vocais durante a fonação. Entretanto, para as vozes afetadas por fenda, não só os pulsos possuem amplitudes irregulares, como também se iniciam acima do ponto 0 no eixo Y. Isso caracteriza o não fechamento total das pregas

vocais, constatando a existência da fenda glótica, e também nos mostra através da irregularidade das amplitudes, a falta de controle do timbre da voz, caracterizando a soproiedade e rugosidade presentes em vozes afetadas por essa patologia.

- **Gráfico do Espectro de Frequências de Resposta:** Se comparar os gráficos das vozes com fenda com o gráfico da voz sem a fenda, somente é notável que nos que possuem a fenda existe mais ruído nas regiões de frequência, caracterizado pelo aspecto mais grosso do espectro de frequências. Entretanto, se compararmos o gráfico de vozes com fenda com implementação do ruído linear com o gráfico de vozes com fenda com implementação do ruído quadrático nota-se uma diferença entre eles, o gráfico do ruído linear apresentam menor densidade de ruído se comparado ao gráfico do ruído quadrático. Isso caracteriza o que foi notado pelos fonoaudiólogos na primeira análise, uma maior soproiedade nas vozes sintetizadas com ruído linear e uma maior rugosidade nas vozes sintetizadas com o ruído quadrático, apesar de ambas possuíres ambas as características.
- **Gráfico do Espectro do Som:** Ao analisarmos esse gráfico é notável a diferença entre a riqueza harmônica das vozes não afetadas por fenda glótica se comparadas às afetadas. Esse gráfico ilustra a perda de brilho na voz ocasionada pela patologia em questão. A perda de qualidade notada pela segunda análise é refletida diretamente aqui. A soproiedade e rugosidade características dessa patologia vocal fazem com que o portador da patologia tenha uma esforço maior durante a fonação em detrimento da perda de intensidade e conseqüentemente o brilho harmônico da voz também é afetado.

5.4 Conclusão

Conclui-se a partir desse trabalho a possibilidade de se implementar uma fenda glótica computacionalmente. Sendo possível atingir um nível de qualidade de síntese computacional suficiente para que as simulações possam ser utilizadas para aprendizado e auxílio no diagnóstico da patologia aqui estudada.

5.5 Trabalhos Futuros

Apesar do módulo implementado aqui para o software SynPath, ainda existem coisas a se desenvolver e melhorar. O atual software ainda não possui uma interface gráfica de controle para usuários, não implementa um controle dinâmico do trato vocal, não implementa consoantes, não possui modelagem das falsas pregas vocais, também chamadas de pregas ventriculares, dentre outros. Além do não implementado, o software ainda necessita de melhorias em algumas áreas de performance. Entretanto, para o presente trabalho, o software mostrou-se suficientemente capaz de executar tudo o que foi proposto.

Referências

- [1] A. Lagier A. Giovanni and N. Henrich. Fisiología de la fonación. *EMC*, 43(3):1–15, 2014. 3
- [2] J. W. Van Der Berg. Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research*, 1:227–244, 1958. 16
- [3] Ingo R. Titze Brad H. Story and Darrell Wong. A simplified model for simulation and transformation of speech. *IEEE*, pages 320–327, 1996. 23
- [4] Kelley C. Stewart Michael W. Plesniak David E. Sommer Sean D. Peterson Byron D. Erath, Matias Zanartu. A review of lumped-element models of voiced speech. *Speech Communication*, 55:667–690, 2013. 21
- [5] J. L. Flanagan and L. L. Landgraf. Self-oscillating source for vocal-tract synthesis. *IEEE Tran. Audio Eletroacoust.*, AU-16(1):57–64, 1968. 1, 16, 18, 21
- [6] Henry Gray. *Anatomy of the Human Body*, <http://www.bartleby.com/107/236.html>, 1918. vii, 4, 5
- [7] Google Images. *Domínio Público*, 2015. vii, 3, 14, 15, 18, 25
- [8] K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal folds. *Bell Syst. Tech. J.*, 51(6):1233–1268, 1972. 1, 18, 21
- [9] K. Ishizaka and N. Isshiki. Computer simulation of pathological vocal-cord vibration. *Journal of Acoustical Society of America*, 60(5):1193–1198, 1976. 21
- [10] J. Schoentgen Jorge C. Lucero and M. Behlau. Physics-based synthesis of disordered voices. *Interspeech*, 105(1):1–5, 2013. 26, 27, 28
- [11] Kelem G. Lourenço Nicolas Ruty Jorge C. Lucero, Laura L. Koenig and Xavier Pelorson. A lumped mucosal wave model of the vocal folds revisited: Recent extensions and oscillation hysteresis. *Journal of Acoustical Society of America*, 129(3):1568–1579, 2010. vii, 26, 27
- [12] Jody Kreiman and Diana Sidtis. *Foundations of Voice Studies*. Wiley-Blackwell, John Wiley and Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom, 2011. vii, 1, 2, 6, 7
- [13] Hans Petter Langtangen. *A Primer on Scientific Programming with Python*. Springer, Springer Heidelberg New York Dordrecht London, 2009. 26

- [14] Jorge C. Lucero. A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset–offset. *Acoustical Society of America*, pages 423–431, 1999. [26](#)
- [15] Jorge C. Lucero and Laura L. Koenig. On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds. *Journal of Acoustical Society of America*, 121(6):3280–3283, 2007. [26](#)
- [16] Jorge C. Lucero and J. Schoentgen. Modeling vocal fold asymmetries with coupled van der pol oscillators. *Journal of Acoustical Society of America*, 19(060165):1–8, 2013. [1](#)
- [17] George R Wodicka Matias Zañartu, Luc Mongeau. Influence of acoustic loading on an effective single mass model of the vocal folds. *Journal of Acoustical Society of America*, 121(2):1119–1129, 2007. [21](#)
- [18] Michael S. Howe Richard S. McGowan. Comments on single-mass models of vocal fold vibration. *Journal of Acoustical Society of America*, 127(5):215–221, 2010. [21](#)
- [19] Jean Schoentgen Samia Fraj and Francis Grenez. Development and perceptual assessment of a synthesizer of disordered voices. *Journal of Acoustical Society of America*, 132(4):2603–2615, 2012. [23](#), [26](#)
- [20] Charles Severance. *Python for Informatics*. Charles Severance, 2009. [26](#)
- [21] Brad H. Story. Tubetalker: An airway modulation model of human sound production. *First International Workshop on Performative Speech and Singing Synthesis*, pages 989–1010, 2011. [23](#)
- [22] Brad H. Story. Phrase-level speech simulation with an airway modulation model of speech production. *Computer Speech and Language*, pages 989–1010, 2012. [23](#), [36](#), [37](#)
- [23] Ingo Titze. *Principles of Voice Production*. Prentice Hall, Eaglewood Cliffs, New Jersey 07632, 1994. [vii](#), [9](#), [10](#), [11](#), [15](#), [16](#), [17](#), [18](#), [19](#), [20](#), [21](#), [22](#), [25](#)
- [24] Ingo Titze. *The Myoelastic Aerodynamic Theory of Phonation*. National Center for Voice and Speech, 2006. [23](#)
- [25] Ingo R. Titze. Parameterization of the glottal area, glottal flow, and vocal fold contact area. *Journal of Acoustical Society of America*, 75(2):570–580, 1984. [28](#)
- [26] Ingo R. Titze. The physics of small-amplitude oscillation of the vocal folds. *Journal of Acoustical Society of America*, 83(4):1536–1552, 1988. [21](#), [22](#), [26](#), [27](#)
- [27] Ewald R. Weibel. *The Pathway For Oxygen*. Harvard University Press, Copyright by the President and Fellows of Harvard College, 1984. [vii](#), [13](#)
- [28] John M. Zelle. *Python Programming: an introduction to computer science*. Franklin, Beedle and Associates Incorporated, 8536 SW St. Helens Drive, Suite D, 2010. [26](#)