

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA - DAELT
ENGENHARIA ELÉTRICA**

**GEORGYA BROGIAN
JACQUELINE PALMAS
VANESSA CHRISTINE VON KRUGER**

**SÍNTESE DE SINAIS DE VOZ USANDO MODELOS
BIOLOGICAMENTE INSPIRADOS**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA

2018

GEORGYA BROGIAN
JACQUELINE PALMAS
VANESSA CHRISTINE VON KRUGER

SÍNTESE DE SINAIS DE VOZ USANDO MODELOS
BIOLOGICAMENTE INSPIRADOS

Trabalho de Conclusão de Curso de Graduação, apresentada à disciplina de Trabalho de Conclusão de Curso II, do Curso de Engenharia Elétrica, Departamento Acadêmico de Eletrotécnica (DAELT), da Universidade Tecnológica Federal do Paraná (UTFPR), como requisito parcial para obtenção do título de Engenheiro(a) Eletricista.

Orientador: Prof. Dr. Marcelo de Oliveira Rosa.

CURITIBA

2018

**GEORGYA BROGIAN
JACQUELINE PALMAS
VANESSA CHRISTINE VON KRUGER**

Síntese de sinais de voz usando modelos biologicamente inspirados

Este Trabalho de Conclusão de Curso de Graduação foi julgado e aprovado como requisito parcial para a obtenção do Título de Engenheira Eletricista, do curso de Engenharia Elétrica do Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR).

Curitiba, 26 de junho de 2018.

Prof. Antonio Carlos Pinho, Dr.
Coordenador de Curso
Engenharia Elétrica

Profa. Annemarlen Gehrke Castagna, Mestre
Responsável pelos Trabalhos de Conclusão de Curso
de Engenharia Elétrica do DAELT

ORIENTAÇÃO

Marcelo de Oliveira Rosa, Dr.
Universidade Tecnológica Federal do Paraná
Orientador

BANCA EXAMINADORA

Marcelo de Oliveira Rosa, Dr.
Universidade Tecnológica Federal do Paraná

Antonio Carlos Pinho, Dr.
Universidade Tecnológica Federal do Paraná

Elder Oroski, Dr.
Universidade Tecnológica Federal do Paraná

RESUMO

BROGIAN, Georgya; PALMAS, Jacqueline; VON KRUGER, Vanessa C.. **SÍNTESE DE VOZ USANDO MODELOS BIOLÓGICAMENTE INSPIRADOS**. 2018. 78p. Trabalho de Conclusão de Curso - Engenharia Elétrica - Universidade Tecnológica Federal do Paraná, 2018.

O uso de sinais de voz tem contribuído para o diagnóstico de patologias na laringe de uma maneira não-invasiva, e para desenvolver os estudos que utilizam-se desses sinais de voz sintetizados adotam-se representações realísticas da laringe. O modelo matemático de duas massas apresentado neste trabalho auxilia na análise de quais são os fatores físicos que afetam a vibração das cordas vocais, estruturas importantes no processo de produção da fala.

Palavras-chave: Laringe, modelo de duas massas, cordas vocais.

ABSTRACT

BROGIAN, Georgya; PALMAS, Jacqueline; VON KRUGER, Vanessa C.. **SYNTHESIS OF VOICED SOUNDS USING BIOLOGICALLY INSPIRED MODELS**. 2018. 78p. Trabalho de Conclusão de Curso - Engenharia Elétrica - Universidade Tecnológica Federal do Paraná, 2018.

The use of voice signals has contributed to the diagnosis of pathologies in the larynx in a non-invasive way and to develop studies using these synthesized voice signals is adopt realistic representations of the larynx. The mathematical model of two masses presented in this paper assists in the analysis of which are the physical factors that affect the vibration of the vocal cords that are important structures in the process of speech production.

Keywords: Larynx, two mass model, vocal cords.

LISTA DE FIGURAS

Figura 1 - Produção do Sinal Glotal.....	11
Figura 2 - Sistema Respiratório.....	20
Figura 3 - Vista posterior da Laringe.....	21
Figura 4 - Vista frontal da Laringe.....	21
Figura 5 - Vista em corte da Laringe.....	22
Figura 6 - Imagens de uma Videolaringoscopia.....	23
Figura 7 - Efeito da Passagem do Ar em Uma Folha.....	25
Figura 8 - Sistema de Produção da Voz.....	26
Figura 9 - Movimentos das Cordas Vocais.....	27
Figura 10 - Modelo de Uma Massa de Flanagan e Landgraf (1968).....	29
Figura 11 - Modelo de Ishizaka e Flanagan (1972).....	31
Figura 12 - Modelo de Três Massas de Titze (1994).....	32
Figura 13 - Modelo de Duas Massas Ishizaka e Flanagan.....	33
Figura 14 - Distribuição de Pressão ao Longo do Sistema de Produção de Voz.....	35
Figura 15 - Circuito Equivalente da Glote.....	39
Figura 16 - Aproximação do trato vocal por tubos cilíndricos.....	41
Figura 17 - Circuito T equivalente a um tubo cilíndrico.....	41
Figura 18 - Circuito acústico.....	42
Figura 19 - Fluxograma para o código de síntese de sinais vocálicos.....	47
Figura 20 - Modelo tridimensional da vogal /a/.....	49
Figura 21 - Dimensões do modelo da vogal /a/ em mm.....	50
Figura 22 - Vazão no tubo glotal e vazão irradiada pela boca para a condição padrão adotada.....	51
Figura 23 - Área glotal 1 e área glotal 2 para a condição padrão adotada.....	52
Figura 24 - Deslocamento x_1 e x_2 para a condição padrão adotada.....	53
Figura 25 - Simulação do efeito da variação das massas.....	54
Figura 26 - Variação da frequência em relação a variação da pressão subglotal.....	55
Figura 27 - Vazão do fluxo irradiado pela boca com $P_s = 2000 \text{ dyn/cm}^2$	55
Figura 28 - Simulação do efeito da variação da pressão subglotal.....	56
Figura 29 - Curvas correspondentes $k_2 = 124 \text{ kdyn/cm}$ (a) e $k_2 = 64 \text{ kdyn/cm}$ (b).....	57
Figura 30 - Curvas correspondentes à simulação na qual $k_2 = 224 \text{ kdyn/cm}$	58

Figura 31 - Curvas correspondentes às simulações nas quais $k_1 = 120$ kdyn/cm (a) e $k_1 = 80$ kdyn/cm (b).	59
Figura 32 - Curvas correspondentes à simulação na qual $k_1 = 280$ kdyn/cm.....	60
Figura 33 - Curvas correspondentes às simulações nas quais $k_c = 160$ kdyn/cm (a) e $k_c = 80$ kdyn/cm.	61
Figura 34 - Curvas correspondentes à simulação na qual $k_c = 360$ kdyn/cm.	62
Figura 35 - Simulação do efeito da variação da área neutra.....	63
Figura 36 - Simulação do efeito da variação da área neutra.....	64
Figura 37 - Simulação do efeito da variação da área neutra.....	65
Figura 38 - Simulação do efeito da variação do comprimento da corda vocal.	66
Figura 39 - Simulação do efeito da variação do comprimento da corda vocal.	67
Figura 40 - Simulação do efeito da variação do comprimento da corda vocal.	67
Figura 41 - Simulação do efeito da variação da espessura das massas.	68
Figura 42 - Simulação do efeito da variação da espessura das massas.	69
Figura 43 - Simulação do efeito da variação da espessura das massas.	69
Figura 44 - Voz humana masculina real gravada utilizando o MATLAB.	70
Figura 45 - Sobreposição dos sinais da vogal /a/ gravada com a vogal /a/ sintetizada.	72
Figura 46 - Laringe para diferentes patologias.....	73

LISTA DE SÍMBOLOS

- F_0 Frequência fundamental, [Hz]
- k Constante de mola linear, [dyn/cm]
- r Constante de amortecimento, [dyn.s/m]
- M Massa do modelo de uma massa, [g]
- P_S Pressão subglotal máxima, [cm/ H_2O]
- P_2 Pressão de saída do modelo de uma massa, [cm/ H_2O]
- l Comprimento da corda vocal do modelo de uma massa [cm]
- F Força do modelo de uma massa, [dyn]
- x Deslocamento lateral, [cm]
- M_1 Massa de uma corda vocal, [g]
- S_1 Força elástica residual, [dyn]
- M_2 Massa de uma corda vocal, [g]
- S_2 Força elástica residual, [dyn/cm²]
- r_1 Amortecedor (equivalente da resistência viscosa da massa M_1), [dyn.s/cm²]
- r_2 Amortecedor (equivalente da resistência viscosa da massa M_2), [dyn.s/cm²]
- P_{11} Pressão na zona de contração do modelo de duas massas, [cm/ H_2O]
- P_{12} Pressão entre a zona de contração e expansão, [cm/ H_2O]
- P_{21} Pressão entre a zona de contração e expansão, [cm/ H_2O]
- P_{22} Pressão expansão, [cm/ H_2O]
- k_c Rigidez da mola, [dyn/cm]
- U_g Vazão volumétrica, [cm³/s]
- P_1 Pressão de entrada utilizada no modelo de duas massas, [cm/ H_2O]
- S Força elástica residual, [dyn/cm²]
- x_1 Deslocamento lateral, [cm]
- x_2 Deslocamento lateral, [cm]
- F_1 Força que age sobre a massa 1, [dyn]

F_2 Força que age sobre a massa 2, [dyn]

A_{g1} Área glotal, [cm²]

A_{g01} Área transversal glotal neutra, [cm²]

l_g Comprimento da corda vocal, [cm]

A_{g2} Área glotal, [cm²]

A_{g02} Área transversal glotal neutra, [cm²]

f_{sj} Força necessária para produzir o deslocamento lateral x , [dyn]

k_j Rigidez linear, [dyn/cm]

x_j Deslocamento lateral, [cm]

η_{kj} Coeficiente que descreve a não-linearidade da mola

S_{kj} Força elástica, [dyn]

P_0 Pressão subglotal, [cm/H₂O]

P_b Pressão de Bernoulli, [cm/H₂O] [N/m²]

ρ Densidade do ar, [kg/m³]

μ Viscosidade dinâmica do ar, [P]

d Espessura da massa do modelo de uma massa, [cm]

A_1 Área, [cm²]

d_1 Espessura da massa M_1 , [cm]

d_2 Espessura da massa M_2 , [cm]

R_C Resistência da zona de contração, [Ω]

L_C Indutância, [H]

R_{v1} Resistência de fluxo independente utilizada no circuito acústico, [dyn.s/cm⁵]

L_{g1} Comprimento da corda vocal, [cm]

R_{12} Resistência na glote, [Ω]

R_{v2} Resistência de fluxo independente utilizada no circuito acústico, [Ω]

L_{g2} Indutância na glote, [H]

R_e Resistência da zona de expansão, [Ω]

η Coeficiente que descreve a não-linearidade da mola no modelo de uma massa

T Período, [s]

C_n Capacitância, [F]

R_n Resistências em série, [Ω]

s_j Perímetro da circunferência do tubo da j-ésima seção, [cm]

ω Frequência, [rad/s]

c Velocidade do som, [cm/s]

R_r Resistência da zona de irradiação, [Ω]

L_r Indutância da zona de irradiação, [H]

A_n Área, [cm²]

L_1 Indutância da malha 1, [H]

U_r Vazão volumétrica da malha de irradiação, [cm³/s]

U_1 Vazão volumétrica malha 1, [cm³/s]

U_2 Vazão volumétrica malha 2, [cm³/s]

U_3 Vazão volumétrica malha 3, [cm³/s]

l_{gd1} Área exposta, [cm]

l_{gd2} Área exposta, [cm]

$x_{1mín}$ Deslocamento da massa 1, [cm]

$x_{2mín}$ Deslocamento da massa 2, [cm]

P_{m1} Pressão média da massa 1, [dyn/cm²]

P_{m2} Pressão média da massa 2, [dyn/cm²]

P_r Pressão radiada, [dyn/cm²]

SUMÁRIO

1. INTRODUÇÃO	11
1.1 TEMA.....	11
1.1.1 Delimitação do Tema.....	13
1.2 PROBLEMAS E PREMISSAS	13
1.3 OBJETIVOS	14
1.3.1 Objetivo Geral	14
1.3.2 Objetivos Específicos	14
1.4 JUSTIFICATIVA	15
1.5 PROCEDIMENTOS METODOLÓGICOS.....	16
1.6 ESTRUTURA DO TRABALHO.....	16
2. FUNDAMENTAÇÃO TEÓRICA	18
2.1 APARELHO FONADOR E A PRODUÇÃO DA VOZ.....	18
2.2 PATOLOGIAS NA LARINGE	27
2.3 MODELOS MATEMÁTICOS DA LARINGE	28
3. ESTRUTURA E PREMISSAS PARA A SIMULAÇÃO COMPUTACIONAL DE SINAIS VOCAIS	40
3.1 OBTENÇÃO DAS EQUAÇÕES.....	40
3.2 MÉTODOS NUMÉRICOS	44
3.3 SIMULAÇÃO COMPUTACIONAL	46
4. RESULTADOS E DISCUSSÕES	48
4.1 SÍNTESE DA VOGAL /A/	49
4.2 ESTUDO DA VARIAÇÃO DOS PARÂMETROS.....	53
4.2.1 Efeito da variação das massas	53
4.2.2 Efeito variação da pressão subglotal	54
4.2.3 Efeito da variação dos coeficientes de elasticidade.....	56
4.2.4 Efeito da variação da área neutra.....	62
4.2.5 Efeito da variação do comprimento da corda vocal	65
4.2.6 Efeito da variação da espessura das massas	68
4.3 SÍNTESE DA VOGAL /A/ APÓS AJUSTES	70
4.4 RELAÇÃO DOS PARÂMETROS COM VOZES PATOLÓGICAS	72
5. CONCLUSÃO	74

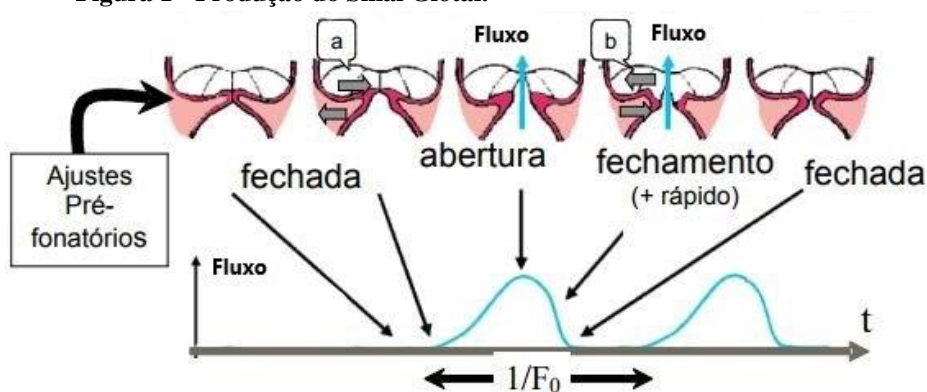
1. INTRODUÇÃO

1.1 TEMA

A laringe humana é um órgão localizado entre a base da língua e a traqueia, na região da garganta. É caracterizada pelas duas funções principais que desempenha: comporta-se como uma via aerífera, que controla a passagem de ar entre a via nasal e os pulmões, e é responsável pela produção do som.

A produção de som é um processo complexo que envolve as cartilagens, músculos e membranas que constituem a laringe. No entanto, o início do processo de produção do som está atrelado à contração e expansão dos pulmões. No momento em que é gerada uma diferença de pressão entre os pulmões e a região externa a boca, ocorre o escoamento (fluxo) do ar pela laringe (CATALDO *et al.*, 2004). Este fluxo é convertido em uma série de pulsos de ar, também denominados trem de pulsos ou sinais glotais, como descrito por Cataldo, Sampaio e Nicolato (2004). Estes pulsos, antes de se tornarem o sinal de voz em si ao fim deste processo, serão ainda modulados pela boca, dentes, lábios, entre outras articulações (CATALDO *et al.*, 2006).

Figura 1 - Produção do Sinal Glotal.



: Fonte: Vieira, 2004 (apud Hirano, 1981).

O trem de pulsos, formado por meio da produção do sinal glotal, é ilustrado e explicado por meio da Figura 1. Este complexo processo de produção de sinais (som), devido às características anatômicas e fisiológicas da laringe, pode ser modelado matematicamente por equações diferenciais ordinárias (EDOs), que descrevem os movimentos viscoelásticos dos tecidos da laringe. Existem várias formas de representar esse sistema por meio de analogias

feitas com sistemas físicos, geralmente sistemas mecânicos, como o modelo de Flanagan e Landgraf (1968), Ishizaka e Flanagan (1972) e Titze (1994). Esses modelos representam as cordas vocais por meio de uma, duas e três massas, respectivamente e em grau crescente de complexidade.

A busca pelo entendimento do processo de produção de som é motivada pelo papel que a voz desempenha na vida de um ser humano, já que esta é um dos principais meios de comunicação (CATALDO *et al.*, 2004). Desta forma, estudos realizados acerca do processo de produção de voz são de extrema importância. Principalmente quando visam desenvolver melhorias na área da medicina, como a criação de métodos alternativos no diagnóstico de patologias que acometem a laringe.

Novas contribuições neste sentido são necessárias, tendo em vista que o processo de identificação de patologias na laringe como afirmado por Souza (2011), é realizado de maneira invasiva por alguns procedimentos. Dentre eles a videolaringoscopia (exame dentre os mais adotados nesta área), podendo causar desconforto ao paciente. O que dificultaria a execução do exame e, por consequência, o processo de diagnóstico da uma dada doença.

Outro ponto a ser destacado, quando se trata de patologias na laringe, é a dependência da execução destes exames de uma solicitação médica. Na maioria das vezes, os procedimentos são feitos quando o paciente já se encontra em um quadro avançado, ou seja, quando a produção da voz já foi afetada e/ou quando o paciente já sente muita dor (SOUZA, 2011). O diagnóstico tardio afeta diretamente o tratamento da patologia.

Segundo Rosa (2002), o uso da voz para analisar as condições da laringe também tem sido empregado como uma forma não-invasiva de diagnóstico. Visto que, disfonias podem ser identificadas através da voz, pois esta carrega informação codificada em seus sinais.

Aproveitando-se desta característica da voz, o foco deste trabalho concentrou -se em descrever a laringe por meio de um sistema composto por massas, molas e amortecedores de uma forma mais completa, baseado no modelo de Ishizaka e Flanagan (1972), que se caracteriza por ser, dentre os modelos mais simples de representação da laringe, o que carrega mais informações e complexidade. E, a partir do novo modelo, esforços foram concentrados na produção de um sintetizador de sinais de voz, que permitiu simular a voz humana.

1.1.1 Delimitação do Tema

O trabalho proposto visa melhorar o equacionamento matemático da interação fluido-estrutura, que ocorre entre a laringe e o escoamento de ar que por ela passa e nela interfere. Juntamente com a identificação e delimitação de um intervalo de parâmetros, que permitiram o aperfeiçoamento do sinal da voz. A solução de cada um desses equacionamentos foi obtida por meio de programas que envolvem cálculo numérico, com objetivo de sintetizar sinais de voz saudáveis. O fonema escolhido para ser reproduzido pelo sintetizador foi o /a/, já que este é um dos mais citados na literatura.

A ferramenta computacional, empregada para obter a solução do modelo matemático, foi o MATLAB. Por meio deste *software* foram realizadas as simulações do modelo obtido a partir do estudo de equações que descrevem o funcionamento da laringe, juntamente com um modelo reduzido de escoamento de ar. O MATLAB, também foi utilizado para auxiliar na análise da qualidade dos resultados obtidos sinais sintetizados, ou seja, foram comparados ao modelo “ideal” (voz humana) e numericamente apresentaram quão longe ou perto esses modelos estão um do outro. A estratégia de comparação utilizada foi a distância vetorial, através da raiz do erro médio quadrático.

1.2 PROBLEMAS E PREMISSAS

Segundo Cataldo *et al.* (2004), as cordas vocais não se movem como corpos rígidos, este comportamento foi identificado por meio de observações do movimento das cordas vocais em câmera lenta. Com objetivo de tentar entender melhor o funcionamento da laringe, alguns modelos matemáticos foram construídos por estudiosos nessa área. Considerando a não rigidez das cordas vocais, dois modelos que envolvem o movimento de uma, duas ou três massas podem ser criados.

O modelo matemático mais significativo para este trabalho é o de Ishizaka e Flanagan (1972), que dada a sua importância será explicado detalhadamente no capítulo II. Os estudos desenvolvidos no presente trabalho têm como referência este modelo, pois a resolução do seu equacionamento é, em relação ao modelo de três massas, mais simples e, dentre dos estudos existentes com duas massas, o que traz mais informações em relação ao comportamento das

cordas vocais durante a fonação.

Ao analisar-se os modelos matemáticos já desenvolvidos, derivados de estudos acerca desse tema, notou-se que os equacionamentos matemáticos pecam em não abordar todos os fenômenos observáveis na laringe, distanciando, portanto, a modelagem do sistema real. No entanto, percebe-se que uma modelagem completa é muito trabalhosa, já que envolveria cálculos computacionais muito complexos e demorados. Portanto, a saída para esse problema, que é um dos objetivos deste trabalho, é deixar o equacionamento matemático, tratando as cordas vocais como um corpo com duas massas, o mais completo possível, conciliando uma forma computacionalmente pouco custosa.

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Estudar os modelos matemáticos que descrevem o funcionamento da laringe, por meio de um sistema constituído de massas, molas e amortecedores, tendo como base equacionamentos simples já existentes para síntese de fonemas vocálicos.

1.3.2 Objetivos Específicos

- ✓ Desenvolver programas em MATLAB para resolver as equações diferenciais ordinárias (EDOs) que descrevem o funcionamento da laringe;
- ✓ Alterar os parâmetros dos modelos matemáticos sistematicamente a fim de relacionar essas alterações à mudanças na frequência fundamental, amplitude e transitório do sinal;
- ✓ Analisar os efeitos que a alteração nos parâmetros dos equacionamentos causa no sinal resultante, em comparação aos sinais de referência;
- ✓ Comparar os efeitos devido às alterações dos parâmetros e observar as mudanças que ocorrem no sinal resultante sintetizado, relacionando-os a comportamentos típicos de vozes patológicas.

1.4 JUSTIFICATIVA

Ao modelar matematicamente a laringe, possibilita-se o estudo do principal meio de comunicação da sociedade, a voz. Sendo a laringe o órgão principal para a geração da fala humana, é importante o estudo das patologias associadas a ela. A ineficiência da laringe no processo de produção da voz não representa um risco a vida do paciente, porém dificulta sua convivência em sociedade. Em alguns casos, quando trata-se de profissionais que dependem da voz, a sua perda temporária pode causar dificuldades na execução do seu trabalho (ROSA,1998).

Para o diagnóstico de doenças na laringe, utilizam-se diversos métodos. Um deles é a laringoscopia, que permite visualizar as cordas vocais através de um pequeno espelho colocado na garganta. Há também a videolaringoscopia e nasofibroscopia, que tem o objetivo de visualizar a região da laringe e cordas vocais e auxiliar no diagnóstico de patologias. Um dos problemas relacionados a esses exames é o desconforto do paciente. Além disso, dependendo do especialista que analisa os exames pode-se obter diferentes diagnósticos.

Tomando as dificuldades apresentadas em diagnosticar patologias na laringe, busca-se um método que utilize os sinais de voz para detectar as doenças de forma não invasiva. Para analisar as condições da laringe precisa-se compreender seu funcionamento, para isso foram criados modelos matemáticos que o descrevem. O modelo de duas massas de Ishizaka e Flanagan (1972) foi o abordado no presente estudo.

Considerando as equações que descrevem essa modelagem, esse estudo as solucionou linearmente em programas criados em MATLAB, que resultaram na síntese dos sinais de voz. Possibilitando que em estudos futuros possa-se modificá-lo a fim de representar doenças na laringe.

Modelar matematicamente a laringe e a aproximar estes modelos de um estado natural, têm como objetivo alcançar, mesmo que a pequenos passos, contribuições para medicina. Já que, por meio do sintetizador de voz, alterações nos parâmetros do equacionamento significariam, na prática médica, mudanças realizadas na laringe por uma cirurgia, por exemplo. Desta forma, seria possível dizer previamente a um médico se as alterações pretendidas por ele, na laringe de um paciente, impactariam ou não na produção de som ao fim do procedimento e se isso seria prejudicial ou não para o paciente.

O objetivo do trabalho em questão, visa colaborar com estudos futuros que possam desenvolver uma forma alternativa de obter diagnósticos de patologias na laringe, que utilizem

apenas sinais de voz para identificar estas doenças.

1.5 PROCEDIMENTOS METODOLÓGICOS

Para o desenvolvimento do estudo da “síntese de sinais de voz usando modelos matemáticos biologicamente inspirados”, primeiramente foi realizada uma busca teórica na literatura para a compreensão do funcionamento do objeto de estudo (laringe) e fundamentação teórica em trabalhos já desenvolvidos neste campo relacionados à sua modelagem.

Em seguida, todos estes modelos foram combinados em um equacionamento, o que configurou o estado de interação fluido-estrutura. Então, foi feita a discretização do modelo matemático por meio de métodos numéricos. Para esta etapa, utilizou-se o programa MATLAB que sintetizou o fonema autossustentado.

Finalmente, testes envolvendo distância vetorial foram feitos por meio do MATLAB para analisar o quão próxima a voz (fonema) sintetizada chega de uma voz natural (humana). Para o estudo, a vogal escolhida foi a /a/. Estes testes também possibilitaram analisar como algumas alterações no equacionamento do modelo da laringe afetam a produção sintética da fala.

1.6 ESTRUTURA DO TRABALHO

Para uma melhor compreensão do tema em questão, o trabalho foi organizado na seguinte sequência de capítulos: o capítulo I tem por finalidade apresentar e delimitar o tema deste trabalho, e também a justificativa, os objetivos, e a metodologia utilizada para a pesquisa nesta área.

No capítulo II, foi realizada a fundamentação teórica acerca deste tema, sobre o qual apresentou-se o estudo da fisiologia da laringe juntamente com o processo de produção da voz por este órgão. A seguir, foram retratadas as principais patologias que acometem a laringe e geram os distúrbios nos sinais de voz. Para a obtenção futura dos parâmetros que se alteram no modelo matemático, que descreve a laringe, foram apresentados ainda neste capítulo os principais modelos propostos até então, com ênfase no modelo de duas massas.

Posteriormente, nos capítulos II e III, são apresentadas as equações diferenciais ordinárias que descrevem o processo da produção da fala, e também os métodos numéricos utilizados para a solução destas.

Os resultados e discussões atingidos, através da análise dos sinais obtidos por meio dos programas feitos e executados pelo MATLAB, foram discutidos no capítulo IV. Por fim, apresenta-se as conclusões e considerações feitas a partir das comparações dos resultados que foram obtidos tendo como referência de sinais de voz humana.

2. FUNDAMENTAÇÃO TEÓRICA

Como a voz é o principal meio de expressão e comunicação do ser humano, qualquer distúrbio que a afete pode trazer sérias implicações na vida social e profissional de uma pessoa. A voz é o resultado de fenômenos fisiológicos que são obtidos através da ação de uma série de estruturas, ou seja, a produção da fala não é formada por apenas um órgão (MORI, 2005, p.1). Sendo assim, para uma melhor compreensão desse processo, será apresentado neste capítulo como essas estruturas atuam para a geração da fala, juntamente com a descrição das patologias mais recorrentes que podem interferir e prejudicar a produção dos sinais de voz.

Para o estudo e desenvolvimento de um sintetizador de sinais vocálicos, inspirado na modelagem da laringe, se faz necessária também a compreensão dos modelos matemáticos que descrevem seu comportamento dinâmico. Para tanto, estes modelos serão abordados ao longo do presente capítulo, com ênfase ao modelo proposto por Ishizaka e Flanagan (1972), cujo trabalho serviu como inspiração e base para a implementação do sintetizador de voz.

2.1 APARELHO FONADOR E A PRODUÇÃO DA VOZ

A voz é produzida por meio de um conjunto de órgãos que têm como principal função a respiração humana. A atividade respiratória se faz necessária para a produção da voz, pois é através do fluxo de ar, que vem dos pulmões, que as ondas sonoras são transportadas (ROSA, 1998, p.8). Tanto o aparelho respiratório, o trato vocal e a laringe fazem parte do aparelho fonador e cada um desempenha um papel importante para a formação da voz.

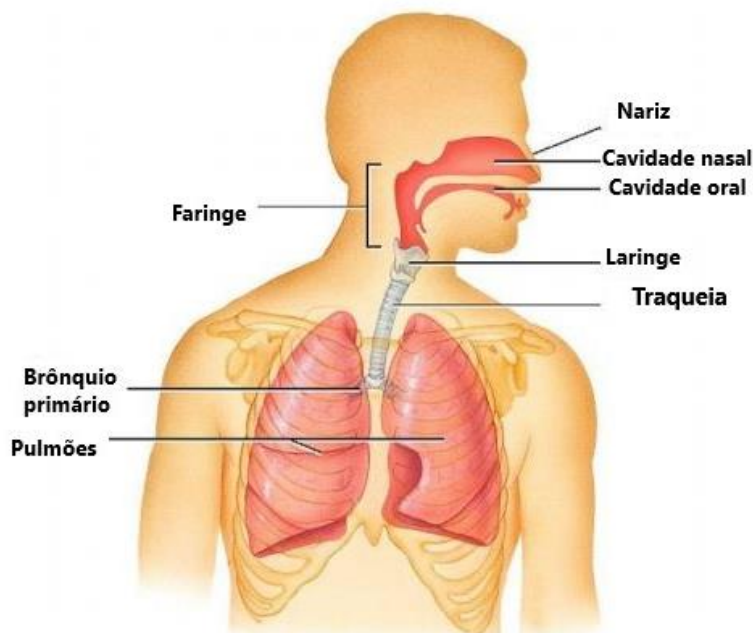
As estruturas que compõem o aparelho fonador podem ser classificadas em três grupos: o primeiro grupo é o da respiração, que se inicia nos pulmões e encerra no final da traqueia e é responsável pela produção do fluxo de ar que excitará o sistema responsável por produzir a fala. Em seguida, tem-se o grupo incumbido de produzir uma onda quase-periódica através do movimento das cordas vocais, esse grupo de órgãos situa-se na faringe. Já no último grupo, que constitui o trato vocal, tem-se o sistema ressonador encarregado de gerar os fonemas, que constituem o vocabulário acústico de uma língua (ROSA, 1998). Para uma melhor compreensão destes grupos, eles serão detalhados nos próximos parágrafos.

O grupo da respiração é formado pelo pulmão e a traqueia, sendo responsável pela

geração de energia sob a forma de um escoamento de ar do pulmão para a laringe. A traqueia é o meio pelo qual o ar é transportado para a laringe, funcionando como um duto para a sua passagem. Os pulmões são órgãos esponjosos que tem a sua base apoiada no diafragma, sua função principal é a hematose, porém contribui no controle de temperatura e participa da produção da voz (RAZERA, 2004). A hematose, também conhecida como troca gasosa, é a troca de gases presentes nos alvéolos pulmonares com o ar atmosférico, proporcionando que a quantidade de oxigênio presente no sangue seja a ideal (VERONEZ, 2012).

A inspiração e a expiração são movimentos de contração e expansão das paredes dos pulmões, que são realizados pelo movimento da caixa torácica. A expiração é caracterizada por um relaxamento natural dos pulmões (colapso passivo do tórax), já que essas estruturas foram forçadas a ter um posicionamento que não é natural quando é realizada a inspiração (ROSA, 1998).

Quando não existe força para manter o ar no pulmão (insuflado), o ar é expelido através da traqueia, os movimentos do diafragma e das costelas fazem com que o pulmão tenha movimentos de expansão e contração (MORI, 2005). Na produção da voz, os pulmões são responsáveis por gerar o fluxo de ar que carrega as ondas sonoras. Além dos pulmões, que armazenam o ar, o sistema respiratório do ser humano é também constituído pelos brônquios que se unem à traqueia e pelo diafragma que separa a cavidade torácica da cavidade abdominal (ROSA, 1998). A Figura 2 ilustra o sistema respiratório, apresentando a posição de cada um dos órgãos no sistema respiratório.

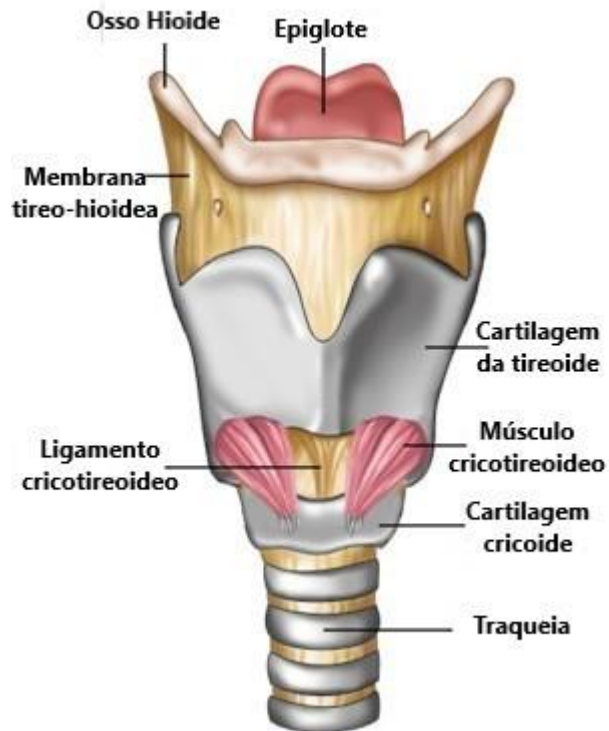
Figura 2 - Sistema Respiratório.

Fonte: Adaptado de Scalassara (2009, p.10).

O segundo grupo é formado pela laringe, a qual possui como função impedir que corpos estranhos atinjam o pulmão, atuando como uma proteção. Sua função secundária é a criação do sinal glotal, que ocorre particularmente devido a estrutura das cordas vocais. É possível dizer que a laringe funciona como um controlador de vazão de ar (ROSA, 2002).

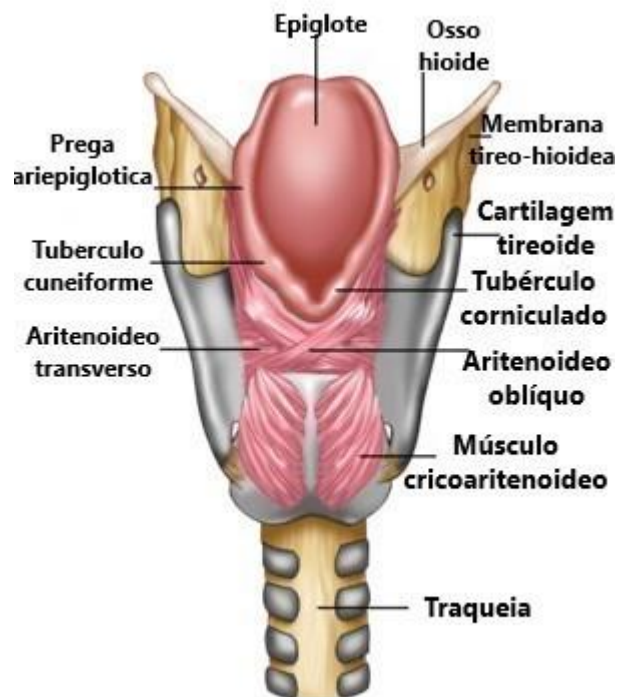
A laringe é composta por cartilagens, osso hioide, músculos intrínsecos: que possuem origem de dentro da laringe e extrínsecos: que se ligam a laringe, mas não se formam dela; nervos, ligamentos e lâminas de aponeurose (RAZERA, 2004). Existem no total nove cartilagens presentes na laringe que são sustentadas pelo osso hioide, sendo classificadas como pares as corniculadas, as cuneiformes e as aritenoideas e como ímpares: a tireoidea, a cricoidea e a epiglote (DAJER, 2010, p.25). As Figuras 3 e 4 e 5 ilustram as principais estruturas da laringe, as posições das cartilagens, cordas vocais verdadeiras e falsas e do osso hioide.

Figura 3 - Vista posterior da Laringe.



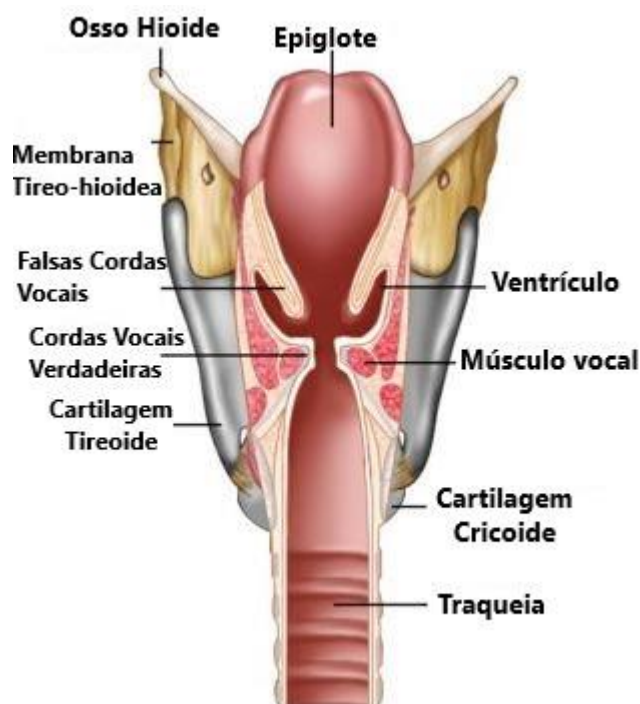
Fonte: Adaptado de <Anatomiaonline.com/respiratorio/laringe/laringe.htm>.

Figura 4 - Vista frontal da Laringe.



Fonte: Adaptado de <Anatomiaonline.com/respiratorio/laringe/laringe.htm>.

Figura 5 - Vista em corte da Laringe.



Fonte: Adaptado de <Anatomiaonline.com/respiratorio/laringe/laringe.htm>.

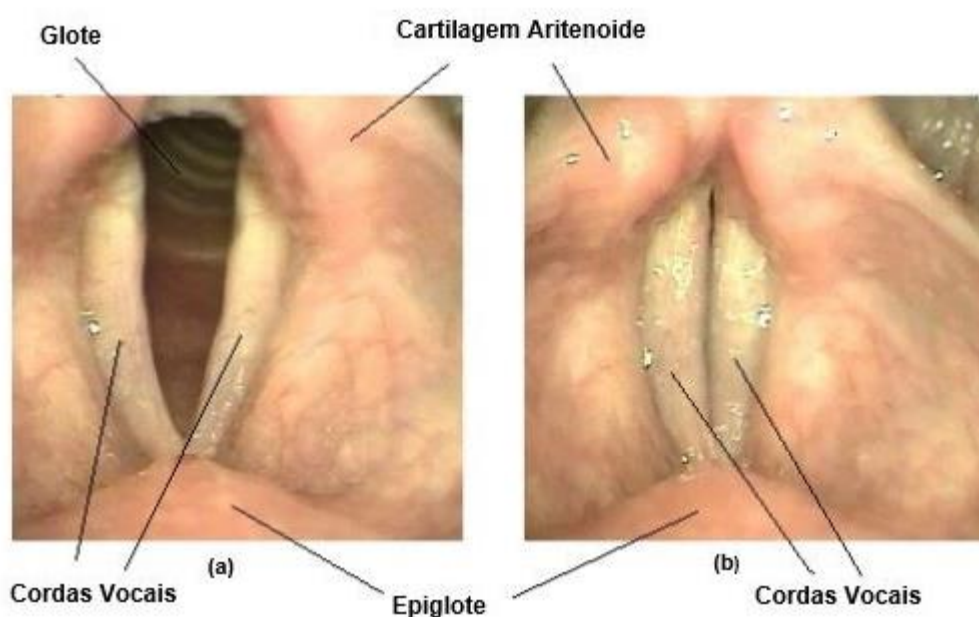
Pode-se dividir a laringe em quatro partes para explicar suas cavidades, que são: vestíbulo, também denominado pregas vestibulares, que está situado desde a epiglote até as falsas cordas vocais; os espaços entre as cordas vocais verdadeiras e falsas, que são chamados de ventrículos larínges; a glote e a cavidade infraglótica que está localizada na extensão entre a traqueia e as cordas vocais verdadeiras, sendo esta última a primeira resistência aerodinâmica da laringe (ROSA, 2002).

Segundo Cassol (2004, p. 26), *apud* Hungria (2000), existem cinco cartilagens principais na laringe: a tiroide, que atua como proteção da laringe; a cricoide, que está situada abaixo da cartilagem tiroide e acima da traqueia; as aritenoides, nas quais as cordas vocais estão posicionadas e as cartilagens epiglóticas, que protegem as vias aéreas inferiores. Os músculos intrínsecos controlam a produção do som e os extrínsecos sustentam a laringe.

As falsas cordas vocais, também conhecidas como cordas vocais vestibulares, diminuem a impedância aerodinâmica no processo de vocalização, e possuem superfície maior que a das cordas vocais verdadeiras (ROSA, 2002). O espaço entre as cordas vocais verdadeiras é denominado glote. Sendo esta responsável pela regulação da quantidade de ar que circula no aparelho fonador, estabelecendo o equilíbrio de oxigênio e gás carbônico (VIEIRA, 2003).

De acordo com Rosa (2011), a ligação da laringe com as demais estruturas do corpo ocorre por meio dos músculos extrínsecos, que são já a movimentação das cartilagens internas da laringe, acontece devido aos músculos intrínsecos que são responsáveis por abrir e fechar o espaço glotal. Esse movimento controla a vibração das cordas vocais, desse modo a fonação é dependente da ação desses músculos (ROSA, 2011). A Figura 6 (a) apresenta as cordas vocais em abdução, ou seja, quando elas estão abertas e há passagem do fluxo de ar através da glote. Na Figura 6 (b) essas encontram-se em adução, movimento de fechamento das cordas vocais, no qual a passagem de ar do pulmão para a boca é interrompida. A glote está ilustrada na Figura 6 (a).

Figura 6 - Imagens de uma Videolaringoscopia.



Fonte: Adaptado de Scalassara (2009 p.10).

É importante salientar, que o escoamento do ar, no estreitamento da glote, e a diferença histológica das cordas vocais influenciam na vibração destas (ROSA, 2002). Na respiração e na fonação, a musculatura intrínseca possui funções importantes, sendo responsável pela abdução e adução das cordas vocais (DAJER, 2010).

Por fim, o terceiro grupo é constituído pelo trato vocal conhecido como sistema ressonador. Ele é composto pela boca, nariz, faringe e também a cavidade torácica (DAJER, 2010), pode ser representado por tubos de ressonância que amplificam o sinal glotal (MORI, 2005, p. 8). O sinal da glote é produzido com baixa intensidade, para que os fonemas sejam caracterizados é necessária a amplificação do sinal, de modo que componentes harmônicas sejam enfatizadas. A língua é importante na variação do tamanho da área de ressonância, ela

faz parte da cavidade bucal que também contém dentes e lábios (ROSA, 1998). Modificando o formato do trato vocal, altera-se os sons básicos gerados pela laringe, o que faz com que existam uma variedade de timbres sonoros (RAZERA, 2004, p. 8).

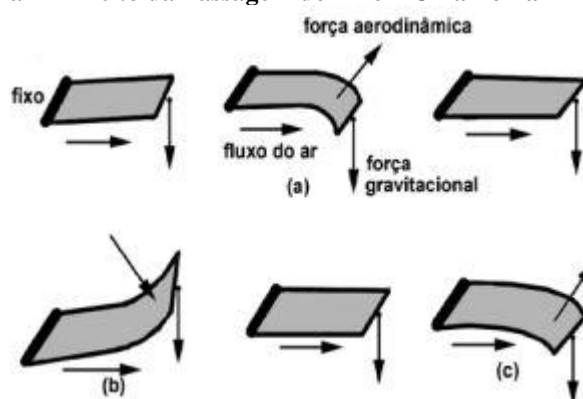
Segundo Scalassara (2009), é possível dividir em três grupos os sons produzidos pelos seres humanos, conforme a variação da forma de excitação. Existem os sons vocálicos, que são os pulsos quase periódicos de ar, no trato vocal, produzidos a partir da tensão da vibração das cordas vocais ocasionados pela força do ar na glote. Scalassara (2009, p. 12) *apud* (Rabiner e Schafer, 1978) afirma que os sons não-vocálicos são produzidos a partir de um estreitamento do trato vocal, geralmente localizado na boca, que faz com que a pressão do ar que está passando aumente, causando uma turbulência que gera um ruído no trato. O terceiro grupo são os sons plosivos, que são gerados no momento da liberação de ar depois do trato vocal ter fechado e ocasionado um aumento de pressão. Os sons vocálicos emitem a produção de vogais, produzem a vibração das cordas vocais e ressonância completa do trato vocal. São ideais para estudo por conterem bastante informação fisiológica do sistema (SCALASSARA, 2009).

O sinal glotal possui características importantes e que, devido a anatomia e fisiologia da laringe, são difíceis de se reproduzir (SAMPAIO *et al.*, 2006, p.107). Nos primeiros estudos realizados nesta área, imaginava-se que os movimentos das cordas vocais eram gerados a partir de estímulos neurológicos. Foi verificado que a laringe possui capacidade de vibração com frequências muito mais rápidas do que os pulsos neurais seriam capazes de estimular. Com isso, a ideia inicial de que os estímulos, para a movimentação das cordas vocais, partissem do sistema neurológico foi descartada (ROSA, 2002). De forma indireta, a ação neural contribui para a movimentação das cordas vocais, controlando a ação dos músculos que contribuem para a criação do movimento oscilatório (SAMPAIO *et al.*, 2006, p.107). Foi proposta então a teoria mioelástica - aerodinâmica da laringe, a qual definiu que as propriedades dos tecidos musculares da laringe e as forças aerodinâmicas, são responsáveis pelo movimento das cordas vocais (ROSA, 2002). A interação da elasticidade dos músculos da laringe com as forças aerodinâmicas da respiração, são combinadas nessa teoria fazendo relação com a descrição do efeito de Bernoulli na laringe. Quando o ar em alta velocidade passa pela laringe, a pressão entre as cordas vocais diminui, com isso há a aproximação entre as cordas vocais. Na sequência ocorre um retrocesso elástico, que causa a adução glótica e o ciclo de vibração recomeça (BEHLAU, 2001).

Rosa (2011) compara o fenômeno que ocorre na laringe com o efeito gerado quando alguém assopra uma folha de papel. O ar, em alta velocidade, atravessa o papel gerando uma pressão menor na extremidade inferior da folha. A diferença de pressão criada, faz com que que

haja uma pressão de baixo para cima sobre a folha, tomando como referência o ponto fixo da Figura 7, movimentando-a neste sentido. Conforme a folha se movimenta, as forças aerodinâmicas são modificadas alterando a direção e intensidade das mesmas sobre o papel. Quando a força aerodinâmica for menor que a força gravitacional, a folha é empurrada para baixo, esse ciclo termina quando surge novamente a força de baixo para cima, que eleva a folha para cima. O efeito da passagem do ar na folha se assemelha com a vibração das cordas vocais. A Figura 7 ilustra esse efeito, (a) quando a força aerodinâmica é maior, (b) quando a força aerodinâmica é menor ou de mesmo sentido, e (c) quando a força aerodinâmica é maior.

Figura 7 - Efeito da Passagem do Ar em Uma Folha



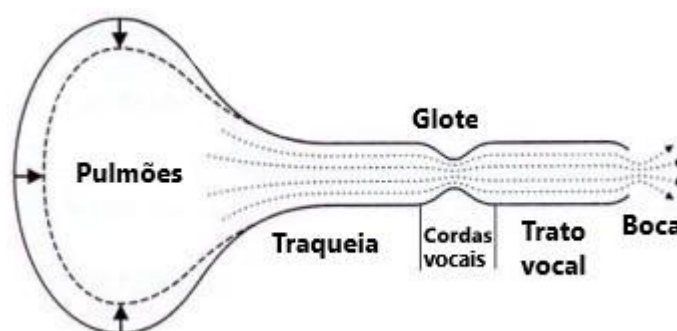
Fonte: Rosa (2011).

Complementarmente, segundo Cveticanin (2012), a produção da voz ocorre através da movimentação de duas cordas vocais opostas localizadas na laringe, decorrente do fluxo de ar gerado pelo pulmão que atravessa a traqueia. Essa movimentação ocorre a partir da interação das forças que são aplicadas na superfície das cordas vocais, com as forças que são geradas no interior do tecido muscular das mesmas, que é contínua e auto-sustentada se as condições de oscilação forem mantidas (ROSA, 2011, p.74). As cordas oscilam, se aproximando para a fonação e se afastando, uma da outra, para a respiração, o que gera pequenas ondas sonoras (CVETICANIN, 2012). A pressão aumenta quando as cordas vocais estão próximas. Quando atingem a pressão suficiente, as cordas se separam (abrem) liberando o ar, que funciona como um impulso para que se inicie as vibrações (SCALASSARA, 2009). Essas vibrações são ressonadas pelo trato vocal durante a passagem do fluxo de ar pelos ciclos de abertura e fechamento (FUKUYAMA, 2001).

Segundo a teoria mioelástica aerodinâmica, o tamanho da área glotal no momento da fonação, a pressão subglotal, as características dos tecidos das fibras das cordas, assim como a tensão aplicada nas cordas vocais verdadeiras e a massa superficial das mesmas, são fatores que afetam a autossustentação da geração das vibrações (ROSA, 2011, p.74).

A diferença de pressão, entre o ar presente no pulmão e o que está fora da boca, produz uma onda quase periódica. Ao longo da passagem do escoamento de ar através da laringe, o fluxo contínuo de ar no pulmão é transformado em uma onda pulsátil, até chegar na boca e na cavidade nasal. Os pulsos gerados são conhecidos como sinal glotal ou trem de pulsos, a modulação do sinal glotal é feita através dos lábios, língua e dentes, gerando o que é conhecido como voz (CATALDO *et al.* , 2004, p. 48). O esquema do sistema de produção da voz está ilustrado na Figura 8.

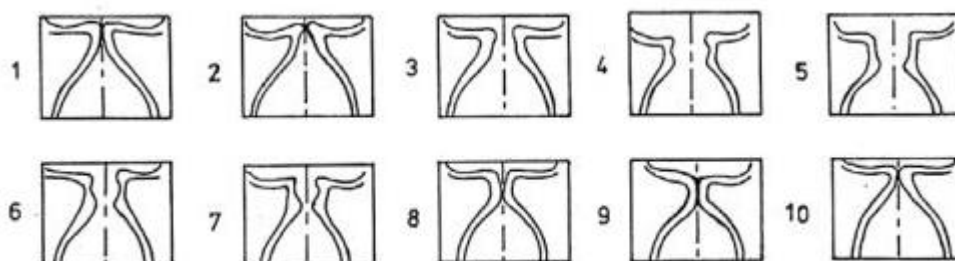
Figura 8 - Sistema de Produção da Voz.



Fonte: Cataldo; Sampaio; Nicolato (2004).

Durante a respiração normal, as estruturas relacionadas a voz permanecem relaxadas. Quando a expiração se inicia, há um aumento da pressão subglótica ocasionado pela aproximação das cordas vocais, representado pela Figura 9 (1). Nas Figuras 9 (2), 9 (3) e 9 (4), é ilustrado a abertura das cordas vocais, decorrente da perda de energia das cordas e do aumento da pressão subglótica, que permitem a passagem do fluxo de ar através da abertura glótica. Há uma diminuição da pressão subglótica causada pela passagem do ar. No movimento de abertura, as cordas vocais obtêm energia elástica, o que pode ser observado por meio da Figura 9 (5), 9 (6) e 9 (7). Com o auxílio do efeito de Bernoulli, a energia acumulada permite que as cordas vocais se aproximem, superando a pressão do ar. Esse movimento está ilustrado nas Figuras 9 (8), 9 (9) e 9 (10). O ciclo vibratório ocorre enquanto durar a vocalização (RAZERA, 2004, p.8).

Figura 9 - Movimentos das Cordas Vocais.



Fonte: Hirano (1981).

2.2 PATOLOGIAS NA LARINGE

De acordo com Hansen *et al.* . (1998, p. 302), a eficiência da voz depende da vibração das cordas vocais e da forma como ocorre o fechamento e a abertura da glote. Patologias na laringe podem impedir que as cordas vocais possuam uma vibração glotal harmônica, e alguns fatores devem ser considerados a respeito das vibrações nas cordas vocais. Dentre estes fatores estão a posição da corda vocal, ou a extensão que as cordas vocais adquirem ao serem estiradas ou contraídas, a mioelasticidade, ou o grau de elasticidade das cordas vocais, e o nível de pressão do ar que passa pelas das cordas vocais (ZWETSCH *et al.*, 2006, p. 110).

Patologias na laringe podem causar problemas com os músculos respiratórios. A paralisia é um exemplo desse tipo de doença, ela ocasiona uma pressão subglotal insuficiente que altera as vibrações das cordas vocais. Sendo assim, o sinal de voz resultante é afetado por patologias na laringe e pode ser usado para medir os desvios ocasionados por ela (SCALASSARA, 2009, p. 19).

No Quadro 1 apresenta-se, de maneira compacta, algumas patologias que acometem mais comumente a laringe. Estas doenças foram apresentadas em Rosa (1998) e são descritas suas características quanto às alterações fisiológicas e acústicas.

Quadro 1 - Patologias e suas alterações.

Patologia	Alteração fisiológica	Alteração acústica
Edema de Reinke	Cobertura da corda vocal menos rígida e vibração como uma estrutura mais massiva.	Frequência fundamental reduzida tornando a voz mais grave e sopro.
Diplofonia	Alteração de massa-tamanho-tensão nas cordas vocais, gerando vibrações diferentes entre elas.	Aparecimento de um sinal com duas frequências distintas.
Nódulos	Protuberância esbranquiçada na borda glótica das cordas vocais. Interfere na vibração dependendo da firmeza e aumento da massa.	Sinal glotal aperiódico.
Pólipos	Aumento da massa vibratória, aumento da rigidez em geral, e vibração assimétrica.	Voz sopro, e dependendo do tamanho do pólipo e da localização pode causar também rouquidão.
Carcinoma (Câncer)	Surgimento de tumores.	Pode variar de uma voz levemente sopro até uma rouquidão severa.
Cistos	Protuberância na corda vocal resultante de um acúmulo de líquidos.	A voz possui aspecto sopro.
Granuloma	Mucosa apresenta granulação e aspecto rígido.	Níveis de rouquidão e sopro elevados.
Paralisia de pregas vocais	Paralisia unilateral ou bilateral das cordas vocais, o que impede a produção adequada de um sinal glotal.	Voz sopro.

Fonte: Adaptado de Rosa (1998).

2.3 MODELOS MATEMÁTICOS DA LARINGE

De maneira sucinta, a produção de voz está associada ao movimento lateral das cordas vocais, localizadas na laringe, quando o ar vindo do pulmão passa pela traqueia (CVETICANIN, 2012). O sinal obtido por meio desse processo, de acordo com Montagnoli e Fonseca (2005), que seja registrado sob condições apropriadas, carregará todas as informações sobre as variações acústicas a que este foi submetido, durante seu curso pelo sistema de produção de voz.

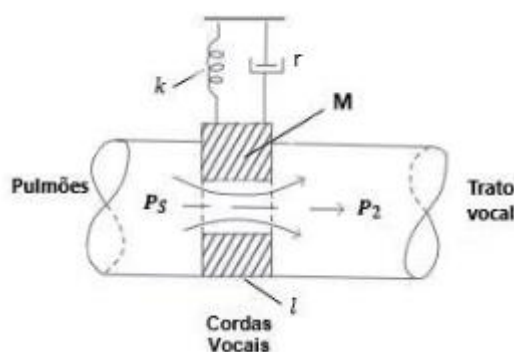
Tendo em vista que os sinais de voz carregam estes detalhes, sua simulação

computacional possibilita o seu estudo e entendimento de forma aprofundada. Já que, como afirma Rosa (2002), a reprodução destes sinais torna possível a observação de diversos fenômenos físicos que ocorrem na laringe, permitindo uma melhor compreensão dos mecanismos que os criaram. A título de exemplo, por meio de simulações, é possível manipular os parâmetros presentes no modelo de base de tal forma a reproduzir-se computacionalmente um sinal de voz real específico. O que dá margem a novas formas de diagnóstico de patologias presentes na laringe pela medicina. Por esse motivo, a modelagem matemática do processo de geração da voz é tão relevante, tendo em vista que estes equacionamentos são os alicerces que tornam possível a simulação destes sinais. Adicionalmente, os resultados obtidos por meio dos modelos matemáticos na reprodução de sinais vocais têm se mostrado realísticos, pois estes equacionamentos incorporam tanto as características anatômicas quanto fisiológicas da laringe (ROSA, 2002).

O estudo acerca da dinâmica das cordas vocais tem sido realizado há muito tempo. No entanto, independentemente do surgimento de novos modelos, estes sempre se mostraram fundamentados em modelos criados anteriormente (CATALDO *et al.*, 2004). Segundo Rosa (2002), os modelos matemáticos que explicam o processo de produção de voz diferenciam-se em grau de complexidade e no entendimento subjetivo do funcionamento das cordas vocais.

O modelo de produção de voz a ser abordado pelo presente trabalho é conhecido mais comumente por modelo de uma massa. Este modelo mecânico foi proposto por Flanagan e Landgraf em 1968. É considerado uma das representações mais simples da dinâmica das cordas vocais, tendo em vista que não leva em consideração as diferentes camadas que constituem as cordas vocais. O esquemático representativo deste modelo mecânico pode ser observado na Figura 10.

Figura 10 - Modelo de Uma Massa de Flanagan e Landgraf (1968).



Fonte: Adaptado de Cveticanin (2012, p.3).

No modelo mostrado na Figura 10, cada uma das cordas vocais é representada por um

sistema que consiste em uma massa M , uma mola linear k e um amortecedor r . O papel de cada uma destas grandezas físicas, concentra-se em representar respectivamente a massa das cordas vocais, sua elasticidade e seu amortecimento: capacidade de perder energia. A excitação das cordas vocais, ou seja, sua vibração, é provocada por uma força F , que está relacionada ao fluxo de ar vindo dos pulmões, pressão variada, que incide sobre a área da superfície intra-glotal.

A dinâmica deste sistema, composto por uma massa, mola e amortecedor, é descrita por meio da Equação (1), no qual x é o deslocamento sofrido pela massa M de uma das cordas vocais e k representa o coeficiente de elasticidade da mola.

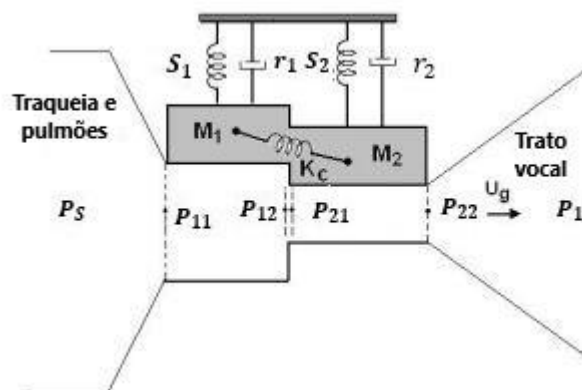
$$M\ddot{x} + r\dot{x} + kx = F. \quad (1)$$

Apesar da simplicidade, uma consideração feita por esse modelo que gerou significativa contribuição para estudos realizados posteriormente, foi a escolha do princípio de Bernoulli como fenômeno responsável pela geração de pressão negativa sob as cordas vocais (ROSA, 2002). Este princípio relaciona a velocidade à pressão estática, sendo estas grandezas inversamente proporcionais.

Apesar do modelo de uma massa, apresentado acima, ter produzido bons resultados na síntese de sinais de voz, juntamente com a reprodução de várias das propriedades do fluxo glotal, este modelo não foi capaz de reproduzir alguns dos comportamentos fisiológicos das cordas vocais (ISHIZAKA; FLANAGAN, 1972). Com o intuito de resolver este problema, Ishizaka e Flanagan deram sequência ao estudo da dinâmica das cordas vocais desenvolvido por Flanagan e Landgraf e, em 1972, propuseram um modelo matemático constituído de duas massas.

Neste modelo, como pode ser observado na Figura 11, cada uma das cordas vocais consiste em duas massas, M_1 e M_2 . Sendo que, a cada uma destas massas, estão associadas diferentes elasticidades, S_1 e S_2 , e amortecimentos, r_1 e r_2 , que representam as características viscoelásticas das cordas.

Figura 11 - Modelo de Ishizaka e Flanagan (1972).



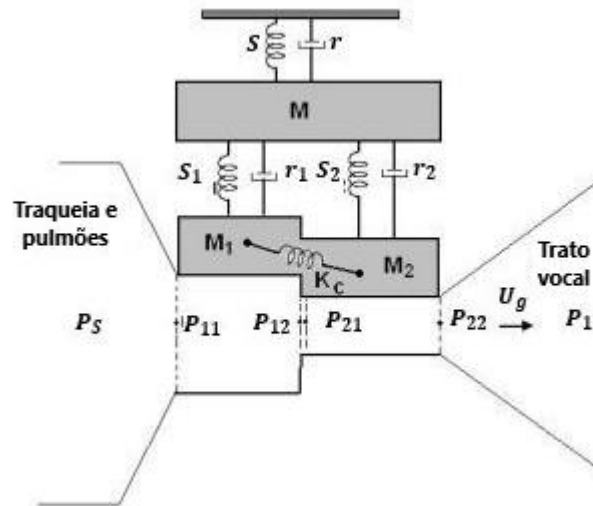
Fonte: Adaptado de Cataldo; Sampaio; Nicolato (2004).

Conforme citação de Ishizaka e Flanagan (1972, p. 1235), trabalhos teóricos comprovam que este modelo é capaz de reproduzir satisfatoriamente grande parte dos detalhes presentes na glote, como a diferença de fase das extremidades superiores e inferiores das cordas vocais e a oscilação para uma impedância de entrada capacitiva do trato vocal. Talvez seja este, o principal motivo deste modelo ser largamente utilizado na literatura como base para o desenvolvimento de outras pesquisas.

Os equacionamentos que expressam a dinâmica das cordas vocais para o modelo de duas massas, assim como outros equacionamentos relevantes para o presente trabalho, serão apresentados e explicados com minuciosidade no item subseqüente.

Outro modelo mecânico, com um grau maior de complexidade, foi proposto por Titze em 1994. Seu trabalho acrescenta uma massa em relação ao modelo apresentado anteriormente, por isso este trabalho também é conhecido por modelo de três massas, como pode ser verificado na Figura 12.

Figura 12 - Modelo de Três Massas de Titze (1994).



Fonte: Adaptado de Cataldo; Sampaio; Nicolato (2004, p.51).

Titze considerou cada uma das cordas vocais como um sistema composto de três massas, M , M_1 e M_2 , de molas não-lineares, S , S_1 e S_2 , que representam as tensões equivalentes nas cordas e de amortecedores r , r_1 e r_2 que simbolizam a inércia, que as cordas vocais apresentam, em voltar ao seu estado original.

Assim como os modelos anteriores, o modelo de três massas também considera somente o deslocamento no sentido perpendicular ao tubo vocal das massas M , M_1 e M_2 . Esses deslocamentos são chamados, respectivamente, de x , x_1 e x_2 . As equações (2), (3) e (4), descrevem a dinâmica das cordas vocais do modelo estudado por Titze.

$$M\ddot{x} + r\dot{x} + S(x) + S_1(x - x_1) + r_1(\dot{x} - \dot{x}_1) + S_2(x - x_2) + r_2(\dot{x} - \dot{x}_2) = 0, \quad (2)$$

$$M_1\ddot{x}_1 + S_1(x_1 - x) + r_1(\dot{x}_1 - \dot{x}) + k_c(x_1 - x_2) = F_1, \quad (3)$$

$$M_2\ddot{x}_2 + S_2(x_2 - x) + r_2(\dot{x}_2 - \dot{x}) + k_c(x_2 - x_1) = F_2, \quad (4)$$

sendo F_1 e F_2 as forças necessárias para deslocar verticalmente as massas em uma distância de x , x_1 e x_2 , respectivamente.

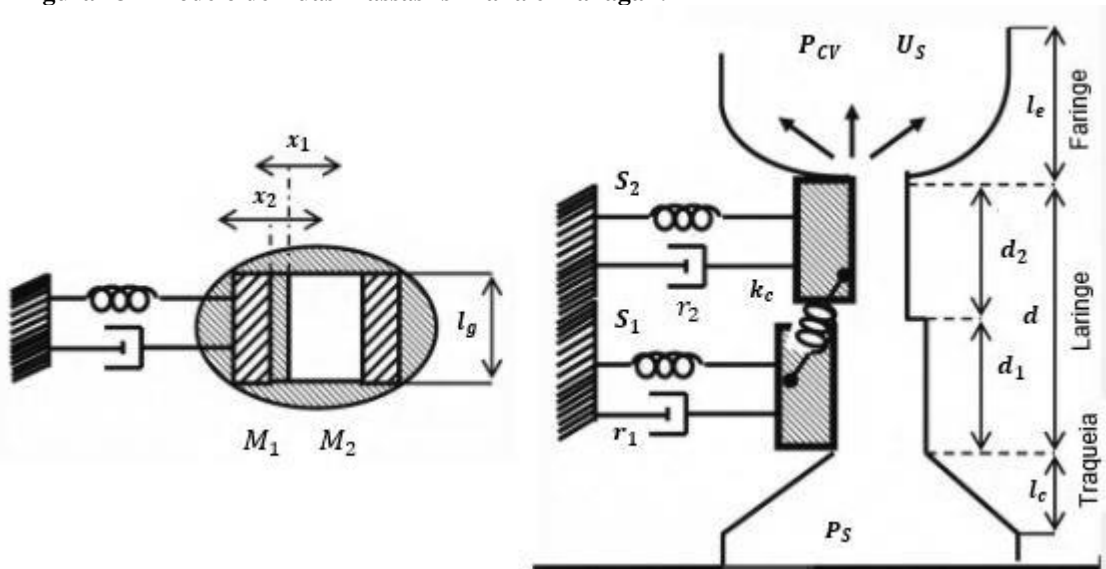
Nota-se que a complexidade dos modelos matemáticos aumenta proporcionalmente ao nível de detalhe e considerações durante a modelagem matemática, como pode ser observado neste último caso abordado. Segundo Cataldo *et al.*, (2004), em um artigo que compara os três modelos mecânicos apresentados acima, a síntese computacional de sinais de voz usando o modelo de três massas requer um tempo maior de CPU, *Central Process Unit*, para sua execução em relação aos outros modelos anteriores. É neste momento, portanto, em que deve ser avaliado até que ponto faz sentido o aumento de complexidade de um sistema para a

obtenção de um resultado mais fidedigno, tendo em vista o custo computacional que essa riqueza de detalhes irá acarretar.

2.4 MODELO DE DUAS MASSAS

O modelo de duas massas proposto por Ishizaka e Flanagan em 1972, considera o sistema para a dinâmica das cordas vocais como um sistema massa, mola e amortecedor com dois graus de liberdade. Este modelo determina que cada uma das cordas vocais é um sistema de duas massas com movimentos simétricos, por isso seus equacionamentos são análogos. Essas massas estão ligadas à laringe por duas molas não lineares S_1 e S_2 , essas massas são acopladas entre si por uma mola linear k_c como mostrado na Figura 13.

Figura 13 - Modelo de Duas Massas Ishizaka e Flanagan.



Fonte: Adaptado de Teffahi (2009, p.1).

A representação do sistema de duas massas na Figura 13 é diferente da realizada por Cataldo, Figura 11. Observa-se ambos os sistemas são iguais, no entanto a figura que ilustra um deles está rotacionada em relação a outra em 90° . Visando evitar possíveis confusões e para uma melhor compreensão acerca da posição das cordas vocais na laringe, a representação do sistema expresso pela Figura 13 foi adaptada, já que neste formato há concordância entre os deslocamentos das massas das cordas vocais, chamados de 'x', com o eixo ao longo do qual esses deslocamentos ocorrem, eixo 'X' ou eixo das abscissas.

Como pode ser observado pela Figura 13, as massas movem-se somente na direção transversal e seus deslocamentos são descritos pelas seguintes equações:

$$M_1\ddot{x}_1 + r_1\dot{x}_1 + s_1 + k_c(x_1 - x_2) = F_1, \quad (5)$$

$$M_2\ddot{x}_2 + r_2\dot{x}_2 + s_2 + k_c(x_2 - x_1) = F_2, \quad (6)$$

em que:

$$s_i(x_i) = k_i(x_i + \eta_{ki}x_i^3), \text{ para } x_i > -\frac{A_{goi}}{2l_g}, \quad (7)$$

$$s_i(x_i) = k_i(x_i + \eta_{ki}x_i^3) + h_i \left\{ \left(x_i + \frac{A_{goi}}{2l_g} \right) + \eta_{ki} \left(x_i + \frac{A_{goi}}{2l_g} \right)^3 \right\}, \text{ para } x_i \leq -\frac{A_{goi}}{2l_g}. \quad (8)$$

As áreas da região existente entre as cordas vocais e a referência, que é a linha de simetria entre elas, são expressas pelas equações (9) e (10) a seguir:

$$A_{g1} = A_{g01} + 2l_g x_1, \quad (9)$$

$$A_{g2} = A_{g02} + 2l_g x_2. \quad (10)$$

As molas S_1 e S_2 são representações equivalentes da tensão nas cordas vocais. E a relação não-linear entre a deflexão, a partir da posição de equilíbrio, e a força requerida para produzir essa deflexão é dada por:

$$f_{sj} = k_j x_j (1 + \eta_{kj} x_j^2), \quad j = 1, 2, \quad (11)$$

em que f_{sj} a força requerida para produzir o deslocamento x_j , k_j é a rigidez linear e η_{kj} é o coeficiente que descreve a não-linearidade da mola S_j .

Considerando que há uma força que deforma as cordas vocais no momento em que há a colisão das massas, fechamento da glote, tem-se que a força de restauração durante esse processo pode ser representada por uma mola S_{hj} . A força requerida para produzir a deformação na massa possui a seguinte relação:

$$f_{hj} = h_j \left(x_j + \frac{A_{goj}}{2l_g} \right) \left\{ 1 + \eta_{kj} \left(x_j + \frac{A_{goj}}{2l_g} \right)^2 \right\}, \quad (12)$$

para $x_j + \frac{A_{goj}}{2l_g} \leq 0 \quad j = 1, 2.$

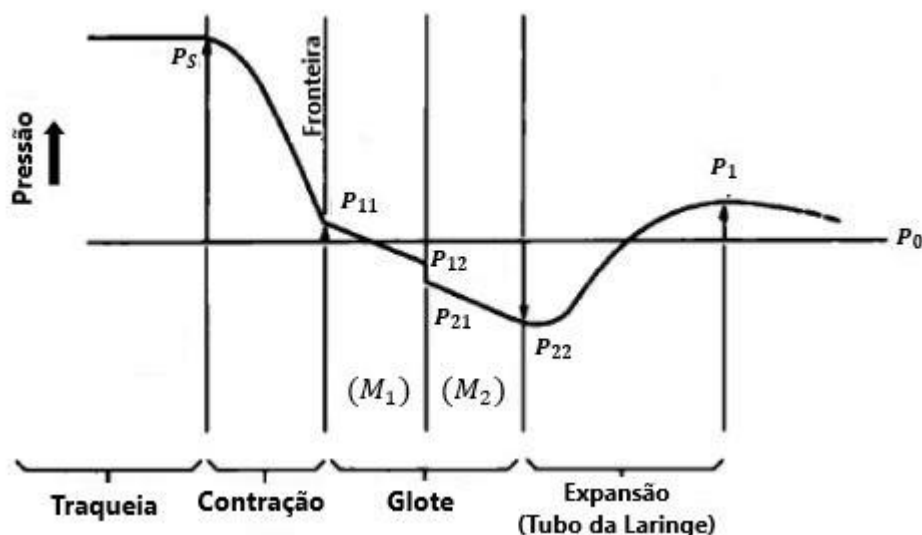
Na Equação (12), h_j é a rigidez linear, η_{kj} é um coeficiente positivo representando a não-linearidade das cordas vocais quando estão em contato, A_{g0j} é a área da região existente entre as cordas vocais quando estão em repouso e l_g é o comprimento das cordas vocais. A força resultante agindo em M_j durante o fechamento será, então, a soma de f_{sj} com f_{hj} .

As forças, às quais as cordas vocais estão submetidas, estão relacionadas com as pressões geradas pelo escoamento do ar ao longo do tubo glotal. Portanto, faz-se necessária a compreensão de como essas pressões se distribuem em cada uma das partes deste sistema e quais fatores influenciam na sua queda, a fim de se buscar e entender a origem daquelas grandezas.

Em seu modelo de duas massas, Ishizaka e Flanagan (p. 1238, 1972) consideram o comportamento do escoamento do ar ao longo do tubo glotal como sendo quase-estático. Essa premissa é feita com base em duas percepções físicas: a pequena dimensão da glote em relação aos comprimentos de onda das frequências relevantes para o tema em questão e a alta velocidade de escoamento do ar pelo sistema de produção de voz.

Segundo Ishizaka e Flanagan (1972) é possível a obtenção das pressões, ao longo do tubo glotal, mediante o uso da equação de Bernoulli para um fluxo unidimensional em cada uma das regiões de interesse. Por meio destes equacionamentos, que serão apresentados em breve no texto, Ishizaka e Flanagan (1972) obtiveram o gráfico de distribuição de pressões pelo tubo glotal, apresentado pela Figura 14.

Figura 14 - Distribuição de Pressão ao Longo do Sistema de Produção de Voz.



Fonte: Adaptado de Ishizaka; Flanagan (1972, p.1241).

A primeira etapa do processo de produção de voz inicia-se nos pulmões (fonte), no qual

a pressão é máxima e representada na Figura 13 por P_s . Esta pressão permanece constante até o momento em que se adentra na zona de contração ou *vena contracta*, região anterior à massa m_1 das cordas vocais, onde ocorre uma queda brusca de pressão, devido à mudança abrupta na área do tubo por onde o escoamento de ar flui.

“A *vena contracta*” (derivada da camada-limite) faz a área de entrada A_{g1} parecer menor do que na verdade é e a queda de pressão maior do que a ditada por uma mudança de área ideal” (ISHIZAKA; FLANAGAN, 1972, p. 1239). Para representar esse fenômeno no equacionamento, estes autores acrescentam um fator de perda de 1,37 à equação de Bernoulli. Portanto, a pressão nesta região (P_b) pode ser verificada pela Equação (13) a seguir.

$$P_b = 1,37 \frac{\rho}{2} \left(\frac{U_g}{A_{g1}} \right)^2. \quad (13)$$

Na etapa seguinte à zona de contração, conforme a Figura 14, encontra-se a região glotal, composta pelas cordas vocais. Para o modelo de duas massas, o escoamento de ar passará primeiramente pela parte inferior da corda vocal, representada pela massa M_1 . A queda de pressão nesta região é determinada pela perda devido à viscosidade do tecido que a compõe. A equação que rege a queda de pressão nesta parte é dada pela Equação (14).

$$P_{11} = 12 \frac{\mu l_g^2 d_1}{A_{g1}^3} U_g, \quad (14)$$

sendo P_{11} a pressão na borda “inicial” da massa M_1 , de uma das cordas vocais, e μ o coeficiente de viscosidade de cisalhamento. Observa-se que o decaimento nesta região é linear, ocasionado por uma resistência ao fluxo do volume de ar.

Em seguida, na região limítrofe entre as massas M_1 e M_2 , apesar da vazão volumétrica U_g permanecer constante, há uma mudança na velocidade das partículas que compõem o ar nesta região. Portanto, haverá uma queda significativa na pressão devido à variação na energia cinética por unidade de volume de fluido. Esta queda é expressa pela Equação (15).

$$P_{12} - P_{21} = \frac{\rho}{2} U_g^2 \left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2} \right). \quad (15)$$

Adentrando-se na região formada pelo limite interno da corda superior, neste caso m_2 ,

o comportamento do fluido nesta região é análogo à análise feita para a massa M_1 , ou seja, a queda de pressão é dada pelas perdas devido a viscosidade do tecido. A Equação (16) descreve tal comportamento, cuja variação é diretamente proporcional à distância d .

$$P_{21} = 12 \frac{\mu l_g^2 d_2}{A_{g2}^3} U_g. \quad (16)$$

A última etapa a ser estudada é a expansão, que está relacionada ao escoamento do ar das cordas vocais até a parte imediatamente exterior a boca. Durante esse intervalo ocorre uma variação positiva abrupta da pressão, já que a tendência da pressão é de se aproximar do valor atmosférico. Para estimar o valor desta recuperação de pressão, Ishizaka e Flanagan se basearam na primeira Lei de Newton, $f = (d/dt)(mv)$, tendo em vista que para pequenas constrições valores experimentais existentes para a recuperação da pressão poderiam distanciar os valores da realidade, por conta da sua dificuldade de medição e seu grau de incerteza. A Equação (17) apresenta como se dá a diferença de pressão para esta última etapa.

$$P_{22} - P_1 = -\frac{\rho}{2} \left(\frac{U_g}{A_{g2}}\right)^2 2 \frac{A_{g2}}{A_1} \left(1 - \frac{A_{g2}}{A_1}\right). \quad (17)$$

Para as equações se aproximarem ao máximo do estado físico real, Ishizaka e Flanagan (1972, p. 1240) consideraram ainda em seus equacionamentos, o efeito da inércia da massa de ar que atravessa o sistema de produção de voz. Portanto, as equações finais que descrevem a distribuição das pressões ao longo do tubo glotal, para cada uma das etapas, podem ser observadas a seguir.

$$P_s - P_{11} = 1,37 \frac{\rho}{2} \left(\frac{U_g}{A_{g1}}\right)^2 + \int_0^{l_g} \frac{\rho}{A_c(x)} dx \frac{dU_g}{dt}, \quad (18)$$

$$P_{11} - P_{12} = 12 \frac{\mu l_g^2 d_1}{A_{g1}^3} U_g + \frac{\rho d_1}{A_{g1}} \frac{dU_g}{dt}, \quad (19)$$

$$P_{12} - P_{21} = \frac{\rho}{2} U_g^2 \left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2}\right), \quad (20)$$

$$P_{21} - P_{22} = 12 \frac{\mu l_g^2 d_2}{A_{g2}^3} U_g + \frac{\rho d_2}{A_{g2}} \frac{dU_g}{dt}, \quad (21)$$

$$P_{22} - P_1 = -\frac{\rho}{2} \left(\frac{U_g}{A_{g2}} \right)^2 2 \frac{A_{g2}}{A_1} \left(1 - \frac{A_{g2}}{A_1} \right). \quad (22)$$

Conforme Sampaio *et al.* (2006), assim como para os sistemas de propagação eletromagnética, os sistemas acústicos também podem ser descritos, por analogia, na forma de circuitos elétricos. Desta maneira, tem-se que a diferença de pressão no sistema acústico representa a tensão elétrica em um circuito e o fluxo de ar, que é o volume de ar deslocado/tempo, equivale à corrente elétrica, quantidade de carga deslocada/tempo.

Fazendo uma analogia com um circuito elétrico, o pulmão comporta-se como uma fonte de tensão elétrica. As demais características para o restante do aparelho fonador, associadas à um circuito RL, são obtidas, por Ishizaka e Flanagan (1972), a partir do equacionamento apresentado anteriormente para as pressões ao longo do tubo glotal. Os autores consideraram os termos, nestas equações, que acompanham o U_g , análogo à corrente que circula pelo circuito, como sendo equivalentes às resistências, se estes forem lineares, e, caso contrário, correspondentes as indutâncias, por conta da presença da não-linearidade na expressão. O equacionamento destes termos citados acima é apresentado pelas equações (23), (24), (25), (26), (27), (28), (29) e (30).

$$R_c = 1,37 \frac{\rho |U_g|}{2 A_{g1}^2}, \quad (23)$$

$$L_c = \int_0^{l_c} \frac{dx}{A_c(x)}, \quad (24)$$

$$R_{v1} = 12 \frac{\mu l_g^2 d_1}{A_{g1}^3}, \quad (25)$$

$$L_{g1} = \frac{\rho d_1}{A_{g1}}, \quad (26)$$

$$R_{12} = \frac{\rho}{2} \left(\frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2} \right) |U_g|, \quad (27)$$

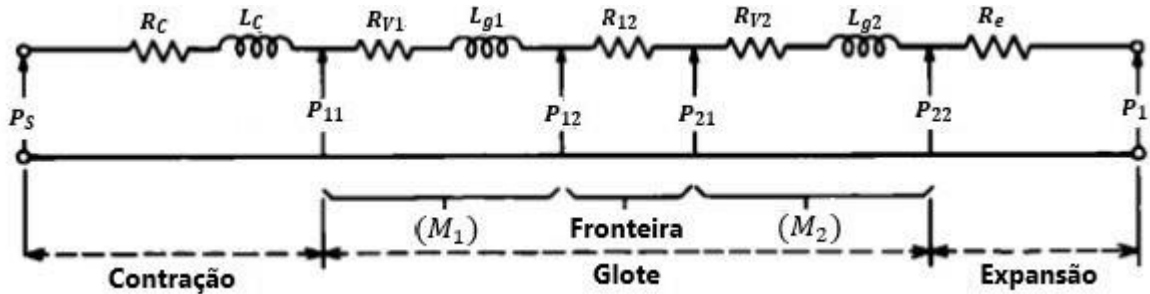
$$R_{v2} = 12 \frac{\mu l_g^2 d_2}{A_{g2}^3}, \quad (28)$$

$$L_{g2} = \frac{\rho d_2}{A_{g2}}, \quad (29)$$

$$R_e = -\frac{\rho}{2} \frac{2}{A_1 A_{g2}} \left(1 - \frac{A_{g2}}{A_1} \right) |U_g|. \quad (30)$$

Tomando como base as equações acima citadas, obtém-se o circuito acústico equivalente apresentado pela Figura 15, representando o escoamento do ar desde o pulmão até a parte externa da boca.

Figura 15 - Circuito Equivalente da Glote.



Fonte: Adaptado de Ishizaka; Flanagan (1972, p.1241).

Nota-se que a Figura 15 estabelece um paralelo com o gráfico da Figura 14, em que foi apresentada a distribuição das pressões ao longo do sistema de produção de voz. Percebe-se que à cada uma das etapas de queda de pressão, correspondente à queda de tensão no circuito, existem resistências e/ou impedâncias responsáveis por essas variações.

3. ESTRUTURA E PREMISSAS PARA A SIMULAÇÃO COMPUTACIONAL DE SINAIS VOCAIS

No presente capítulo serão apresentadas as equações consideradas para a simulação computacional de um sistema vocal. A representação desse sistema é dada por um circuito elétrico, introduzido no Capítulo II. Sendo assim, serão apresentadas as equações que regem este circuito e que, por conseguinte, descrevem o movimento das cordas vocais e o escoamento de ar, os quais já foram descritos na seção precedente.

Ainda neste capítulo, serão mostradas as considerações acerca da transferência destas equações para o ambiente computacional. Como a modelagem matemática, que descreve o processo de produção da voz, fornece um sistema de equações íntegro-diferenciais, as considerações feitas para a resolução das mesmas estão definidas a seguir.

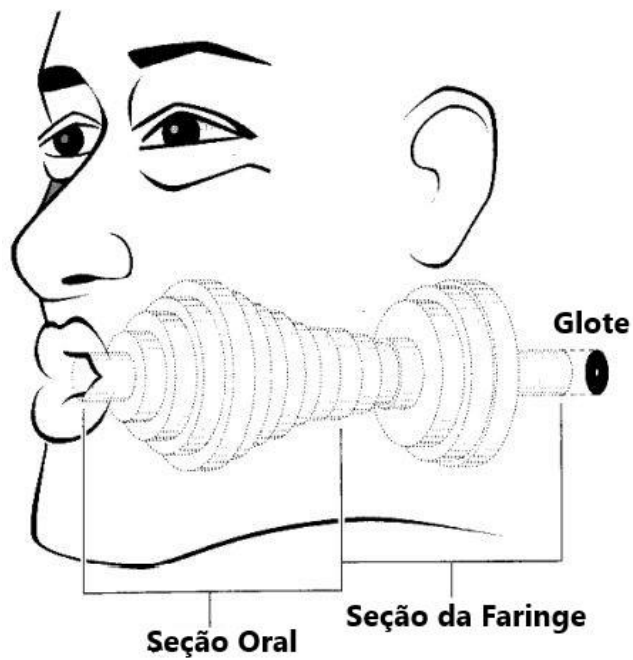
Por fim, será apresentado o fluxograma da rotina de implementação do código que resultará na síntese de sinais de voz inspirada no modelo de duas massas de Ishizaka e Flanagan (1972).

3.1 OBTENÇÃO DAS EQUAÇÕES

Foi acoplado ao circuito que representa a glote (circuito RL) o circuito RLC para o trato vocal. Sendo o trato vocal, conforme Figura 16, um conjunto ‘ n ’ de tubos cilíndricos que têm seus valores fixos, já que pretendeu-se reproduzir o sinal emitido de uma única vogal, fonema /a/. O presente trabalho não entrou no mérito de sons nasalados, devido a sua complexidade, já que para este caso uma modelagem adicional seria necessária para representar a cavidade nasal.

Pode-se observar, por meio da Figura 16, o conjunto de ‘ n ’ tubos cilíndricos que compõem o trato vocal.

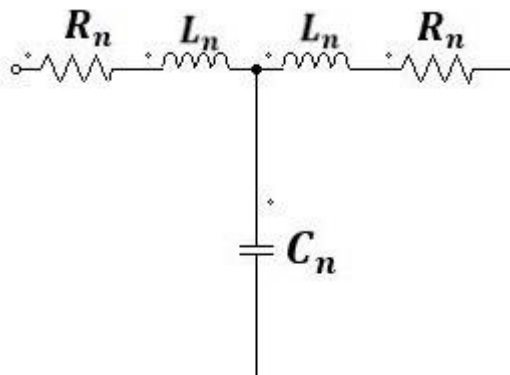
Figura 16 - Aproximação do trato vocal por tubos cilíndricos.



Fonte: Adaptado de SAMPAIO; CATALDO; BRANDÃO (2006).

Fazendo uma analogia com circuitos elétricos, esses tubos podem ser representados como uma linha de transmissão, em que o circuito equivalente para cada um deles é representado por um circuito em 'T', como pode ser visto na Figura 17.

Figura 17 - Circuito T equivalente a um tubo cilíndrico.



Fonte: SAMPAIO; CATALDO, BRANDÃO (2006).

No qual os valores das resistências, indutâncias e capacitâncias são dados por:

$$L_j = \frac{\rho l_j}{2A_j}, \quad (31)$$

$$C_j = \frac{l_j A_j}{\rho c^2}, \quad (32)$$

$$R_j = \frac{s_j}{A_j^2} \sqrt{\rho \mu \omega / 2} \quad j = 1, 2, \dots, n, \quad (33)$$

sendo que A_j e l_j representam a seção transversal e o comprimento do tubo respectivamente. A densidade do ar é ρ , c é a velocidade do som, s_j o perímetro da circunferência do tubo da j -ésima seção, μ é a viscosidade dinâmica do ar, e ω é a frequência em radianos por segundo.

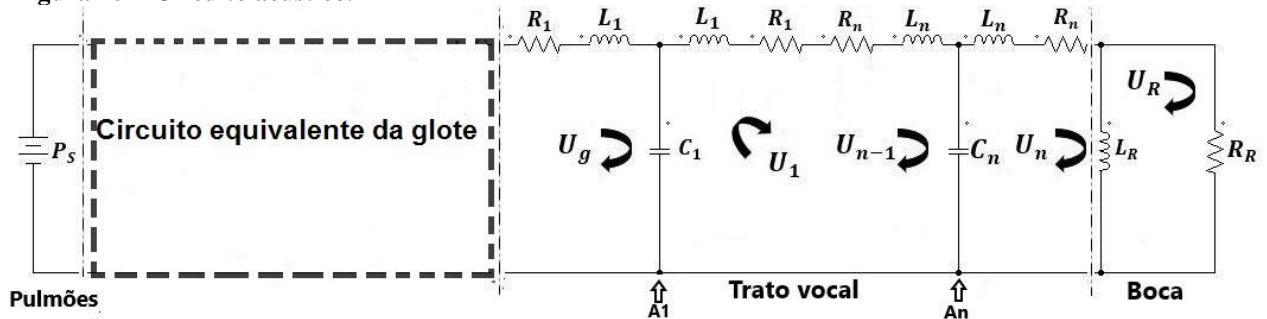
Já a irradiação do som, que ocorre pela boca, é representada como uma impedância de radiação. Pode-se fazer um comparativo dessa impedância com a modelagem de uma antena em sistemas elétricos, a qual é formada por uma resistência R_r e por uma indutância L_r cujas equações estão a seguir. Sendo assim, a diferença de pressão entre os terminais dessa impedância é o que corresponde a voz sintetizada (SAMPAIO *et al.*, 2006).

$$R_r = \left(\frac{128 \rho c}{9 \pi^2 A_n} \right). \quad (34)$$

$$L_r = \frac{8 \rho}{3 \pi} \sqrt{\pi A_n}. \quad (35)$$

Feitas as devidas considerações, e tomando como base as relações entre as diferenças de pressões na glote explicadas no item 2.4, o circuito acústico completo, que representa o sistema vocálico, desde o pulmão até a parte externa a boca, é apresentado na Figura 18.

Figura 18 - Circuito acústico.



Fonte: Adaptado de Ishizaka; Flanagan (1972).

A partir da 2ª Lei de Kirchhoff, lei das tensões, obtém-se as equações que regem o circuito acústico acima. Essas equações estão expressas a seguir por meio das equações (36), (37), (38), (39), (40) e (41), em que a malha g corresponde à seção das cordas vocais, as malhas 1, 2, 3 e 4 são referentes às seções do trato vocal e a malha 5 corresponde à irradiação da boca, supondo um trato

vocal formado por 4 tubos cilíndricos, mesma premissa estabelecida por Ishizaka e Flanagan (1972) em seu artigo.

$$(Malha g) \quad (R_{k1} + R_{k2})|U_g|U_g + (R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\frac{dU_g}{dt} + L_1\frac{dU_g}{dt} + R_1U_g + \frac{1}{c_1}\int_0^t(U_g - U_1)dt - P_s = 0, \quad (36)$$

$$(Malha 1) \quad (L_1 + L_2)\frac{dU_1}{dt} + (R_1 + R_2)U_1 + \frac{1}{c_2}\int_0^t(U_1 - U_2)dt + \frac{1}{c_1}\int_0^t(U_1 - U_g)dt = 0, \quad (37)$$

$$(Malha 2) \quad (L_2 + L_3)\frac{dU_2}{dt} + (R_2 + R_3)U_2 + \frac{1}{c_3}\int_0^t(U_2 - U_3)dt + \frac{1}{c_2}\int_0^t(U_2 - U_1)dt = 0, \quad (38)$$

$$(Malha 3) \quad (L_3 + L_4)\frac{dU_3}{dt} + (R_3 + R_4)U_3 + \frac{1}{c_4}\int_0^t(U_3 - U_L)dt + \frac{1}{c_3}\int_0^t(U_3 - U_2)dt = 0, \quad (39)$$

$$(Malha 4) \quad (L_4 + L_R)\frac{dU_L}{dt} + R_4U_L - L_R\frac{d(U_R)}{dt} + \frac{1}{c_4}\int_0^t(U_L - U_3)dt = 0, \quad (40)$$

$$(Malha 5) \quad L_R\frac{d(U_R - U_L)}{dt} + R_R U_{boca} = 0. \quad (41)$$

As massas oscilantes, M_1 e M_2 , são conduzidas por pressões médias, que são representadas pelas seguintes expressões:

$$P_{M1} = \frac{1}{2}(P_{11} + P_{12}) = P_s - 1,37\frac{\rho(U_g)^2}{2A_{g1}} - \frac{1}{2}\left(R_{v1}U_g + L_{g1}\frac{dU_g}{dt}\right). \quad (42)$$

$$P_{M2} = \frac{1}{2}(P_{21} + P_{22}) = P_{M1} - \frac{1}{2}\left\{(R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\frac{dU_g}{dt}\right\} - \frac{\rho}{2}U_g^2\left(\frac{1}{A_{g2}^3} - \frac{1}{A_{g1}^3}\right). \quad (43)$$

Após obter as pressões médias, pode-se concluir em qual condição as cordas vocais estão operando. As forças F_1 e F_2 agem sobre as massas M_1 e M_2 , respectivamente, através das faces correspondentes l_{gd1} e l_{gd2} , gerando os deslocamentos x_1 e x_2 . Estes deslocamentos são apresentados pelo Quadro 2 para cada uma destas condições.

Quadro 2 - Deslocamentos x_1 e x_2 .

x_1	x_2	$F_1/l_g d_1$	$F_2/l_g d_2$
$x_1 > x_{1min}$	$x_2 > x_{2min}$	P_{M_1}	P_{M_2}
$x_1 \leq x_{1min}$	$x_2 > x_{2min}$	P_s	0
$x_1 > x_{1min}$	$x_2 \leq x_{2min}$	P_s	P_s
$x_1 \leq x_{1min}$	$x_2 \leq x_{2min}$	P_s	0

Fonte: Ishizaka; Flanagan (1972, p.1244).

No Quadro 2, $x_{1min} = -(A_{g01}/2l_g)$ e $x_{2min} = -(A_{g02}/2l_g)$, e A_{g0} representa a área glotal em condição de repouso. Feitas essas considerações pode-se obter o valor de F_1 e F_2 e, por conseguinte determinar o valor de x_1 e x_2 através das equações (5) e (6). Essas equações são acopladas às equações que descrevem o fluxo, dado que x_1 e x_2 determinam as áreas glotais A_{g1} e A_{g2} .

3.2 MÉTODOS NUMÉRICOS

A modelagem matemática, que descreve o processo de produção da voz baseada no modelo de duas massas apresentada no capítulo precedente, é formada por um conjunto de equações íntegro-diferenciais. Como este sistema de equações apresenta não-linearidades, portanto, não admite soluções analíticas, assim métodos numéricos foram empregados para as suas resoluções.

O método numérico utilizado foi o método de “Euler Regressivo”, no qual as diferenciações, de primeiro e de segunda ordem e as integrações, no domínio do tempo, podem ser descritas respectivamente da seguinte maneira:

$$\frac{d}{dt} x(nT) \approx \frac{x[n]-x[n-1]}{T}, \quad (44)$$

$$\frac{d^2}{dt^2} x(nT) \approx \frac{1}{T} \left(\frac{x[n]-x[n-1]}{T} - \frac{x[n-1]-x[n-2]}{T} \right) = \frac{x[n]-2x[n-1]+x[n-2]}{T^2}, \quad (45)$$

$$\int_0^{nT} x(\tau) d\tau \approx T \sum_{i=0}^n x[i]. \quad (46)$$

Desta forma, aproxima-se o sistema de equações diferenciais por um sistema de equações de diferenças e para as equações integrais a aproximação ocorre por um somatório de

valores. Sendo T o período de amostragem utilizado, $x[n]$ é a amostra do sinal $x(t)$ no instante $t = nT$.

Após o emprego do método de “Euler Regressivo”, a próxima etapa foi a resolução dos conjuntos de sistemas de equações lineares por meio da simulação computacional. Uma das formas de se obter o vetor de soluções, para cada instante no intervalo de tempo considerado, é dispor estas equações na forma matricial $Ax = B$ e resolvê-las. Como exemplo desta estratégia, utilizou-se as equações (47) e (49) da dinâmica das cordas vocais. Portanto, colocando-as na forma matricial citada anteriormente e aplicando o método de Euler para a eliminação das derivadas, obtém-se a seguinte igualdade:

$$\begin{bmatrix} \frac{M_1}{T^2} + \frac{r_1}{T} & -k_c \\ -k_c & \frac{M_2}{T^2} + \frac{r_2}{T} \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \end{bmatrix} = \begin{bmatrix} F_1 + \frac{M_1}{T^2} (2x_1[n-1] - x_1[n-2]) + \frac{r_1}{T} (x_1[n-1]) - s_1 - k_c(x_2[n-1]) \\ F_2 + \frac{M_2}{T^2} (2x_2[n-1] - x_2[n-2]) + \frac{r_2}{T} (x_1[n-1]) - s_2 - k_c(x_1[n-1]) \end{bmatrix}. \quad (47)$$

Observa-se pela primeira matriz desta equação que o sistema é acoplado, ou seja, o cálculo, neste caso, de x_1 depende do valor de x_2 atual e vice-versa. O MATLAB é um software capaz de realizar esse cálculo, contudo esse tipo de operação, do ponto de vista computacional, tem um alto custo. Uma alternativa para resolver este empecilho estaria no uso de algum artifício matemático capaz de desacoplar o sistema de equações. A saída mais simples encontrada foi a introdução de atrasos em termos presentes nas equações que pudessem prejudicar o desacoplamento, tornando a matriz A uma matriz diagonal.

Portanto, a fim de se eliminar o termo $-k_c$ da primeira linha da matriz A , considerou-se a seguinte aproximação para o termo x_2 :

$$x_2[n] \approx x_2[n-1]. \quad (48)$$

Desta forma, o cálculo de x_1 é dependente exclusivamente de termos calculados anteriormente, o que facilitou a resolução de tais equações. O sistema de equações neste formato é apresentado pela Equação (49).

$$\begin{bmatrix} \frac{M_1}{T^2} + \frac{r_1}{T} & 0 \\ 0 & \frac{M_2}{T^2} + \frac{r_2}{T} \end{bmatrix} \begin{bmatrix} x_1[n] \\ x_2[n] \end{bmatrix} = \begin{bmatrix} F_1 + \frac{M_1}{T^2} (2x_1[n-1] - x_1[n-2]) + \frac{r_1}{T} (x_1[n-1]) - s_1 + k_c(x_2[n-1] - x_1[n-1]) \\ F_2 + \frac{M_2}{T^2} (2x_2[n-1] - x_2[n-2]) + \frac{r_2}{T} (x_1[n-1]) - s_2 + k_c(x_1[n-1] - x_2[n-1]) \end{bmatrix} \quad (49)$$

A aproximação feita pela Equação (48) também foi empregada, em alguns casos, com

o objetivo de se eliminar termos cúbicos que surgiram em pontos específicos dos equacionamentos.

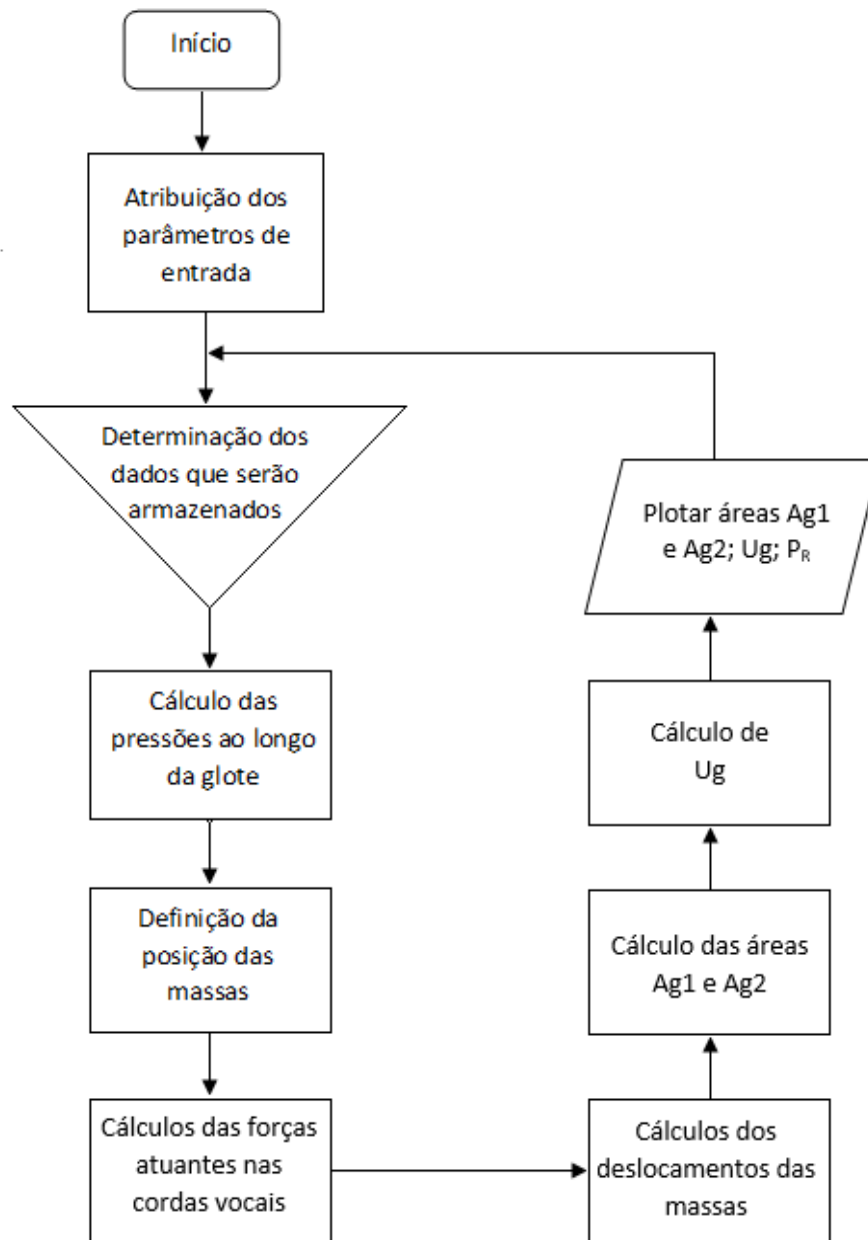
3.3 SIMULAÇÃO COMPUTACIONAL

O algoritmo implementado foi realizado em MATLAB para a sintetização de sinais de voz vocálicos. De forma sintética, primeiramente foram inseridos os dados de inicialização do modelo da laringe. Valores estes como o período de amostragem, pressão subglotal, perfil da formação neutra, massa, coeficiente de elasticidade, entre outros. Posteriormente foram determinados os dados a serem armazenados matricialmente.

Em seguida, os cálculos referentes às equações que descrevem o movimento das cordas vocais, foram resolvidos juntamente com as equações que descrevem o escoamento de ar também presente nesse processo.

As resoluções destas equações resultaram na vazão volumétrica U_g , para cada instante, e através disso foi possível obter os sinais de voz sintetizados. A análise desses dados se deu pelos sinais plotados de U_g , as áreas da glote A_{g1} e A_{g2} e a pressão irradiada, pressão na parte externa da boca, que será representada por P_R . O fluxograma que descreve de forma visual a sequência lógica do código pode ser observado na Figura 19.

Figura 19 - Fluxograma para o código de síntese de sinais vocálicos.



Fonte: Autoria própria.

4. RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados os resultados obtidos com a simulação desenvolvida para a síntese de sinais de voz, baseada no modelo matemático de duas massas, proposto por Ishizaka e Flanagan em 1972. Além da apresentação dos resultados obtidos por meio das simulações, será realizado um estudo sobre os parâmetros que compõem as equações do modelo utilizado.

Após realizadas as considerações para a solução das equações, que descrevem o movimento das cordas vocais atreladas ao fluxo que por elas escoam, estas foram transcritas para o software MATLAB. Essas considerações, que envolveram soluções numéricas, foram explicadas previamente no capítulo II. Desta forma, foi possível obter o sinal do som produzido por esse sistema para a vogal /a/, juntamente com gráficos para visualizar esses sinais.

A fim de se estudar aspectos relacionados a vibração das cordas vocais e compreender como os parâmetros das equações do modelo proposto refletem no formato final dos sinais de voz, realizou-se um estudo sobre estes parâmetros. Este estudo foi feito por meio da variação dos valores dos parâmetros e posterior análise do seu efeito sobre as curvas resultantes.

Inicialmente, para obtenção do sinal voz correspondente ao fonema desejado, foram usados valores padrões para os parâmetros, discriminados no Quadro 3. Estes parâmetros foram baseados nos dados utilizados por Ishizaka e Flanagan (1972) em suas simulações. É importante ressaltar que estes autores adotaram em seu artigo valores fisiologicamente realísticos, ou seja, valores que caracterizam de fato cordas vocais reais.

Quadro 3 - Parâmetros para a síntese de voz.

Parâmetro	Valor	Unidade
lg	1,40	cm
A_{g01}	0,05	cm ²
A_{g02}	0,05	cm ²
d₁	0,20	cm
d₂	0,10	cm
M₁	0,10	g
M₂	0,05	g
k₁	80000	dyn/cm
k₂	64000	dyn/cm
k_c	80000	dyn/cm
P_s	8000	dyn/ cm ²

Fonte: Autoria própria.

Após a síntese do sinal de voz correspondente aos parâmetros padrões, iniciou-se o estudo do comportamento de cada um destes parâmetros. A variação deles não foi realizada de forma simultânea, ou seja, variou-se apenas o parâmetro de interesse e para os demais foram mantidos seus valores padrões fixos.

Para a obtenção de resultados significativos, foram adotados aproximadamente 30 valores diferentes para cada parâmetro. Das simulações executadas, serão apresentadas apenas as mais significativas para justificar a influência de cada parâmetro na síntese de voz.

4.1 SÍNTESE DA VOGAL /A/

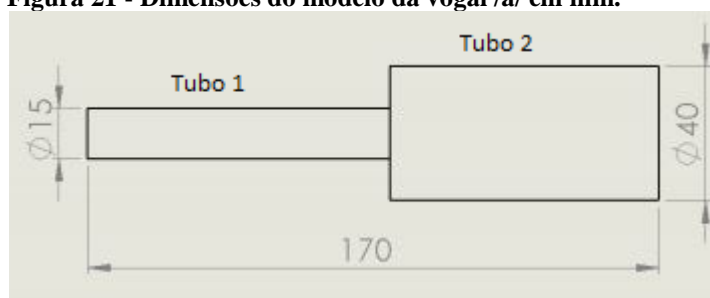
Para a obtenção do sinal de voz considerou-se que o conjunto de 'n' tubos cilíndricos, que constituem o trato vocal, seria formado por dois cilindros. Esta consideração foi baseada no modelo tridimensional utilizado por SILVA Jr. (2015). Desta maneira, como o código de simulação é composto por quatro tubos, foram atribuídos a estes mesmos valores dois a dois.

É importante ressaltar que as dimensões dos cilindros variam de acordo com a vogal que se deseja obter. No caso da vogal /a/, os tubos foram modelados como apresentado na Figura 20, e suas dimensões estão presentes na Figura 21.

Figura 20 - Modelo tridimensional da vogal/a/.



Fonte: SILVA Jr. (2015).

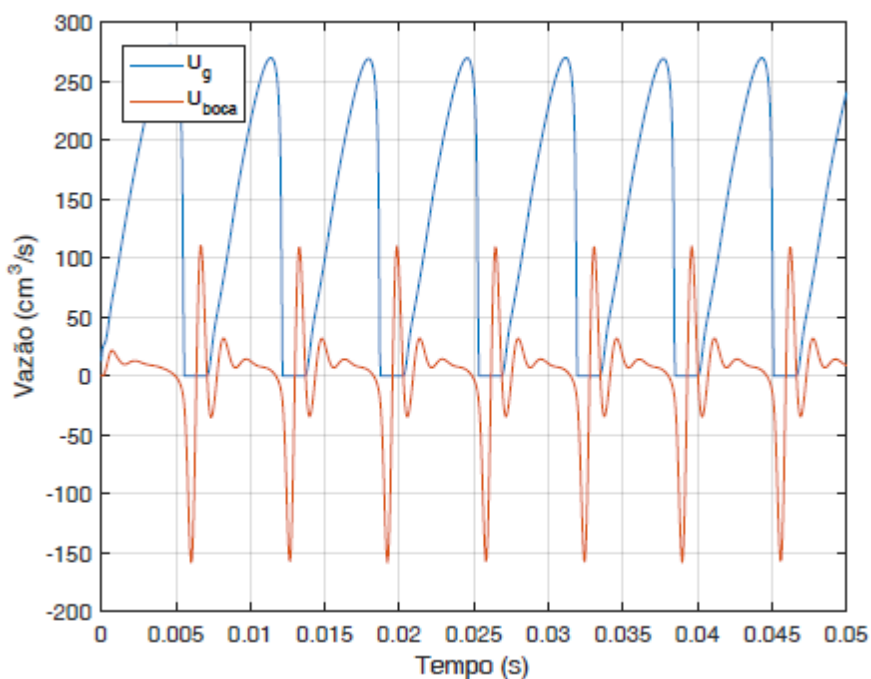
Figura 21 - Dimensões do modelo da vogal /a/ em mm.

Fonte: SILVA Jr. (2015).

Considerando-se o tubo (1) e o tubo (2), conforme Figura 21, com comprimento de 80 mm e 90 mm respectivamente, obteve-se as curvas, que descrevem as condições típicas da glote nos primeiros 0,05 s de vocalização, por meio das simulações. Curvas essas que correspondem a vazão do ar no tubo glotal, as áreas glotais para as massas M_1 e M_2 , e aos deslocamentos x_1 e x_2 . Essas curvas estão contidas nas figuras 22, 23 e 24, respectivamente e serão detalhadas nos parágrafos a seguir.

A representação gráfica da vazão da voz obtida na saída do sistema, apresentada pela Figura 22, corresponde a curva U_{boca} , que é equivalente à U_r do circuito acústico. Essa curva é a porção do fluxo U_g , iniciado na glote, que passa pelo trato vocal e é irradiada pela boca. Esse sinal é fruto do efeito de um componente RL, conforme Figura 18, e corresponde a vazão responsável pelo som característico para a vogal sintetizada. Observa-se que a onda do fluxo glotal irradiado pela boca é caracterizada pela sua assimetria e um declive acentuado.

Figura 22 - Vazão no tubo glotal e vazão irradiada pela boca para a condição padrão adotada.

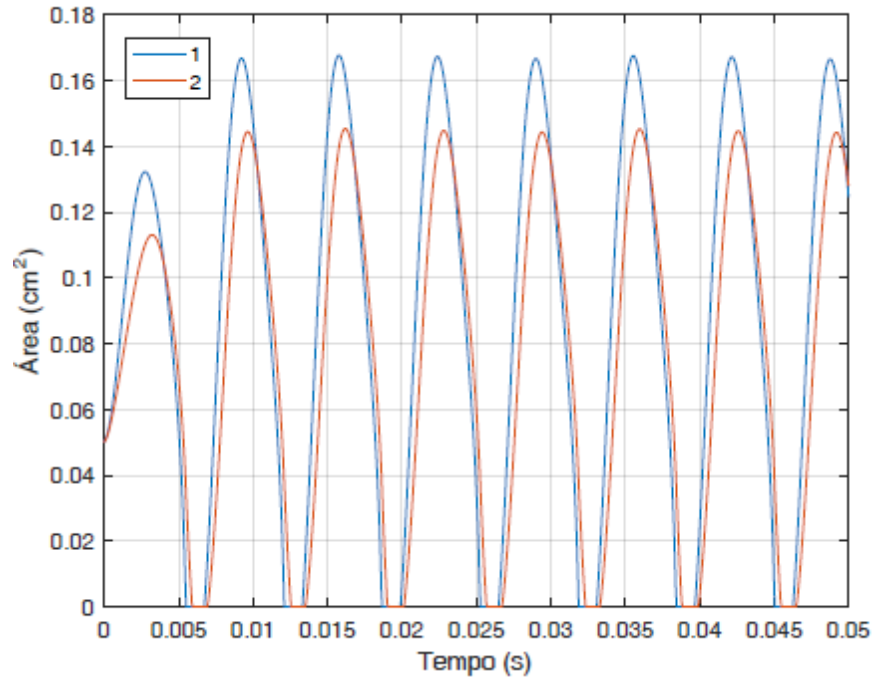


Fonte: Autoria própria.

O sinal que de fato representa a voz falada ou sintetizada, como explicado no capítulo III, é dado pela grandeza P_r . Os gráficos dela, no entanto, não foram incluídos no presente trabalho pois a facilidade de comparar sinais de mesma natureza, neste caso U_g e U_{boca} , é muito maior. Essa escolha, no entanto, não afeta o processo analítico, já que a relação entre U_{boca} e P_r é proporcional, dada por R_r . Portanto, a distinção entre eles ocorre somente em função da amplitude de onda.

Na Figura 23, observam-se as curvas resultantes das áreas glotais A_{g1} e A_{g2} . Os valores negativos dessas ondas indicam o momento de fechamento glótico, ou seja, quando as cordas vocais entram em contato uma com a outra. Sendo assim, para a vogal /a/ o tempo de abertura da glote para o período total, após atingir o regime permanente, é cerca de 0,005 s, como pode ser visto na figura abaixo.

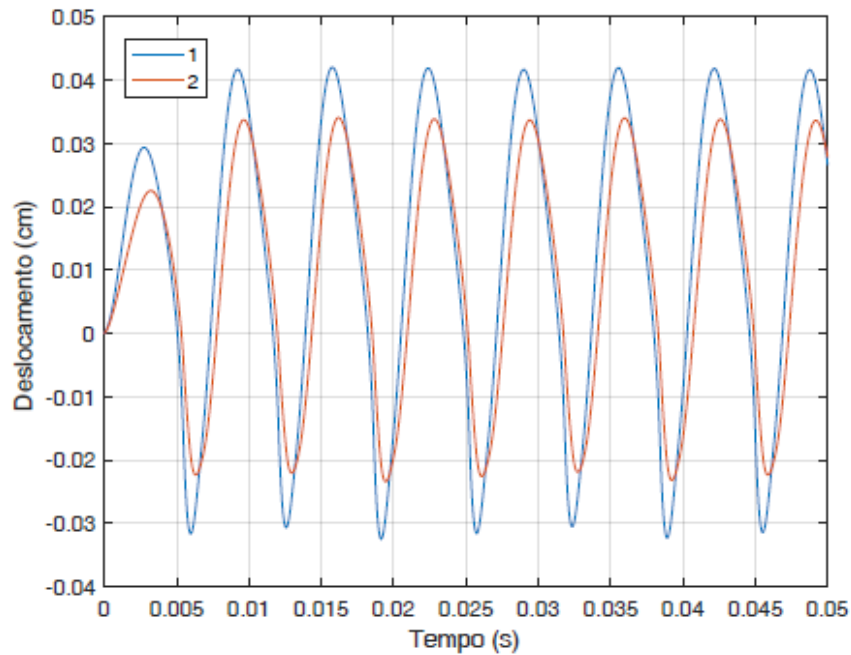
Figura 23 - Área glotal 1 e área glotal 2 para a condição padrão adotada.



Fonte: Autoria própria

Os deslocamentos aos quais as massas são submetidas, x_1 e x_2 , são apresentados na Figura 24. Segundo Ishizaka e Flanagan (1972, p.1252) as cordas vocais são maciças e, geralmente, submetidas a uma frequência acima da frequência natural, por isso o deslocamento mecânico delas não reflete o detalhe da interação acústica que o fluxo glotal exige.

Figura 24 - Deslocamento x_1 e x_2 para a condição padrão adotada.



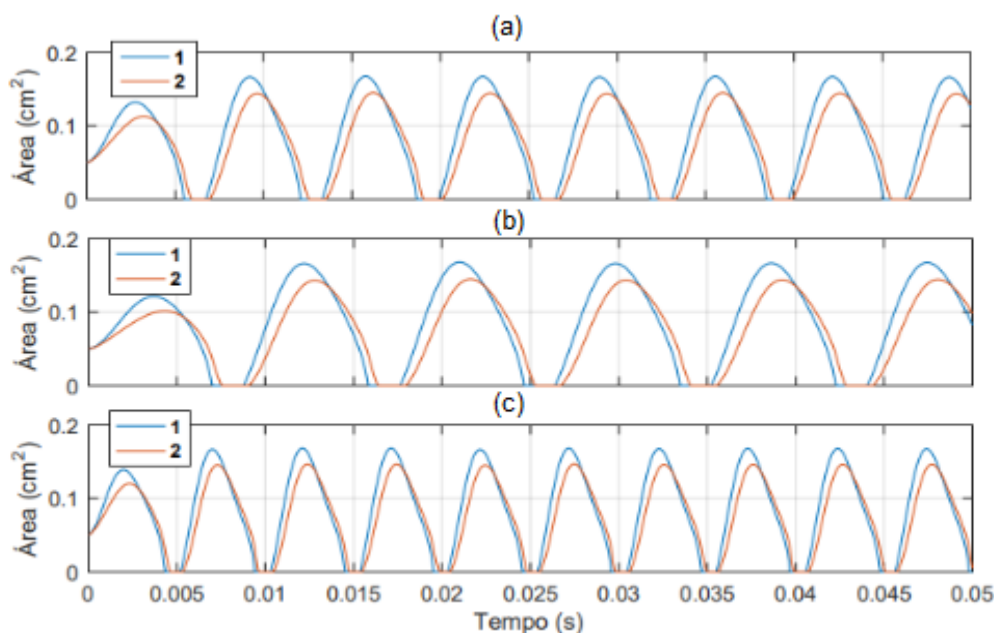
Fonte: Autoria própria.

4.2 ESTUDO DA VARIAÇÃO DOS PARÂMETROS

4.2.1 Efeito da variação das massas

Os resultados da simulação, quanto às massas M_1 e M_2 , mostram que o aumento destas causa uma diminuição na frequência do sinal de voz. É possível visualizar essa alteração através de uma análise feita nas curvas das áreas glotais para diferentes valores das massas. Modificando proporcionalmente os valores das massas M_1 e M_2 verificou-se que quanto maior o valor das massas maior a inércia a qual elas estão sujeitas, sendo assim a frequência de vibração das cordas vocais é mais lenta. Estas considerações são análogas para a diminuição das massas M_1 e M_2 , neste caso as cordas vibram mais rapidamente gerando um sinal de voz com frequências mais altas, essas comparações podem ser vistas na Figura 25.

Figura 25 - Simulação do efeito da variação das massas.



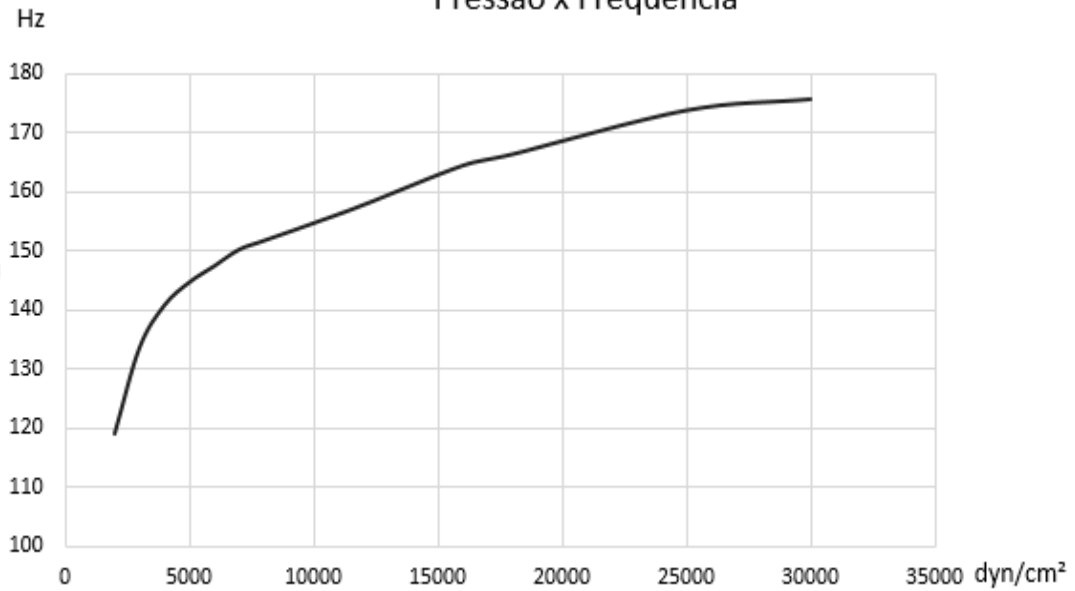
Fonte: Autoria própria.

No caso (a), da Figura 25, apresentam-se as curvas das áreas glotais para as massas $M_1 = 0,10$ g e $M_2 = 0,05$ g, que correspondem às condições normais da glote, ou seja, valores para uma corda vocal saudável. Quando as massas foram variadas para duas vezes seus respectivos valores, portanto $M_1 = 0,2$ g e $M_2 = 0,1$ g, a frequência de abertura e fechamento da glote diminuiu, como pode ser visto no caso (b). Já no caso (c), no qual $M_1 = 0,05$ g e $M_2 = 0,025$ g, a frequência de fechamento da glote aumentou com a diminuição das massas. Pode-se verificar que as alterações nas massas não geram mudanças significativas nas amplitudes das ondas, portanto as cordas vocais nos três casos geram as mesmas áreas glotais.

4.2.2 Efeito variação da pressão subglotal

Outro aspecto importante na produção da voz é a influência da pressão subglotal (P_s) sobre a frequência da vibração da corda vocal e sobre a amplitude do sinal de voz. A Figura 26 relaciona a variação da pressão com a frequência da voz sintetizada. A variação da pressão foi feita iniciando em 1000 dyn/cm² até 30000 dyn/cm², com um passo de 2000 dyn/cm². Através da variação da pressão subglotal percebeu-se que a pressão mínima para a produção da vogal /a/ é a cerca de 2000 dyn/cm².

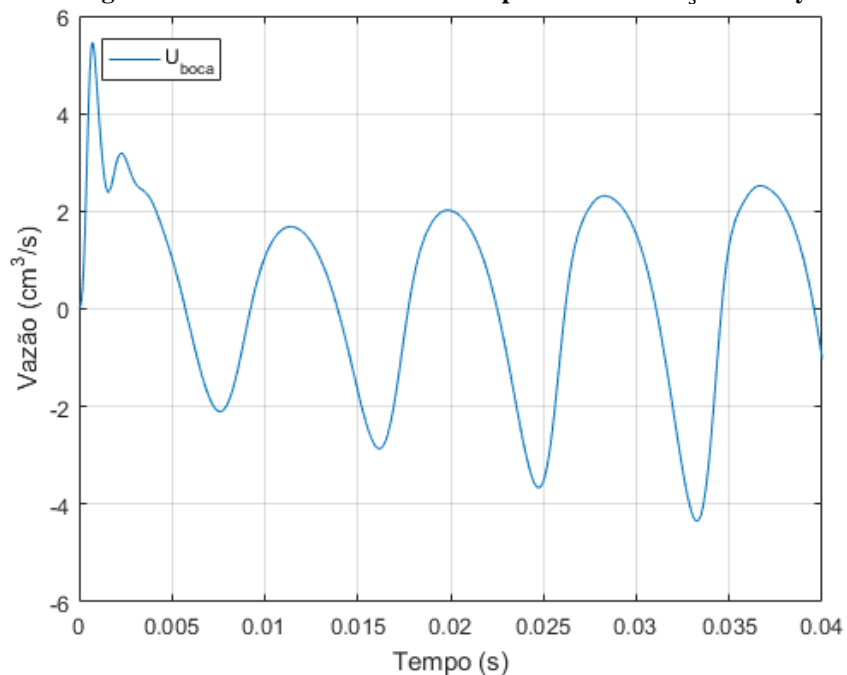
Figura 26 - Variação da frequência em relação a variação da pressão subglotal.
Pressão x Frequência



Fonte: Autoria própria.

Como pode ser observado na Figura 27, para valores abaixo de 2000 dyn/cm² a curva que representa a síntese da voz já não corresponde à configuração da onda inicialmente simulada para uma pressão subglotal de 8000 dyn/cm². O que significa que para valores abaixo de 2000 dyn/cm² as cordas vocais não se fecham.

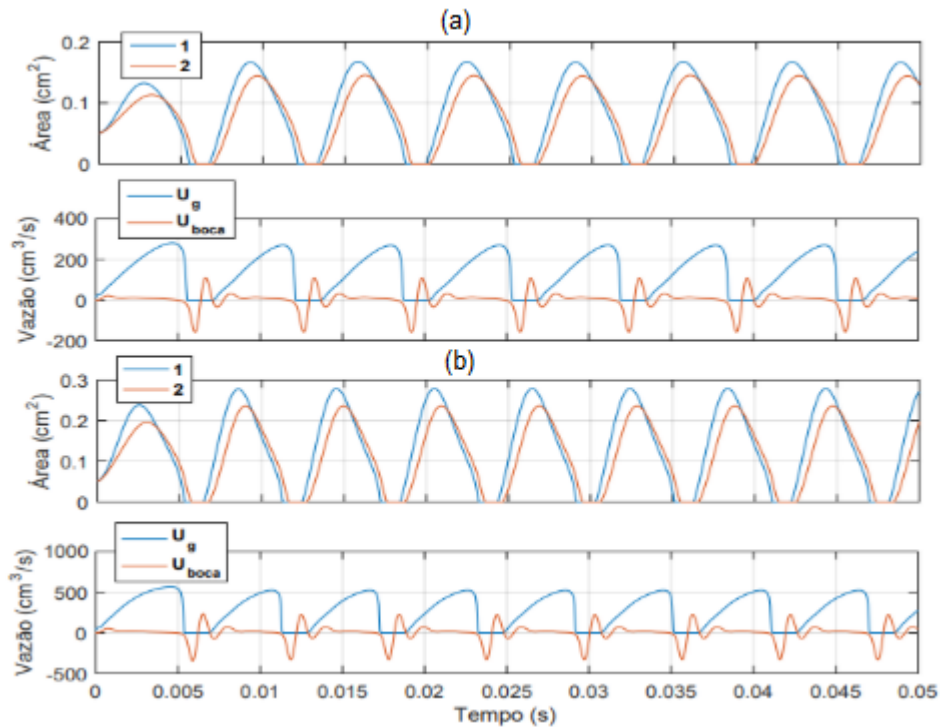
Figura 27 - Vazão do fluxo irradiado pela boca com $P_s = 2000$ dyn/cm².



Fonte: Autoria própria.

Constatou-se também que um aumento na pressão subglotal induz um aumento no fluxo glotal e da amplitude da área glotal como pode ser visto na Figura 28. Para o caso das condições normais (Figura 28 a), no qual $P_s = 8000 \text{ dyn/cm}^2$, o gráfico da vazão do fluxo glotal apresenta um pico de cerca de $270 \text{ cm}^3/\text{s}$, e uma área glotal de cerca de $0,15 \text{ cm}^2$. Já quando ocorre o aumento desta pressão para 20000 dyn/cm^2 , caso (b), o fluxo glotal chega a $530 \text{ cm}^3/\text{s}$ e a área glotal atinge cerca de $0,22 \text{ cm}^2$.

Figura 28 - Simulação do efeito da variação da pressão subglotal.



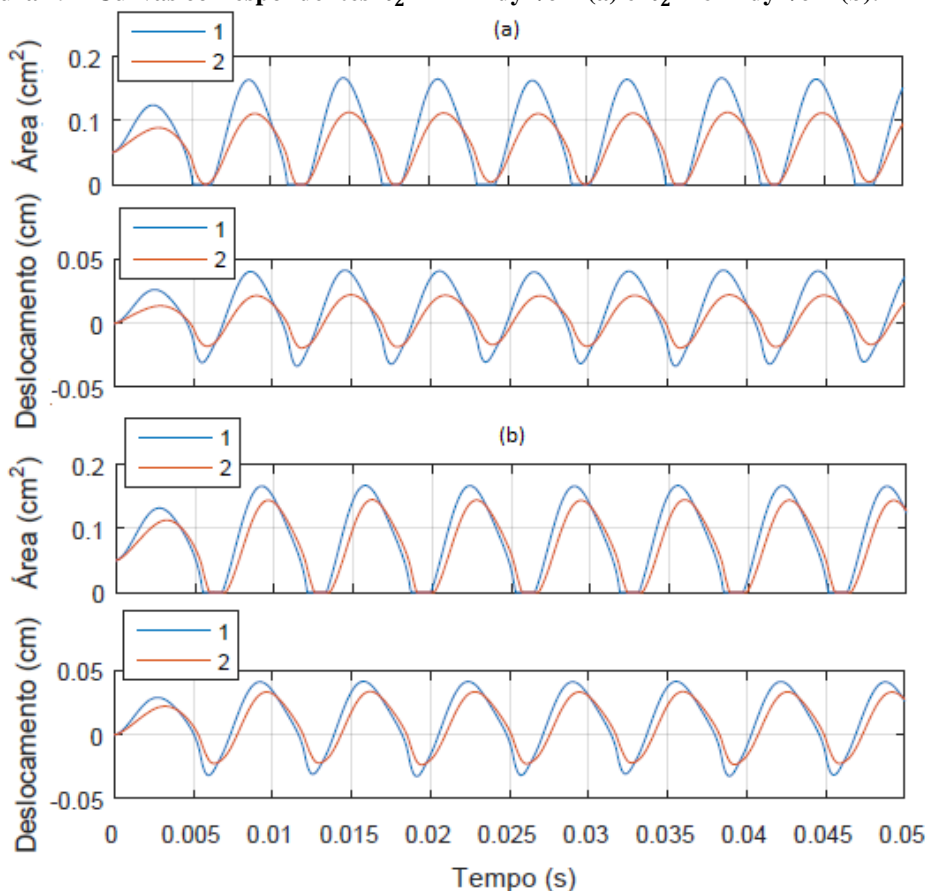
Fonte: Autoria própria.

4.2.3 Efeito da variação dos coeficientes de elasticidade

O coeficiente de elasticidade, tanto de cada uma das massas M_1 e M_2 (simbolizados por k_1 e k_2) quanto entre as massas (chamado de k_c), representa a resposta das cordas vocais às deformações geradas quando submetidas a forças externas. Portanto, tendo como base a Lei de Hooke ($F = kx$), quanto maior o coeficiente de elasticidade de um corpo, mantendo-se a distância constante, maior deve ser a força exercida sobre este para deslocá-lo ou deformá-lo em x cm. Tendo este fenômeno físico em mente, foram realizadas simulações variando-se os parâmetros k_1 , k_2 e k_c a fim de se entender os seus efeitos sobre as formas de onda de interesse.

A Figura 29, caso (a), apresenta uma situação na qual o valor de k_2 foi variado para 124 kdyn/cm, em relação ao valor de referência (64 kdyn/cm) apresentado como caso (b).

Figura 29 - Curvas correspondentes $k_2 = 124$ kdyn/cm (a) e $k_2 = 64$ kdyn/cm (b).

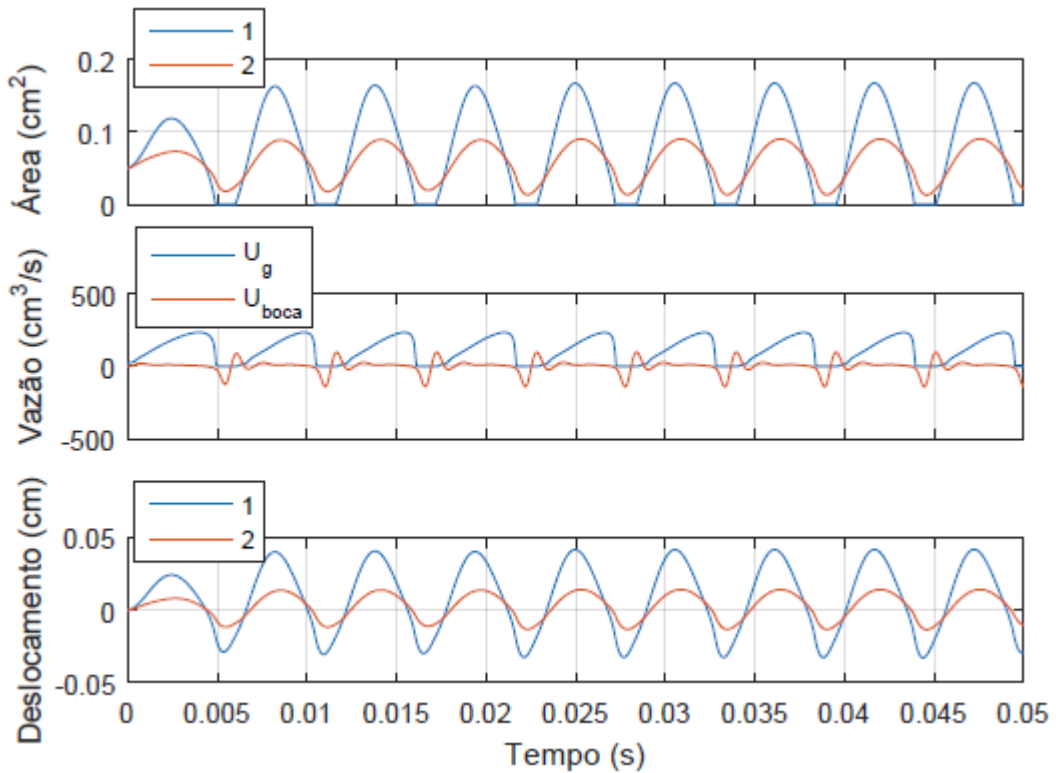


Fonte: Autoria própria.

Nota-se que o aumento de k_2 limitou o movimento da massa M_2 , tendo em vista que a amplitude da curva 2 nestes dois gráficos diminuiu em relação ao caso (b). Esta constatação está atrelada ao fato da força exercida sobre as massas ter permanecido a mesma em ambos os casos (tanto para 64 dyn/cm quanto para 124 dyn/cm), porém o aumento do coeficiente de elasticidade fez com que a variação x (na horizontal) fosse menor. Ainda em relação ao primeiro gráfico (área) do caso (a), observa-se que as massas M_2 das cordas vocais estão na iminência de deixar de colidir uma contra a outra, já que a curva 2 toca levemente o ponto zero, que é a linha de simetria entre as cordas vocais.

A título de comparação, para um caso extremo (valor bem acima dos valores típicos de uma corda vocal saudável) no qual k_2 atinja 224 kdyn/cm, torna-se evidente o efeito do coeficiente de elasticidade no movimento da massa M_2 . Este resultado pode ser visto por meio do gráfico da área da Figura 30.

Figura 30 - Curvas correspondentes à simulação na qual $k_2 = 224$ kdyn/cm.

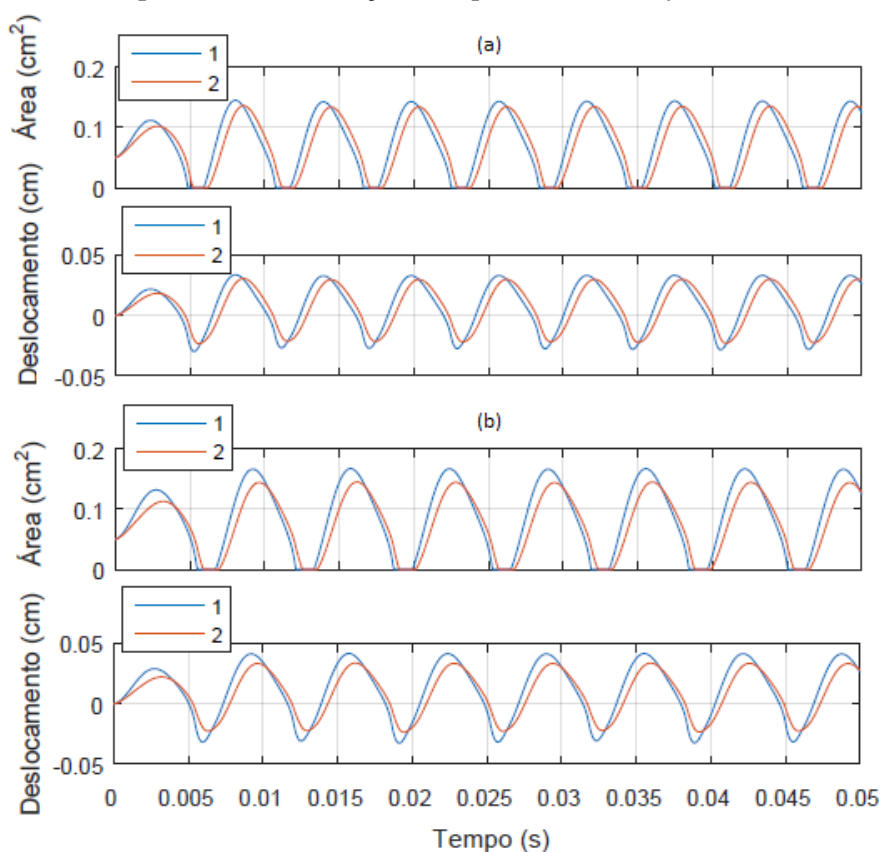


Fonte: Autoria própria.

Neste caso, o deslocamento de M_2 é muito menor em relação ao coeficiente simulado anteriormente, Figura 29 (a). Outro resultado interessante está no fato de que as massas M_2 não chegam a colidir nesta configuração, portanto esta porção da corda vocal ainda irá se movimentar, contudo não fechará em nenhum instante durante o escoamento de ar através dela.

Após as alterações feitas para k_2 , este foi novamente fixado em 64 kdyn/cm, como apresentado na Figura 29 (b), e k_1 foi submetido à diversas variações acompanhadas das gerações de novos sinais de onda. Ao aumentar em 40 kdyn/cm o valor de k_1 , ou seja, $k_1 = 120$ kdyn/cm, as formas de onda já demonstram o impacto desse coeficiente no resultado final. O resultado desta configuração é apresentado na Figura 31 (a).

Figura 31 - Curvas correspondentes às simulações nas quais $k_1 = 120$ kdyn/cm (a) e $k_1 = 80$ kdyn/cm (b).

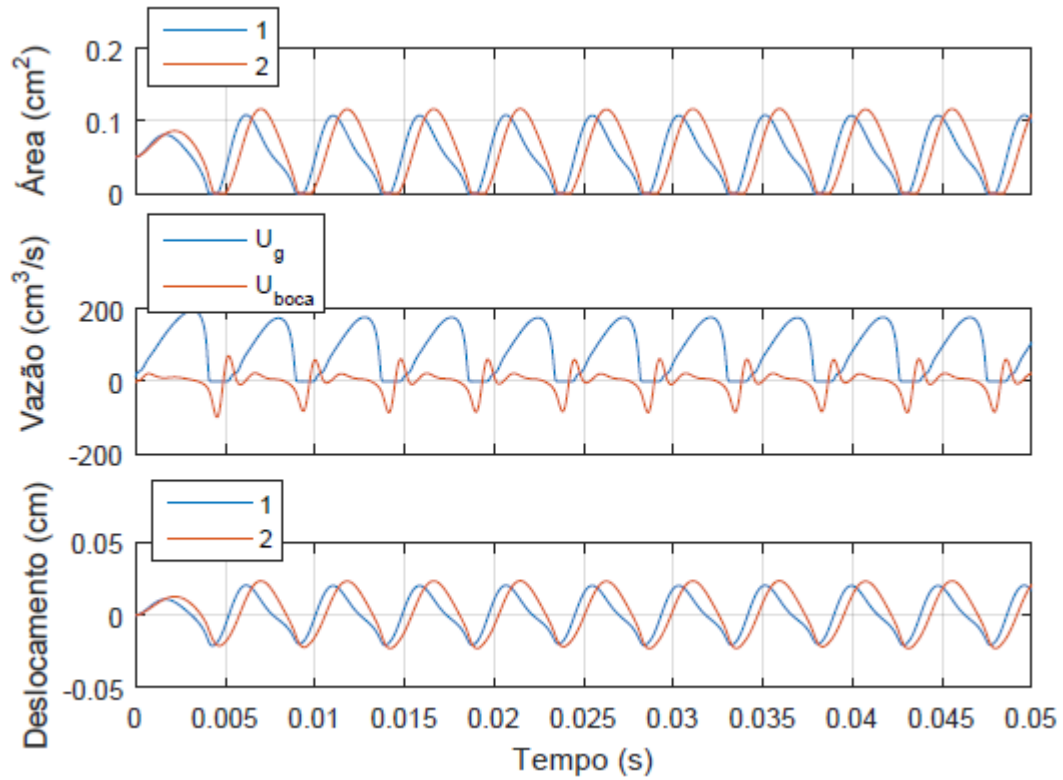


Fonte: Autoria própria.

Comparativamente aos gráficos gerados na simulação de referência, caso (b) – Figura 31, situação em que $k_1 = 80$ kdyn/cm, constata-se que a amplitude dos sinais da área e do deslocamento diminuem para a massa 1, mas também para a massa 2, no entanto de forma mais sutil. A diferença de fase entre as massas não sofre significativa alteração.

Para uma nova situação, na qual aumenta-se o coeficiente k_1 para 280 kdyn/cm, obteve-se os gráficos apresentados na Figura 32. Nota-se que houve uma grande redução em relação à amplitude do movimento das massas M_1 e M_2 em relação ao caso (a) da Figura 31. Portanto, para valores significativamente maiores que os valores típicos, k_1 terá efeito no movimento da massa M_2 . Isso acontece, porque, M_1 possui uma massa bem maior que M_2 , portanto seu movimento terá uma inércia maior e como ambas estão acopladas, M_1 influenciará no deslocamento de M_2 .

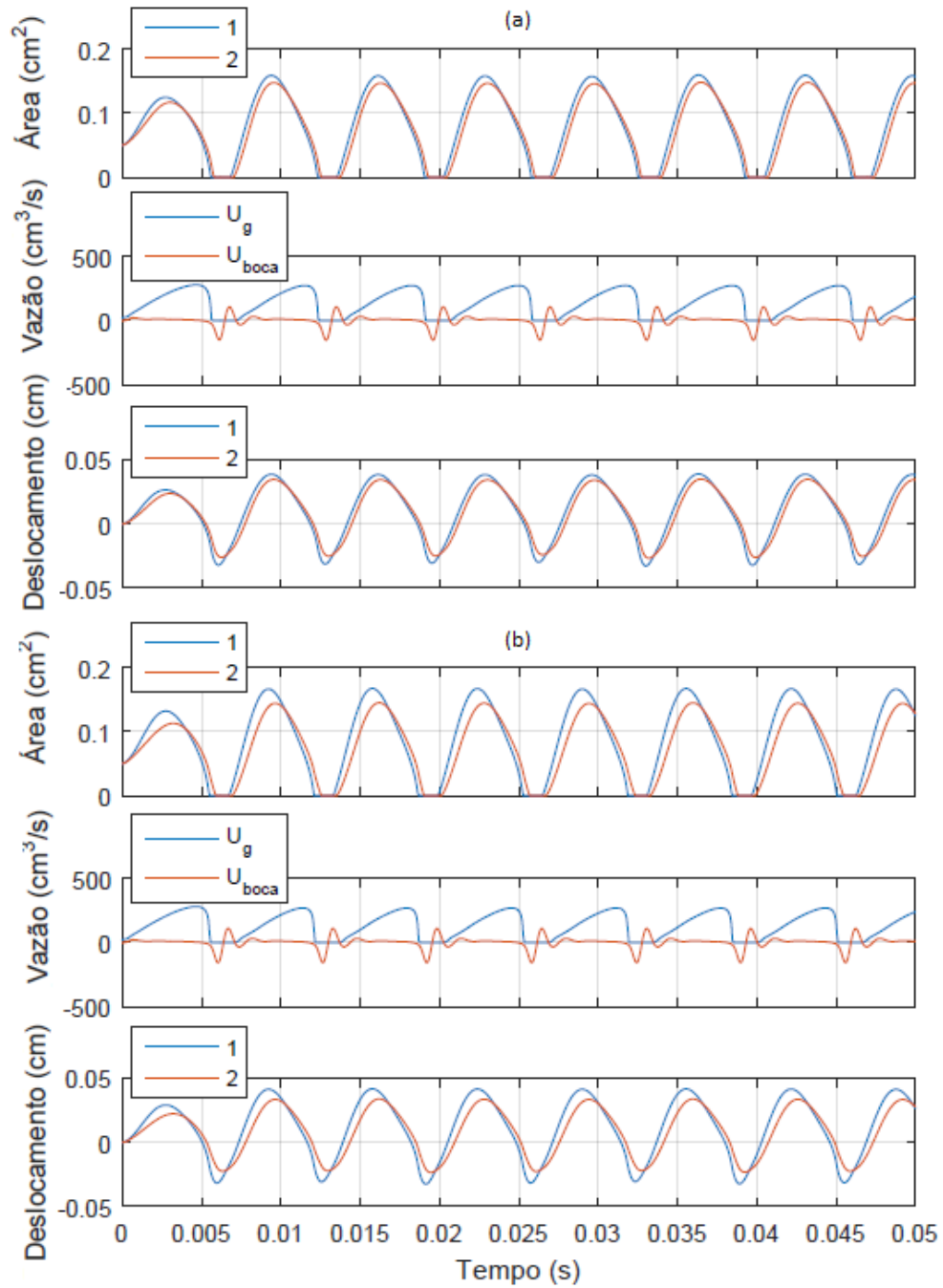
Figura 32 - Curvas correspondentes à simulação na qual $k_1 = 280$ kdyn/cm.



Fonte: Autoria própria.

Por fim, simulou-se várias situações para a variação dos valores do coeficiente de elasticidade k_c entre as massas M_1 e M_2 . Por meio dos resultados obtidos percebe-se que, o aumento de k_c resulta na diminuição da diferença de fase entre as ondas que representam tanto a área quanto o deslocamento percorrido pelas massas M_1 e M_2 . Essa mudança pode ser observada comparando-se os casos (a) e (b) na Figura 33, no qual o caso (b) é o sinal de referência.

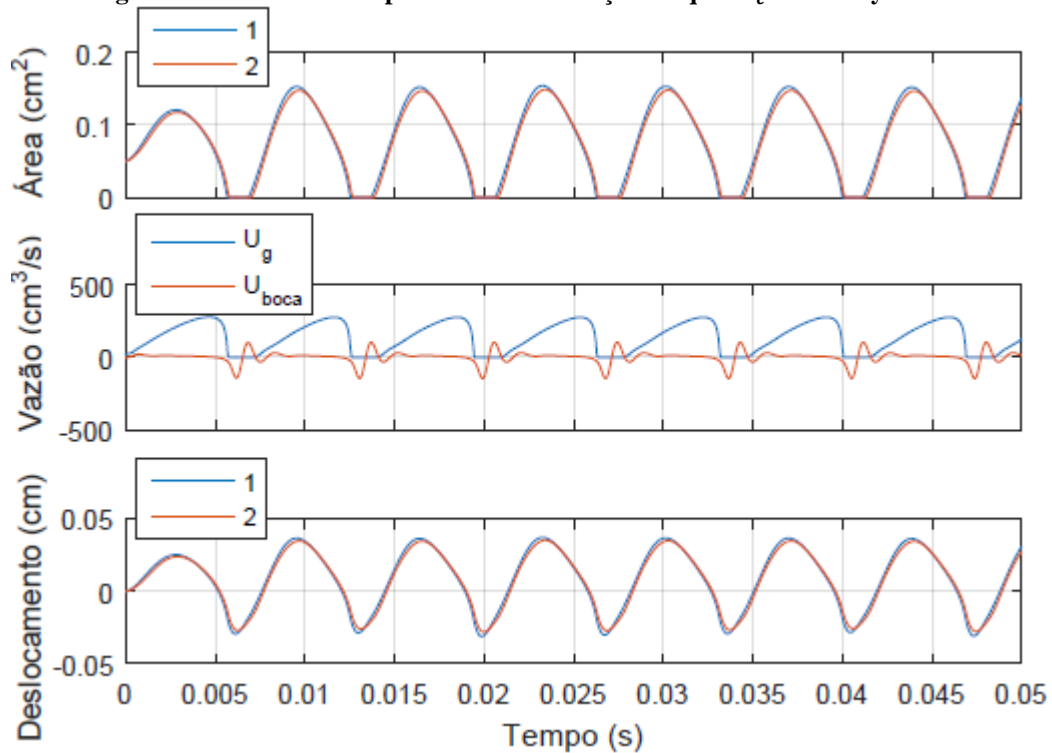
Figura 33 - Curvas correspondentes às simulações nas quais $k_c = 160$ kdyn/cm (a) e $k_c = 80$ kdyn/cm.



Fonte: Autoria própria.

Para o caso em que k_c sofra um aumento significativo, acima dos valores típicos, nota-se que o movimento das massas M_1 e M_2 acontece praticamente em fase, como se estas formassem uma unidade. Este comportamento pode ser visto por meio da Figura 34.

Figura 34 - Curvas correspondentes à simulação na qual $k_c = 360$ kdyn/cm.



Fonte: Autoria própria.

Este fenômeno ocorre devido ao coeficiente de elasticidade entre as massas ser tão grande ao ponto de limitar o movimento de uma massa em relação à outra e, desta forma, influenciar diretamente na dependência dos seus deslocamentos. Segundo Ishizaka e Flanagan (1972, pág. 1253), nesta condição o modelo de duas massas chega muito próximo à resultados obtidos por meio do modelo de uma massa.

Comparando as Figuras 33 (a) e 34, mais especificamente os gráficos referentes à vazão do fluxo de ar, torna-se evidente que o sinal de voz obtido na saída do sistema não sofre significativa interferência da variação do coeficiente de elasticidade k_c entre as massas.

4.2.4 Efeito da variação da área neutra

A partir da análise do comportamento do modelo das cordas vocais, em relação às áreas neutras, é possível realizar a correlação entre as vibrações do modelo estudado e as cordas vocais humanas (ISHIZAKA; FLANAGAN, 1972).

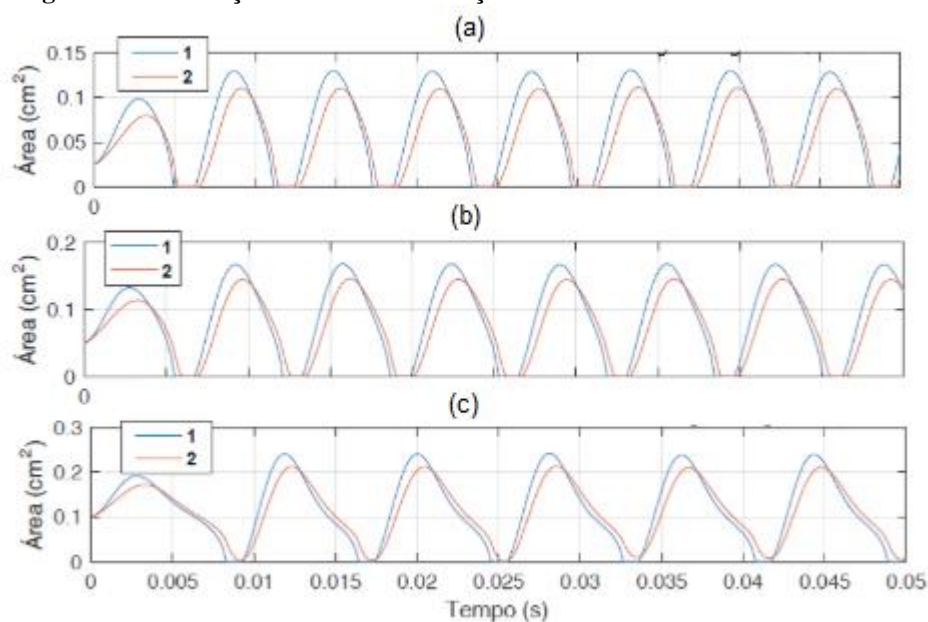
Variando os valores das áreas iniciais, sempre mantendo A_{g01} e A_{g02} iguais, para a simulação da vogal /a/, obteve-se diferentes resultados para avaliação. Foram analisados os

sinais da área glotal, deslocamento das massas e vazão.

Utilizando o valor das áreas neutras igual a $0,05 \text{ cm}^2$, foram obtidos os sinais que representam a área glotal, o deslocamento das cordas vocais e as vazões volumétricas. As quais foram apresentadas anteriormente nas Figuras 22, 23 e 24, respectivamente.

A Figura 35 ilustra os sinais obtidos para a área glotal considerando no caso (a) $A_{g01} = A_{g02} = 0,025 \text{ cm}^2$, no caso (b) $A_{g01} = A_{g02} = 0,05 \text{ cm}^2$, e no caso (c) $A_{g01} = A_{g02} = 0,1 \text{ cm}^2$.

Figura 35 - Simulação do efeito da variação da área neutra.

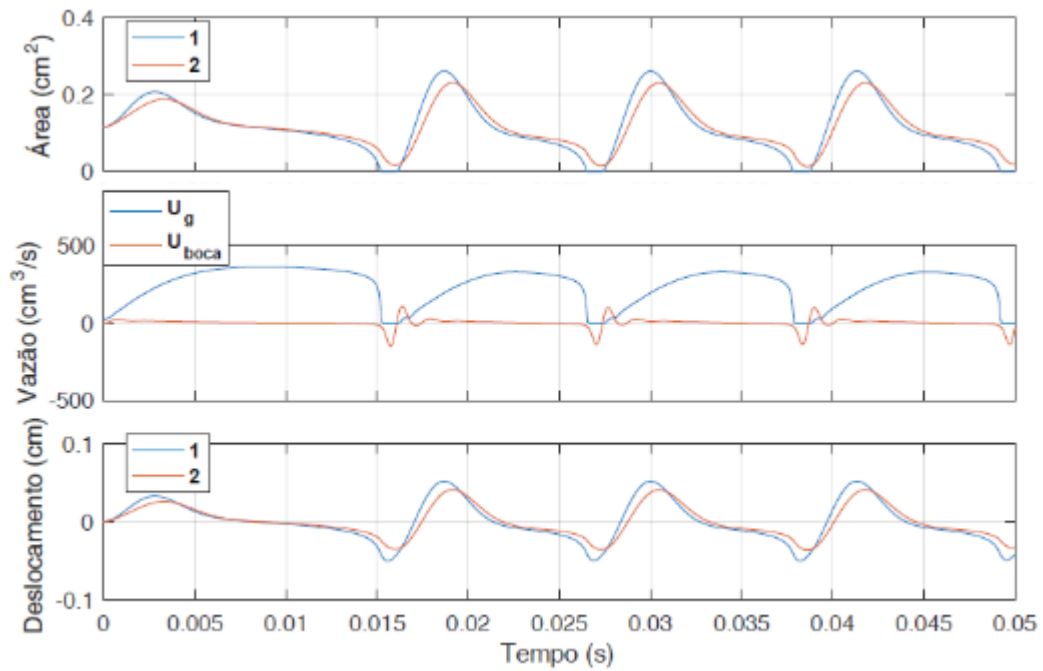


Fonte: Autoria própria.

Analisando o caso (a), apresentado Figura 35, nota-se uma diminuição na amplitude quando a área inicial foi reduzida. Com áreas neutras maiores, conforme caso (c), o valor das áreas glotais, o tempo para estabilizar, as amplitudes e o período de oscilação aumentaram e consequentemente a frequência diminuiu. Outra característica possível de se observar por meio dos resultados, é que quanto menor a área neutra, menor é o tempo necessário para estabilização do movimento.

Em seguida, essas áreas iniciais foram aumentadas para $0,115 \text{ cm}^2$ e $0,12 \text{ cm}^2$ respectivamente. E, conforme as variações foram sendo feitas, foi possível observar a existência de um limite crítico para o tamanho das áreas neutras. Como pode ser observado na Figura 36, com áreas neutras no valor de $0,115 \text{ cm}^2$ o sistema ainda consegue, mesmo que precise de um tempo maior, manter a oscilação típica do modelo.

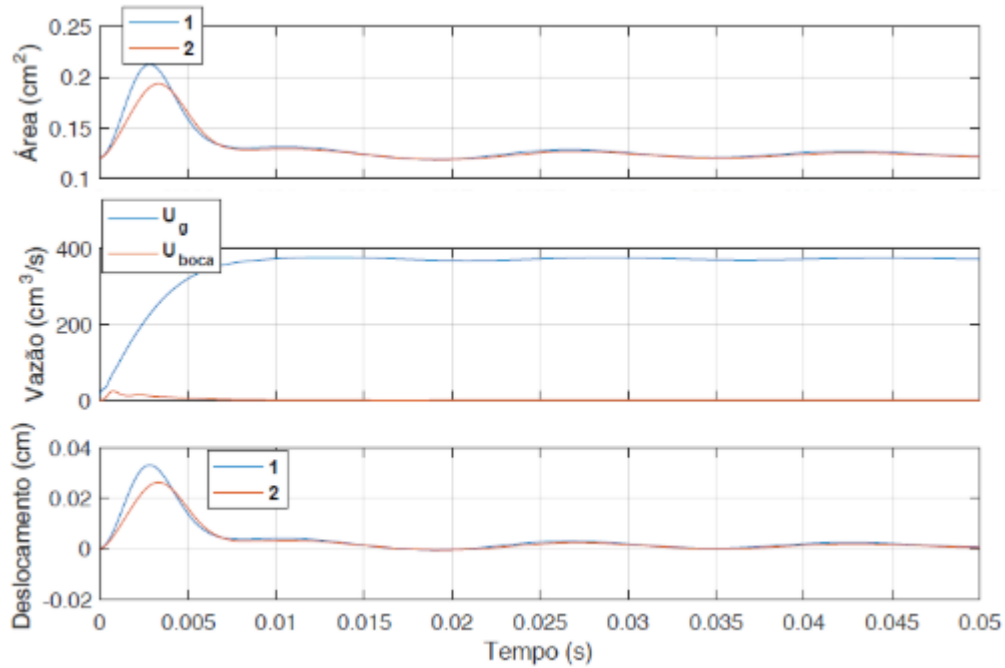
Figura 36 - Simulação do efeito da variação da área neutra.



Fonte: Autoria própria.

Foi verificado que quando ultrapassado o limite crítico, o sistema não corresponde às curvas características do movimento das cordas vocais. Esse efeito pode ser observado quando A_{g01} e $A_{g02} = 0,12 \text{ cm}^2$, que está apresentado na Figura 37. Nota-se neste caso que U_{boca} sofre um pequeno estímulo e, em seguida, retorna à zero. Isso acontece porque esses valores de área configuram uma abertura que permite a passagem direta do ar através da glote, sem possibilitar a interação do fluido com a estrutura das cordas vocais, ou seja, as cordas não vibram e, portanto, não geram o sinal de voz.

Figura 37 - Simulação do efeito da variação da área neutra.



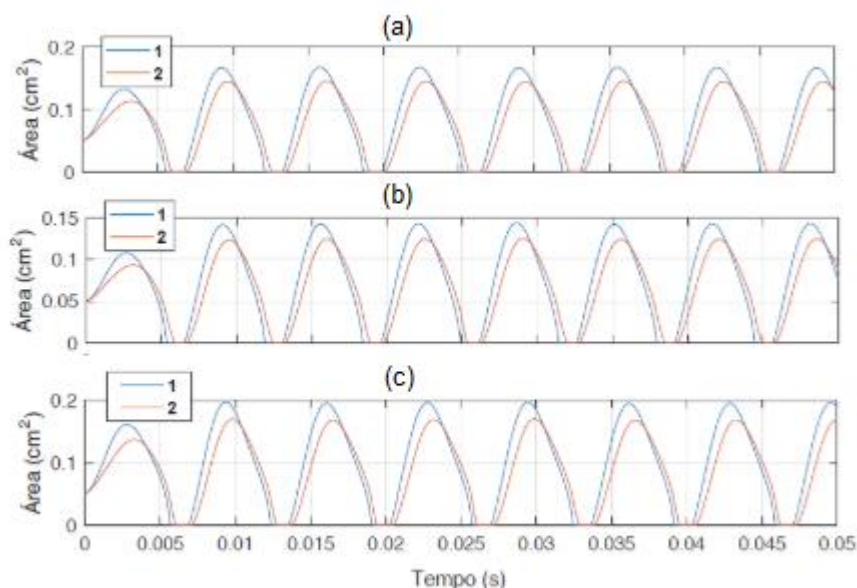
Fonte: Autoria própria.

4.2.5 Efeito da variação do comprimento da corda vocal

A partir da variação do tamanho das cordas vocais na simulação, é possível analisar os efeitos ocasionados por estes parâmetros. Como realizado nas outras variações, foram mantidos todos os outros valores com os valores de referência do Quadro 3, variando apenas o valor de l_g , comprimento das cordas vocais, proporcionalmente para mais e para menos.

Partindo de $l_g = 1,4$ cm, a primeira alteração foi a diminuição para $l_g = 1,2$ cm e depois aumento para 1,6 cm, as curvas obtidas em relação a área glotal estão na Figura 38.

Figura 38 - Simulação do efeito da variação do comprimento da corda vocal.



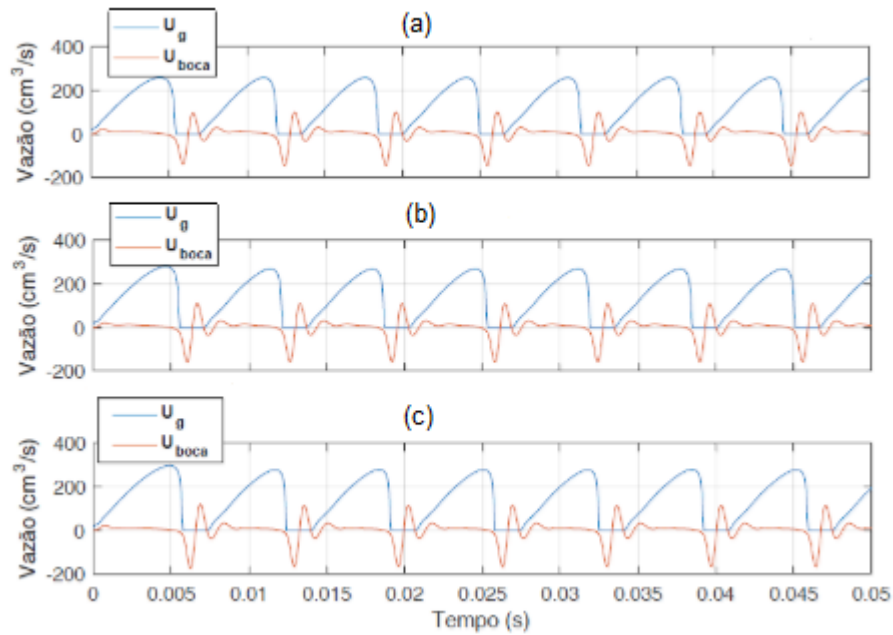
Fonte: Autoria própria.

Analisando os resultados obtidos em relação a área glotal, é constatado que quando o comprimento da corda vocal é reduzido, conforme caso (a), a amplitude da área e o tempo necessário para estabilizar diminui. Se o comprimento é maior, conforme apresentado no caso (c), a amplitude e o tempo aumentam. Observando a Figura 38, tem-se que a área para o caso (b) é em torno de pouco mais de $0,15 \text{ cm}^2$; para o caso (a) a área é menor que $0,15 \text{ cm}^2$ e no caso (c) a área glotal é quase $0,2 \text{ cm}^2$.

Não há mudanças significativas no período e na frequência das curvas devido a variação do tamanho do comprimento da corda vocal. Os efeitos ocorrem apenas em relação ao tamanho da área glotal e o tempo que o sistema necessita para estabilizar.

As mesmas alterações são verificadas na vazão volumétrica, aumento da amplitude e do tempo que o sistema leva para estabilizar a oscilação quando o comprimento da corda vocal é maior. Quando o comprimento é menor, obtém-se valores menores de amplitude e de tempo necessário para o sistema entrar em regime permanente. A Figura 39 apresenta as curvas obtidas das simulações na seguinte ordem: caso (a) no qual $l_g = 1,2 \text{ cm}$, caso (b) no qual $l_g = 1,4 \text{ cm}$ e caso (c) no qual $l_g = 1,6 \text{ cm}$.

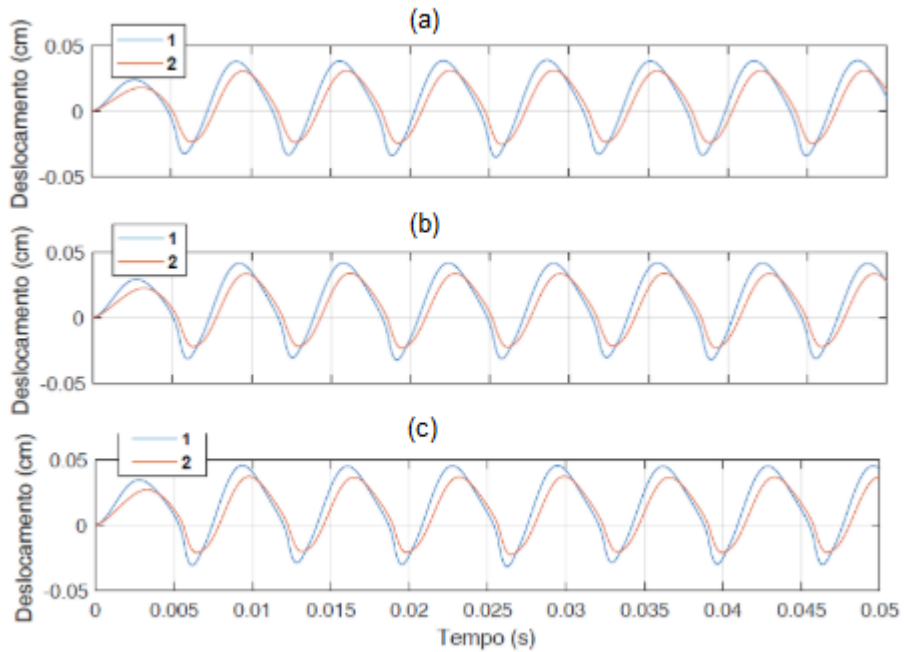
Figura 39 - Simulação do efeito da variação do comprimento da corda vocal.



Fonte: Autoria própria.

A Figura 40 apresenta os resultados obtidos para o deslocamento das massas, sendo o caso (a) com $l_g = 1,2$ cm, caso (b) com $l_g = 1,4$ cm e caso (c) com $l_g = 1,6$ cm.

Figura 40 - Simulação do efeito da variação do comprimento da corda vocal.



Fonte: Autoria própria.

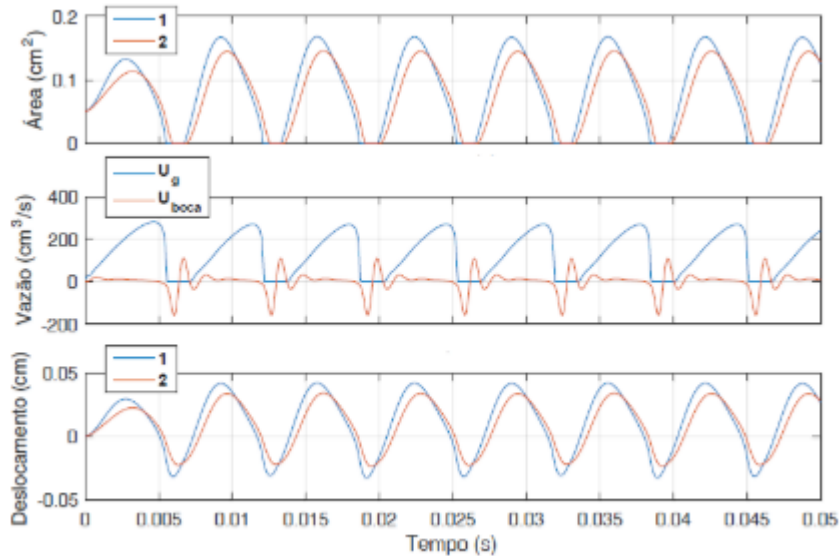
Por fim, analisando os efeitos no deslocamento das massas, observam - se os mesmos efeitos verificados na área glotal e na vazão. A amplitude e o tempo necessário para

estabilização da oscilação, são maiores quando comprimento da corda vocal é maior e menor quando o comprimento da corda vocal é reduzido.

4.2.6 Efeito da variação da espessura das massas

Variando a espessuras das massas M_1 e M_2 , identificaram-se algumas mudanças nos sinais simulados. A Figura 41 apresenta os resultados para a configuração original, em que a espessura da massa d_1 é 0,2 cm e a espessura da massa d_2 é 0,1 cm.

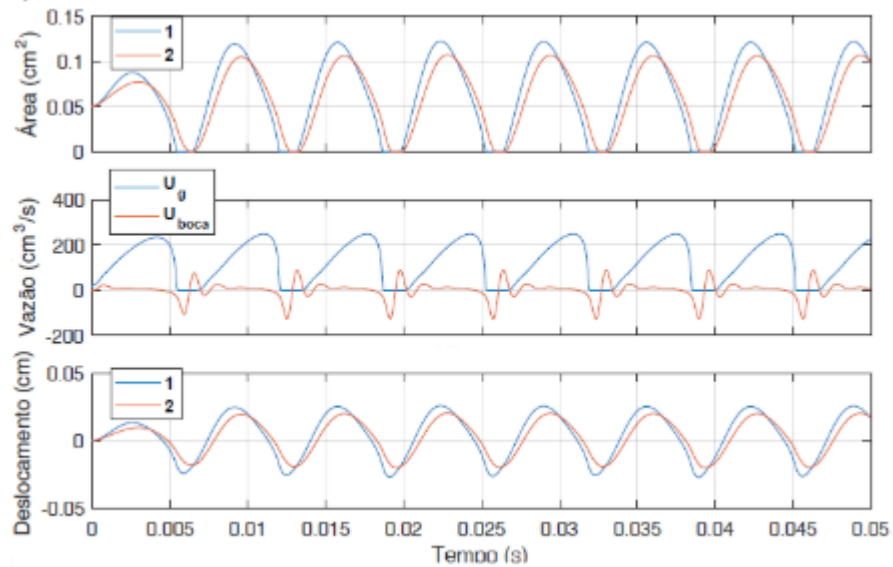
Figura 41 - Simulação do efeito da variação da espessura das massas.



Fonte: Autoria própria.

Reduzindo a espessura das massas obtivemos os resultados da Figura 42, tendo $d_1 = 0,1$ cm e $d_2 = 0,05$ cm. É possível perceber um aumento da diferença de fases das áreas glotais nas massas M_1 e M_2 , as cordas vocais mais finas ficam mais distantes uma da outra, desse modo o espaço entre elas é maior. A área, o deslocamento e a vazão, obtiveram redução da amplitude dos sinais.

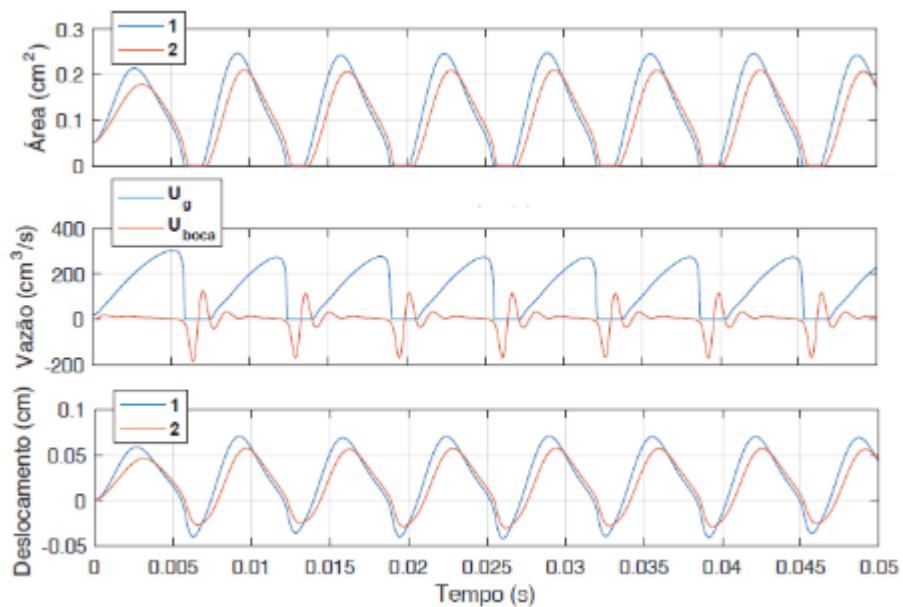
Figura 42 - Simulação do efeito da variação da espessura das massas.



Fonte: Autoria própria.

Para uma nova situação, tendo as espessuras maiores, $d_1 = 0,4$ cm e $d_2 = 0,2$ cm, observa-se na Figura 43, que a diferença entre fases da área é menor em comparação com o caso anterior em que as espessuras eram menores. Isso mostra que, quando as espessuras são maiores, as cordas vocais possuem um espaço menor entre elas, ficando mais próximas uma da outra. As amplitudes, da área e do deslocamento, obtiveram aumento significativos, enquanto que a vazão teve um aumento de amplitude sutil.

Figura 43 - Simulação do efeito da variação da espessura das massas.



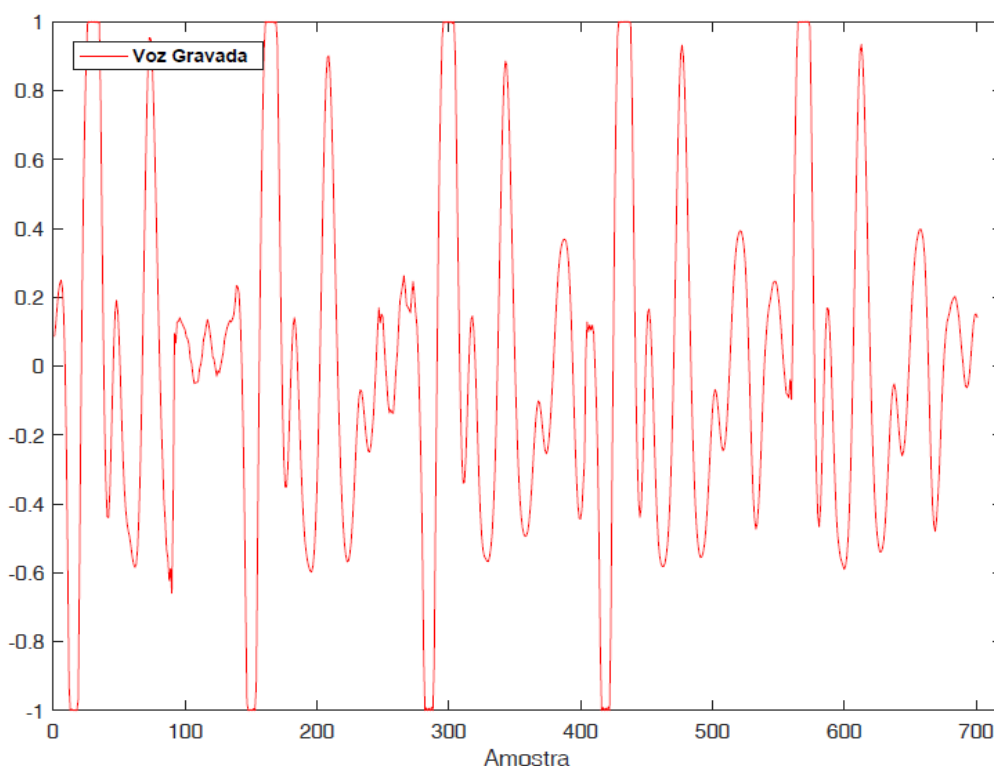
Fonte: Autoria própria.

4.3 SÍNTESE DA VOGAL /A/ APÓS AJUSTES

Após a obtenção dos resultados para o modelo de duas massas, com base nas simulações feitas por Ishizaka e Flanagan (1972) e realizado o estudo acerca da interferência de cada parâmetro presente no equacionamento no sinal de voz, tornou-se viável o refinamento do sinal de saída produzido pelo sintetizador, neste caso, da vogal /a/.

Para a realização dos ajustes, primeiramente foi necessário obter um sinal de voz real (voz humana) de base, a fim de aproximar o sinal sintetizado de uma condição mais natural possível. Para isso, optou-se por utilizar um sinal de voz masculino (25 anos), visando manter a concordância com o que foi feito por Ishizaka e Flanagan (1972). Este sinal foi gravado e sua onda gerada por meio do software MATLAB, conforme Figura 44.

Figura 44 - Voz humana masculina real gravada utilizando o MATLAB.



Fonte: Autoria própria.

A partir do sinal de voz padrão sintetizado, foram feitas variações simultâneas nos parâmetros estudados anteriormente, de modo a aproximá-los do sinal de voz gravado. Além disso, mudanças nos tubos cilíndricos que modelam a vogal foram necessárias para obter o menor erro possível. Partindo do modelo utilizado por Silva Jr. (2015), modelou-se cada um dos quatro cilindros, para a melhoria da síntese, e passou-se a usar os quatro tubos com as

dimensões diferentes. Essa alteração se deu de forma que a transição de um tubo para o outro ficasse mais suave.

Para determinar qual sinal sintetizado mais se aproximou do real foram sobrepostas as curvas da voz sintetizada e do sinal gravado por meio do MATLAB. E então, para mensurar o quão perto a voz sintetizada estava da voz gravada utilizou-se o cálculo da raiz do erro quadrático médio, conhecido como RMSE, que calcula a distância vetorial entre os sinais plotados. Este erro foi calculado através da Equação (50).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_1^n (\text{Voz teórica}_n - \text{Voz Real}_n)^2}. \quad (50)$$

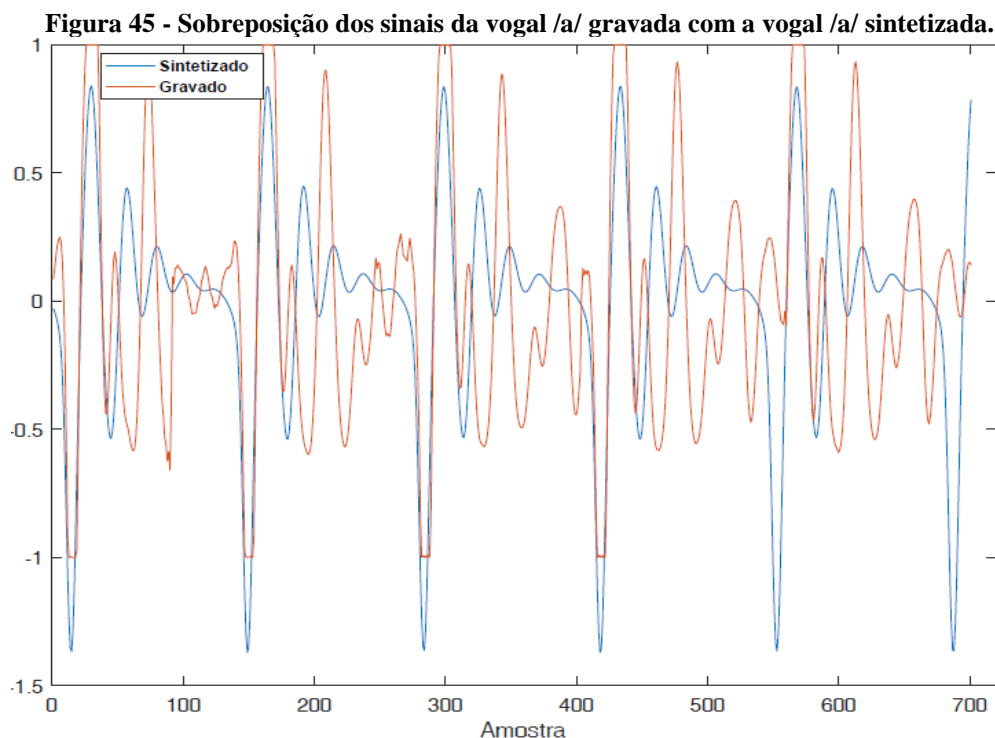
Os melhores casos obtidos com as variações, ou seja, com as menores distâncias vetoriais (erro) são apresentados no Quadro 4.

Quadro 4 - Variações dos parâmetros e seus respectivos RMSE.

		Original	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6	Caso 7
		Valor	Valor	Valor	Valor	Valor	Valor	Valor	Valor
Tubo 1	Comprimento da seção 1 [cm]	4	4,5	4,5	4,5	4,5	4,5	4,5	4,5
	Raio da seção 1 [cm]	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2
Tubo 2	Comprimento da seção 2 [cm]	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0
	Raio da seção 2 [cm]	0,2	0,45	0,45	0,45	0,45	0,45	0,45	0,45
Tubo 3	Comprimento da seção 3 [cm]	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5
	Raio da seção 3 [cm]	0,75	0,3	0,3	0,3	0,3	0,3	0,3	0,3
Tubo 4	Comprimento da seção 4 [cm]	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5
	Raio da seção 4 [cm]	0,75	0,4	0,4	0,4	0,4	0,4	0,4	0,4
Laringe	lg [cm]	1,4	1,4	1,6	1,4	1,5	1,6	1,6	1,6
	Ag01 [cm ²]	0,05	0,05	0,06	0,05	0,05	0,05	0,04	0,05
	Ag02 [cm ²]	0,05	0,05	0,06	0,05	0,05	0,05	0,04	0,05
	d1 [cm]	0,2	0,25	0,25	0,25	0,25	0,25	0,25	0,25
	d2 [cm]	0,1	0,05	0,05	0,05	0,05	0,05	0,05	0,05
	M1 [g]	0,1	0,1	0,1	0,15	0,15	0,15	0,15	0,15
	M2 [g]	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05
	k1 [dyn/cm]	80000	60000	60000	55000	55000	55000	55000	55000
	k2 [dyn/cm]	64000	50000	45000	45000	45000	45000	45000	46000
kc [dyn/cm]	80000	60000	25000	55000	55000	55000	55000	55000	
	Ps [dyn/ cm ²]	8000	8000	8000	8000	8000	8000	8000	8000
RMSE	Raiz do Erro Médio Quadrático	0,5648	0,5477	0,5757	0,4732	0,4618	0,4608	0,4863	0,4592

Fonte: Autoria própria.

Como pode ser observado no Quadro 4, o melhor resultado foi obtido no caso 7. Para efeito de comparação, é apresentada na Figura 45 o sinal de voz sintetizado, correspondente ao caso 7 (curva azul), sobreposto à curva do sinal de voz gravado (curva vermelha).

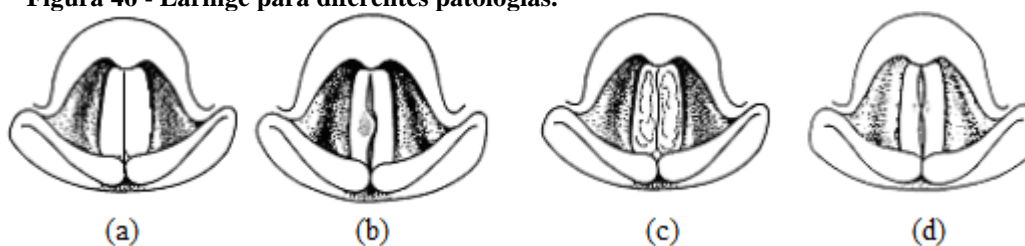


Fonte: Autoria própria.

4.4 RELAÇÃO DOS PARÂMETROS COM VOZES PATOLÓGICAS

Pode-se ainda relacionar o estudo das variações feitas com os parâmetros com alterações causadas no sinal de voz quando há alguma anormalidade nas cordas vocais. Tomando como base estudos feitos por Rosa (1998) e Zwetsch (2006) pode-se afirmar, por exemplo, que um cisto na laringe cria um efeito nas cordas que faz com que haja um aumento da massa vibratória Figura 46 (b), causada por uma protuberância na corda vocal resultante de um acúmulo de líquidos. Com o aumento da massa, no caso do modelo descrito, pode-se analisar como se as massas M_1 e M_2 fossem maiores do que para uma laringe saudável, Figura 46 (a). O que resulta significativamente em frequências de vibrações menores.

Figura 46 - Laringe para diferentes patologias.



Fonte: Adaptado de Zwetsch (2006).

Para o caso em que há algum edema nas cordas vocais, Figura 46 (c), a cobertura da corda vocal se torna menos rígida e a estrutura das cordas fica mais massiva. Ou seja, a presença de edemas está ligada com a alteração de parâmetros como as constantes de elasticidades das molas e o aumento das massas que compõem o modelo de duas massas para a descrição do movimento das cordas vocais no período de vocalização. Estes efeitos podem ser semelhantes ao caso de quando há nódulos, Figura 46 (d). A presença de nódulos interfere na vibração dependendo de sua firmeza e causa um aumento de massa, além de que os nódulos impedem que as cordas vocais se fechem no momento da vocalização.

5. CONCLUSÃO

O modelo de duas massas, adotado para a síntese dos sinais de voz, mostrou-se apropriado na reprodução dos sinais desejados, tanto do ponto de vista de computacional quanto no nível de detalhe das curvas obtidas. A facilidade na resolução das equações que modelam a dinâmica da interação fluido-estrutura foi possível devido à adoção de simplificações matemáticas, necessárias dado o alto nível de complexidade deste sistema. Portanto, percebe-se que, a resolução de um problema de engenharia, muitas vezes exige uma abordagem inicial básica (simplificada), para que gradualmente, conforme o grau de entendimento e domínio do assunto se estabelecem, este incorpore detalhes mais profundos e soluções mais sofisticadas possam ser implementadas.

O estudo da variação dos parâmetros, que compõem o equacionamento do modelo de duas massas, foi importante para alcançar uma voz sintetizada que mais se assemelhava a voz humana gravada. Analisou-se qual característica da onda sintetizada não correspondia a voz real e, tendo como base as análises feitas sobre cada parâmetro, iniciou-se os ajustes até encontrar-se o modelo que mais se assemelhava a voz humana gravada. Para chegar neste sinal, além das alterações nos tubos, foram reduzidos os coeficientes de elasticidade das massas, para que as cordas vocais possuíssem mais liberdade de movimento. Por se tratar de uma voz masculina, aumentou-se o comprimento da corda vocal para 1,6 cm, aumentando assim a amplitude do sinal sintetizado. A espessura da massa M_1 e o seu valor (em gramas) também foram consideradas maior, se comparadas ao modelo padrão estabelecido, para que obtivesse uma amplitude maior no sinal de voz e se aproximasse mais da voz gravada.

Outro ponto importante, se refere às simulações realizadas no MATLAB. Por meio destas, utilizando-se um intervalo de variação para os valores de cada parâmetro que compõem as EDOs, foi percebida a influência que cada um deles possui nos sinais resultantes. A partir disso, constatou-se o quão delicada é a situação de uma pessoa acometida por qualquer alteração em suas cordas vocais e que, dependendo do caso (doença), a comunicação desta pessoa será de fato prejudicada.

Para tanto, o desenvolvimento deste estudo pode ser utilizado como ponto de partida para trabalhos futuros com este direcionamento. Estes poderão utilizar métodos mais eficazes tanto em relação à discretização das equações que modelam o sistema como em relação ao estudo da variação dos parâmetros. Para este último, poderá ser usada metodologia computacional que permita a variação dos parâmetros de uma forma mais sofisticada, tendo em

vista que o presente trabalho baseou-se em tentativa e erro, o qual apesar da simplicidade é uma forma trabalhosa e demorada de obtenção dos resultados.

Esse trabalho também poderá contribuir para estudos futuros na área, que se concentrariam na identificação de padrões que permitam reconhecer uma dada doença em um sinal de voz. Para isso seria necessário obter um banco de dados com vozes saudáveis e patológicas. A forma de identificação poderia ser realizada por comparação pelo método dos mínimos quadrados ou outro método mais preciso. Essas identificações possibilitariam o diagnóstico de patologias na laringe de um paciente de uma forma indireta, ou seja, não invasiva. Além de permitir a visualização, por meio do sinal resultante (simulação), do impacto que uma cirurgia ou algum outro procedimento na laringe traria para a produção vocal de um paciente se alguma alteração na fisiologia deste órgão fosse realizada.

REFERÊNCIAS

BEHLAU, Mara. **Voz: O Livro do Especialista**. Rio de Janeiro, 2001. Vol 1.

CASSOL, Mauriceia. **Benefícios do Canto Coral para Indivíduos Idosos**. 2004. Tese de Doutorado, Universidade Católica do Rio Grande do Sul, Porto Alegre, 2004. Acesso em: 10 nov. 2017.

CATALDO, E; SAMPAIO, R; NICOLATO, L. **Uma discussão sobre modelos mecânicos de laringe para síntese de vogais**. ENGEVISTA, v. 6, n. 1, p. 47-57, abr. 2004. Acesso em: 27 abr. 2017.

CATALDO, E; LUCERO, J.C; NICOLATO, L; SAMPAIO, R. **Comparison of Some Mechanical Models of Larynx in the Synthesis of Voiced Sounds**. 2006. J. Braz. Soc. Mech. Sci. & Eng. vol.28 no.4 Rio de Janeiro Oct./Dec. 2006. Acesso em: 23 abr. 2017.

CVETICANIN, L. **Review on Mathematical and Mechanical Models of the Vocal Cord**. Journal of Applied Mathematics. doi: 10.1155/2012/928591, 2012.

DAJER, Maria Eugenia. **Análise de sinais de voz por padrões visuais de dinâmica vocal**. Tese (Doutorado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

FUKUYAMA, E. **Análise acústica da voz captada na faringe próximo à fonte glótica através de microfone acoplado ao fibrolaringoscópio**. Revista Brasileira Otorrinolaringol.vol.67, n.6, São Paulo, Nov/Dez. 2001.

FLANAGAN, James L.; LANDGRAF, Lorinda L. **Self Oscillating Source for Vocal Tract Synthesizers**. IEEE Transactions on Audio and Electroacoustics. Vol. AU-16, n.1.1968.

HANSEN, J. H.; GAVIDIA-CEBALLOS, L.; KAISER, J.F.A **Nonlinear operator - based speech feature. Analysis method with application to vocal fold pathology assessment**. Fellow, IEEE. 1998.

HIRANO, M. **Clinical Examination of Voice**. Springer, New York, 1981.

HUNGRIA, H. **Otorrinolaringologia**. 8. ed. Rio de Janeiro: Guanabara Koogan, 2000.

ISHIZAKA, K; FLANAGAN, J. L. **Synthesis of Voiced Sounds From a Two- Mass Model of the Vocal Cords**. The Bell System Technical Journal. Vol. 51. N. 6. July- August, USA, 1972.

MONTAGNOLI, Arlindo N.; FONSECA, Everthon S. **Larynx virtual surgery**. Seventh IEEE International Symposium on Multimedia, p.1. 2005.

MORI, Ricardo N. **Perfis acústicos para caracterização de patologias da laringe**. Tese (Doutorado em Engenharia Elétrica) - Escola de Engenharia São Carlos, Universidade de São Paulo, São Carlos, 2005.

RABINER, L. R.; SCHAFER, R. W. **Digital processing of speech signals**. Prentice Hall, Englewood Cliffs, NJ, EUA, 1978.

RAZERA, Daniel E. **Determinadores de pitch**. 2004. 97f. Dissertação (Mestrado em Engenharia Elétrica) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Paulo, 2004.

ROSA, Marcelo de Oliveira. **Análise acústica da voz para pré-diagnóstico de patologias da laringe**. 1998. Dissertação (Mestrado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 1998. doi:10.11606/D.18.1998.tde-11122015-144509. Acesso em:17 abr. 2017.

ROSA, Marcelo de Oliveira. **Laringe digital**. 2002. Tese (Doutorado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2002. doi:10.11606 /T.18.2002.TDE-19112015-110520. Acesso em:23 abr. 2017.

ROSA, Marcelo de Oliveira. **Modelagem da Laringe: da biologia ao computador**. 2011. Revista de Letras. V.30, 1/4, 70-81, jan.2010/dez.2011.

SAMPAIO, R.; CATALDO, E.; BRANDÃO, A. **Análise e processamento de sinais**. Sociedade Brasileira de Matemática Aplicada e Computacional. São Paulo, 2006.

SCALASSARA, Paulo Roberto. **Utilização de Medidas de Previsibilidade em Sinais de Voz para Discriminação de Patologias de Laringe**. Tese (Doutorado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2009. Acesso em: 27 set. 2017.

SILVA Jr., Celso Donizete. **Modelagem Acústica do Trato Vocal Humano Pelo Método dos Elementos de Contorno**. 2015. Projeto de Graduação em Engenharia Mecânica, Faculdade de

Tecnologia da Universidade de Brasília, Brasília, 2015.

SOUZA, Leonardo Mendes de. **Detecção inteligente de patologias na laringe baseada em máquinas de vetores de suporte e na transformada.** 2011. Dissertação (Mestrado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2011. DOI: 10.11606/D.76.2011. TDE-13042011-112203. Acesso em: 17 abr. 2017

TITZE, Ingo. R. **Principles of voice production.** 1994. Prentice-Hall, NJ: Englewood Cliffs, NJ.

TEFFAHI, H. **A Two-Mass Model of the Vocal Cords: Determination of Control Parameters.** Houari Boumedién University of Sciences and Technologies. Algiers, Algeria, 2009.

VIEIRA, Júlio Doin. **Laringe: falsas Cordas Vocais e as Cordas Verdadeiras.** Arquivos Catarinenses de Medicina, v.32, n.1, 2003.

VIEIRA, Maurílio Nunes. **Uma Introdução à Acústica da Voz Cantada.** I Seminário Música Ciência Tecnologia: Acústica Musical. Universidade Federal de Minas Gerais, 2004.

VERONEZ, Djanira A. da Luz. **Abordagem Morfofuncional Do Sistema Respiratório.** Universidade Tecnológica Federal do Paraná. 2012.

ZWETSCH, I.C.; FAGUNDES, R. D.R.; RUSSOMANO, T.; SCOLARI, D. **Processamento Digital de Sinais no Diagnóstico Diferencial de Doenças Laríngeas Benignas.** Scientia Medica, Porto Alegre: PUCRS, v. 16, n. 3, jul. /Set. 2006.