

UNIVERSIDADE FEDERAL DE ALAGOAS  
FACULDADE DE LETRAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA

JULIO CESAR CAVALCANTI DE OLIVEIRA

**LARINGALIZAÇÃO NO PORTUGUÊS BRASILEIRO: UMA ANÁLISE EM TORNO  
DO FENÔMENO LARÍNGEO E IMPLICAÇÕES PARA A COMPARAÇÃO DE  
LOCUTOR**

MACEIÓ  
2017

**JULIO CESAR CAVALCANTI DE OLIVEIRA**

**LARINGALIZAÇÃO NO PORTUGUÊS BRASILEIRO: UMA ANÁLISE EM TORNO  
DO FENÔMENO LARÍNGEO E IMPLICAÇÕES PARA A COMPARAÇÃO DE  
LOCUTOR**

Dissertação apresentada ao Programa de  
Pós-Graduação em Letras e Linguística  
da Universidade Federal de Alagoas  
como requisito parcial para obtenção  
do grau de Mestre em Linguística.

**Orientadora: Profa. Dra. LUCIANA LUCENTE**  
**Co-orientador: Prof. Dr. PLINIO ALMEIDA BARBOSA**

Maceió  
2017

**Catálogo na fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**

Bibliotecária Responsável: Helena Cristina Pimentel do Vale

- O481 Oliveira, Julio Cesar Cavalcanti de  
Laringalização no português brasileiro: uma análise em torno do fenômeno laringeo e implicações para a comparação de locutor / Julio Cesar Cavalcanti de Oliveira. – 2017.  
93 f. : il.
- Orientadora: Luciana Lucente.  
Co-orientador: Plínio Almeida Barbosa  
Dissertação (mestrado em Letras e Linguística : Linguística) – Universidade Federal de Alagoas. Faculdade de Letras. Programa de Pós-Graduação em Letras e Linguística. Maceió, 2017.
- Bibliografia: f. 85-89.  
Anexos: f. 90-95.
1. Linguística. 2. Laringalização. 3. Prosódia. 4. Fonética acústica. 5. Fonética forense. I. Título

CDU: 801.4

 UFAL	UNIVERSIDADE FEDERAL DE ALAGOAS FACULDADE DE LETRAS PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA	 PPGLL
---	---	--

## TERMO DE APROVAÇÃO

JULIO CESAR CAVALCANTI DE OLIVEIRA

Título do trabalho: "LARINGALIZAÇÃO NO PORTUGUÊS BRASILEIRO: *uma análise em torno do fenômeno laríngeo e implicações para a comparação de locutor*"

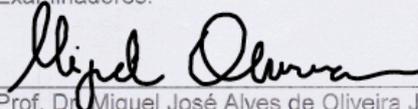
Dissertação aprovada como requisito para obtenção do grau de MESTRE em LINGUÍSTICA, pelo Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas, pela seguinte banca examinadora:

Orientadora:



Prof. Dra. Luciana Lucente (PPGLL/Ufal)

Examinadores:



Prof. Dr. Miguel José Alves de Oliveira Júnior (PPGLL/Ufal)



Prof. Dra. Avelina Mantovan Lima-Gregio (UnB)

Maceió, 20 de fevereiro de 2017.

Dedico este trabalho aos  
meus pais, Sérgio e Givoneide,  
por ampararem os meus sonhos.

## AGRADECIMENTOS

A "Deus" pelo dom da vida e pelo amor infinito. Por todas as graças que me tem concedido, das quais o direito de questionar, questioná-lo e conhecê-lo, a partir de uma experiência pessoal.

À Profa. Dra. Luciana Lucente, minha orientadora, pelo acolhimento inicial, pela amizade e incentivo. Por acreditar neste trabalho e investir nele. Pelo tempo dedicado, pela paciência e pela leveza com que conduziu cada etapa. Grato por tudo.

Ao Prof. Dr. Plinio Almeida Barbosa, meu co-orientador, alguém sem o qual este trabalho não seria possível. Obrigado pelo suporte técnico, científico e pessoal. A você toda a minha gratidão.

Ao Prof. Dr. Miguel de Oliveira Jr, professor querido por quem tenho muita admiração. Obrigado por suas avaliações criteriosas que ajudaram a aperfeiçoar este trabalho e contribuíram definitivamente para a minha formação.

À Profa. Dra. Luzia Miscow da Cruz Payão, por me introduzir aos estudos da fonologia e da prosódia, pelo incentivo acadêmico, por tudo que me permitiu aprender enquanto seu orientando de graduação.

Agradeço aos pesquisadores Profa. Dra. Cristina Felipeto e Prof. Dr. Eduardo Calil, por terem me apresentado o universo encantador dos estudos linguísticos e despertado em mim o espírito investigativo, quando eu ainda bolsista de iniciação científica no primeiro ano de graduação.

Aos meus colegas do PPGLL-UFAL, pelo companheirismo e apoio, em especial a Daniela Tavares, Ana Maria e Maraísa Espíndola.

Aos colegas do Fonufal, pelas discussões enriquecedoras e pelos bons encontros.

À Universidade Federal de Alagoas (UFAL) pelo espaço e pelo suporte técnico.

À Universidade Estadual de Campinas (UNICAMP), na pessoa do Prof. Dr. Plinio Almeida Barbosa, pela abertura e suporte técnico-científico.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro a esta pesquisa.

Agradeço aos meus pais, Sérgio Muniz e Givoneide Cavalcanti, meus maiores incentivadores. Meus verdadeiros alicerces.

Ao meu irmão Caio Victor Cavalcanti, pela existência.

A Francisco Ribeiro, pelo apoio, incentivo e companheirismo durante todo o tempo; à sua irmã, Aline Rocha, pela amizade e carinho, e a Petrócia pelas orações.

A Paula Delfino, pela amizade sincera e presente.

Às minhas primas Tamara Cavalcanti, Bruna Santos Oliveira, Priscilla Oliveira e demais primos.

À minha tia Salezia Magna, pela inspiração.

À minhas avós Maria Olívia e Maria Cavalcanti, e avôs Cícero Oliveira e Audálio Badega.

À minha madrinha Ranúzia Badega e tio Nivaldo Badega.

À toda minha família e amigos: obrigado.

## RESUMO

Este estudo investiga a ocorrência da laringalização no Português Brasileiro, avaliando o seu status enquanto pista prosódica na tarefa de comparação de locutor, à medida em que também considera as influências das variáveis sexo, tipo de segmento e estrutura prosódica da língua. Busca-se avaliar de que forma as variações nas taxas de produção do fenômeno, a exemplo das taxas de laringalização consonantal, vocálica e total, poderiam fornecer evidências acerca do perfil locucional de diferentes falantes, como também, analisar de que forma os parâmetros fonético-acústicos e a estrutura prosódica estão relacionados à sua ocorrência, de modo a possibilitar uma caracterização. A análise desenvolvida é de base semiautomática, obedecendo a um critério que considera a percepção auditiva, inspeção manual e análise do sinal de fala, no entanto, a utilização de *scripts* específicos garantiu maior automatização na tarefa de segmentação prosódica e extração das medidas de interesse, assegurando uma maior consistência na tarefa de comparação dos resultados. A análise do registro de fala semi-espontâneo de dez sujeitos, cinco homens e cinco mulheres, revelou o potencial distintivo do fenômeno para a diferenciação dos perfis de elocução, mediante consistentes variações nas taxas de laringalização vocálica e laringalização total. No que diz respeito a comparação dos grupos, constatou-se maiores taxas do fenômeno referentes ao grupo feminino, o que demonstrou ser estatisticamente significativo. Dentre os parâmetros analisados, a porcentagem da queda de intensidade entre os registros modal e laringalizado foi a medida mais modificada pela ocorrência do fenômeno. A análise dos contextos de ocorrência da laringalização permitiu comprovar o papel da estrutura prosódica enquanto motivadora dos eventos laríngeos, sinalizando para uma provável relação entre laringalização e a função de marcar fronteiras prosódicas no Português Brasileiro.

**Palavras-chave:** Laringalização; Prosódia; Fonética acústica; Fonética Forense

## ABSTRACT

This study investigates the occurrence of laryngalization in Brazilian Portuguese, assessing its status as a relevant prosodic clue in the speaker comparison task, while considering the influences of sex, segment type and prosodic structure. We evaluate how variations in the production of the phenomenon related to different segments, such as consonantal, vowel and total laryngalization rates, could provide evidence about the locutional profile of different speakers. It is also intended to analyze how phonetic-acoustic parameters are related to its occurrence, in order to allow an objective characterization. The analysis developed is based on a semiautomatic method, following a criterion that considers the perception, manual inspection and the analysis of the speech signal, however, the use of specific scripts guaranteed greater automation in the task of prosodic segmentation and extraction of the measures of interest. The analysis of the semi-spontaneous speech of ten subjects, five men and five women, revealed the distinctive potential of the phenomenon for the differentiation of speaking styles, through consistent variations in the vowel laryngalization and total laryngalization rates. It was also found higher rates of the phenomenon concerning to the female group in comparison to the male group, which proved to be a significant result. Among the analyzed parameters, the percentage of the decrease in intensity between modal and laryngeal events was the parameter most modified. The verification of the laryngalization contexts allowed us to prove the role of the prosodic structure as a motivator of the laryngeal phenomenon, indicating a link between laryngalization and demarcation of prosodic boundaries in Brazilian Portuguese.

**Keywords:** Laryngalization; Prosody; Acoustic phonetics; Forensic Phonetics

## LISTA DE FIGURAS

- FIGURA 1: Visão anterior da laringe. Estruturas cartilagosas, membranosas e óssea (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P; 2006)..... 18
- FIGURA 2: Visão posterior da laringe. Estruturas cartilagosas, membranosas e óssea (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P; 2006)..... 19
- FIGURA 3: Visualização da glote em posição de repouso e em coaptação (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P; 2006). ..... 21
- FIGURA 4: Representação do trato vocal, cavidades e articuladores. .... 22
- FIGURA 5: O contínuo representando os tipos de fonação segundo Ladefoged (1971).  
..... 24
- FIGURA 6: Parâmetros de controle laríngeo para a descrição das configurações articulatórias de acordo com Laver (1980).....26
- FIGURA 7: Representação do modo de fonação modal. Observa-se a coaptação completa das pregas vocais e a não adução (constricção) das pregas ariepiglóticas (válvula 3). Exemplo retirado do trabalho de Edmondson & Esling (2006)..... 32
- FIGURA 8: Representação do estado da glote durante a laringalização (creak voice). Observa-se a adução (constricção) das pregas ariepiglóticas, ou da terceira válvula; bem como, a adução das pregas aritenoides, repercutindo na movimentação ântero-posterior das pregas vocais. Exemplo retirado do trabalho de Edmondson & Esling (2006). ..... 32
- FIGURA 9: Representação dos estados da glote durante a ativação da válvula 1 (registro modal), válvula 2 (registro laringalizado) e válvula 3 (constricção laríngea) registrados por Edmondson & Esling (2006). A = aritenóides; AE = pregas

ariepiglóticas; C = cartilagem cuneiforme; E = epiglote; V = pregas vocais (glote).  
 ..... 33

FIGURA 10: Disposição das camadas de segmentação no software Praat, de acordo com nível de análise (segmental e prosódico). ..... 49

FIGURA 11: Exemplo de laringalização vocálica durante o trecho final da palavra “pêras”. Na primeira camada os rótulos classificam os segmentos em vogal ou consoante, na segunda identifica-se o segmento com o rótulo modal (m) ou laringalizado (l). Observa-se na vogal ‘a’ a irregularidade no sinal acústico, reforçada pela ocorrência de pulsos glotais espaçados no espectrograma. .... 56

FIGURA 12: Exemplo de laringalização consonantal no primeiro segmento da frase “na árvore”. Na primeira camada os rótulos classificam os segmentos em vogal ou consoante, na segunda identifica-se o segmento com o rótulo modal (m) ou laringalizado (l). Observa-se na vogal ‘a’ a irregularidade no sinal acústico, reforçada pela ocorrência de pulsos glotais espaçados no espectrograma. .... 56

FIGURA 13: Relação entre taxa da laringalização em vogais e taxa de laringalização em consoantes. Diferenças estatisticamente significativas na comparação das proporções  $p = 0.006$ . ..... 57

FIGURA 14: Taxas individuais de laringalização em vogais. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 1.255 \cdot 10^{-11}$ ), diferença estatisticamente significativa na comparação entre os grupos F- feminino e M- masculino ( $p = 5.539 \cdot 10^{-8}$ ). ..... 58

FIGURA 15: Taxas individuais de laringalização em consoantes. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 0.0003147$ ), diferença estatisticamente significativa na comparação entre os grupos F- feminino e M- masculino ( $p = 0.01269$ ). ..... 59

FIGURA 16: Taxas individuais de laringalização total. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 4.289 \cdot 10^{-13}$ ),

diferença estatisticamente significativa na comparação entre os grupos F-  
feminino e M- masculino ( $p = 5.225.10^{-9}$ )..... 59

FIGURA 17: Proporções da ocorrência de laringalização em fronteiras de grupos  
acentuais vs. não fronteira. Valores individuais. .... 65

**LISTA DE TABELAS**

TABELA 1 - Combinações significativas ou marginalmente significativas taxa de laringalização vocálica. ....	61
TABELA 2 - Combinações significativas ou marginalmente significativas taxa de laringalização consonantal. ....	62
TABELA 3 - Combinações significativas ou marginalmente significativas taxa de laringalização total. ....	63
TABELA 4 - Proporções da ocorrência de laringalização em fronteiras de grupos acentuais vs. não fronteira. ....	66
TABELA 5 - Modificações nos parâmetros fonético-acústicos envolvidos na laringalização. ....	67

## SUMÁRIO

INTRODUÇÃO .....	16
1. SOBRE A LARINGALIZAÇÃO .....	18
1.1. O subsistema laríngeo .....	18
1.2. Registros vocais: diferentes tipos de fonação .....	22
1.3. A Laringalização nas línguas naturais: correlatos acústicos, fisiológicos e perceptuais .....	30
1.4. Laringalização: Motivações prosódicas .....	35
1.5. Relação entre taxa de laringalização e sexo .....	39
1.6. Implicações forenses .....	40
2. MÉTODO .....	44
2.1 Participantes .....	44
2.2. Composição do Corpus .....	45
2.3. Análise dos dados .....	46
2.3.1. Análise da laringalização .....	46
2.3.2. Scripts de análise .....	52
3. RESULTADOS .....	56
3.1. Taxas de laringalização vocálica e consonantal.....	57
3.2. Variação nas taxas de laringalização entre-sujeitos.....	57
3.3. Laringalização e sexo .....	64
3.4. Ocorrência da laringalização em relação ao grupo acentual.....	64
3.5. Parâmetros fonético-acústicos envolvidos na laringalização .....	66
4. DISCUSSÃO .....	69
4.1. Laringalização vocálica e consonantal .....	69
4.2. Implicações para a comparação de locutor .....	70
4.3. Variável Sexo e Laringalização .....	73
4.4. Sobre a relação entre laringalização e prosódia .....	75
4.5. Correlatos acústicos e fisiológicos da laringalização e fronteiras prosódicas .....	78
5. CONCLUSÃO .....	85

6. REFERÊNCIAS..... 87

## INTRODUÇÃO

A busca pelo aprimoramento de técnicas de análise acústica da voz e da fala, com a finalidade de contribuir para as áreas de fonética e fonoaudiologia forense, a partir de experimentos que utilizam a comparação de perfis de falantes, é o que motiva este trabalho.

O fato de que exista uma variação considerável na frequência e no modo com que os falantes fazem uso dos diferentes tipos de registro vocal durante o fluxo da fala, despertou a atenção dos pesquisadores por entender de que forma o uso dos diferentes modos de fonação podem estar associados a um comportamento linguístico particular e se há na sua transição alguma motivação prosódica.

A fim de permitir uma observação com maior controle experimental, optou-se inicialmente pela execução de um estudo piloto prévio, realizado com uma amostra de 6 sujeitos, os quais apresentavam idades e grau de escolaridade similares. Neste experimento, foi possível constatar variações no que diz respeito à incidência e à forma com que os indivíduos realizaram a mudança de registros vocais durante a fala espontânea, permitindo constatar que tais mudanças demonstravam ocorrer em ambientes linguísticos específicos, sinalizando alguma regularidade no contexto de sua realização, como já referido para algumas línguas naturais, a exemplo do inglês, chinês e dinamarquês (REDI & SHATTUCK-HUFNAGEL, 2001; BELOTEL-GRENIÉ & GRENIÉ; GRØNNUM, 2014). A partir dessas observações, foram levantadas as seguintes hipóteses.

A hipótese inicial era de que o fenômeno da laringalização, tipo de fonação não-modal, estivesse relacionado ao estilo de fala no português brasileiro (doravante PB) e, portanto, pudesse suscitar diferenças significativas no que diz respeito às taxas de ocorrência do fenômeno entre diferentes locutores. Tal variação, tanto na ocorrência do fenômeno, quanto no contexto de sua incidência, tornaria a laringalização uma pista prosódica relevante na tarefa de comparação de locutor.

Com base na literatura consultada e nas observações realizadas a partir do estudo piloto, hipotetizou-se ainda que a laringalização, embora não configure um traço contrastivo relevante no sistema fonológico do PB pudesse estar relacionada a

algum tipo de motivação prosódica que justificasse a sua ocorrência, reforçada pelos seus altos índices de ocorrência em finais de enunciado.

Aliado a isso, os seguintes objetivos foram traçados: i) analisar a ocorrência da laringalização no PB; ii) avaliar o fenômeno enquanto uma pista prosódica em potencial para a tarefa de comparação de locutor; iii) verificar de que forma a laringalização incide em diferentes tipos de segmentos; iv) analisar pontos recorrentes da laringalização por parte dos falantes do PB e a sua relação com a estrutura prosódica da língua; v) verificar a influência da variável sexo sobre as medidas de laringalização; e por fim, vi) avaliar os parâmetros fonético-acústicos associados à ocorrência do fenômeno, obedecendo a um critério que considera a percepção auditiva, a inspeção visual e a análise acústica do sinal de fala.

A escassez de trabalhos que analisem o fenômeno da laringalização no PB, bem como, de estudos que levem em conta esta medida como uma potencial pista na diferenciação de falantes, justifica a necessidade por melhor compreendê-lo, assim como verificar de que forma a organização prosódica e a variável sexo podem estar associados à sua ocorrência.

## 1. SOBRE A LARINGALIZAÇÃO

Para uma melhor compreensão dos processos e estruturas envolvidas na produção da voz e da fala, serão feitas algumas considerações a cerca da anatomia e fisiologia responsáveis pela dinâmica fonatória.

### 1.1. O subsistema laríngeo

A laringe é uma estrutura constituída por cartilagens, ligamentos, músculos e tecido membranoso. Encontra-se ligada a estruturas ósseas superiores e inferiores por meio de musculatura extrínseca. O esqueleto cartilaginoso da laringe é constituído sobretudo pelas cartilagens tireoide, cricoide e o par de cartilagens aritenoides (BARBOSA E MADUREIRA, 2015). As Figuras 1 e 2 ilustram o arcabouço laríngeo e permite a visualização dessas estruturas.

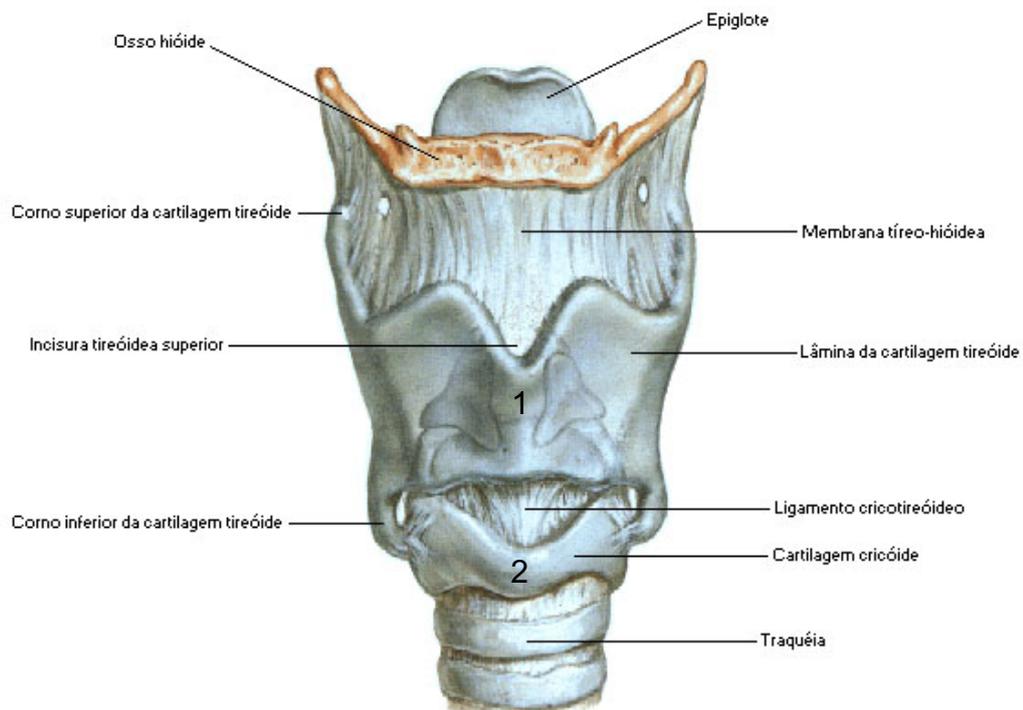


FIGURA 1: Visão anterior da laringe. Estruturas cartilaginosas, mebranas e óssea (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P; 2006).

A cartilagem tireoidea é uma cartilagem única, a maior cartilagem da laringe, e possui o formato de um escudo, sendo composta de duas lâminas laterais, de forma quadrangular, e dois pares de cornos (indicada na Figura 1 e 2 pelo número 1). O

ângulo de união entre as lâminas direita e esquerda, chamado de proeminência laríngea, varia de acordo com o sexo. Na população masculina, observa-se um ângulo ao redor de 90° e na população feminina o ângulo é mais aberto, com cerca de 120°. Esta variação gera grande impacto entre os sexos na fisiologia vocal, como por exemplo na medida que define o tamanho das pregas vocais e contribui na definição da frequência fundamental emitida (BEHLAU, 2001; ANDREWS, 2009)

A cartilagem cricoide também é uma cartilagem única, a segunda maior da laringe, ela articula-se com a cartilagem tireoidea (indicada na Figura 1 e 2 pelo número 2). O possível impacto da variação anatômica desta cartilagem na fonação entre homens e mulheres é desconhecido e pouco estudado (BEHLAU, 2001).

As cartilagens aritenóides (indicadas na Figura 2 pelo número 3), por sua vez, são um par de pequenas cartilagens móveis, consideradas a unidade funcional da laringe pela sua importância nas funções fonatória e respiratória. De forma resumida, são responsáveis pela aproximação e afastamento das pregas vocais. A variabilidade dessas cartilagens entre os sexos é muito pequena, representando os elementos cartilagosos de configuração mais estável da laringe (BEHLAU, 2001).

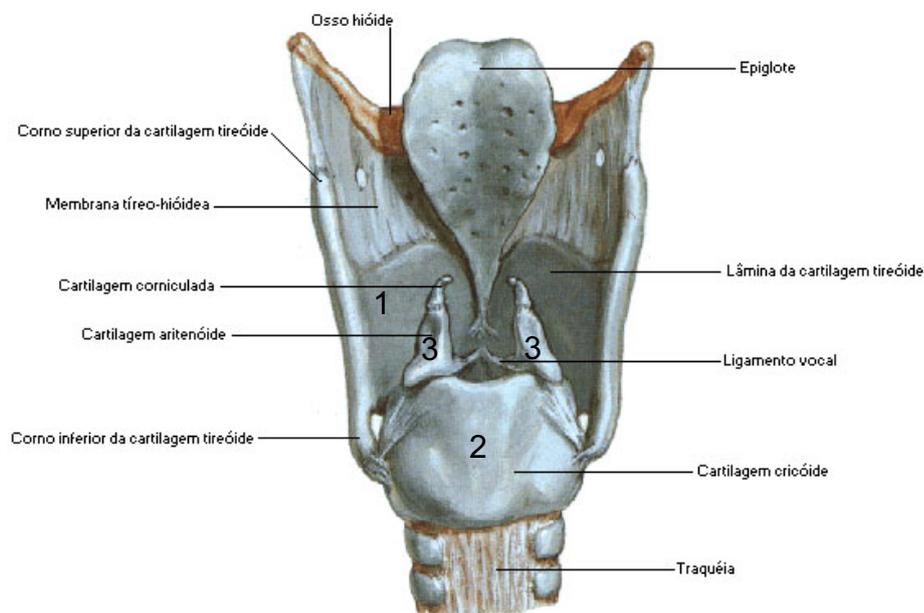


FIGURA 2: Visão posterior da laringe. Estruturas cartilagosas, membranosas e óssea (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P.; 2006).

A laringe como um todo divide-se em três espaços: supraglote, glote e infraglote. A rima glótica, ou simplesmente glote, é o espaço entre as pregas vocais, com altura de cerca de 1 cm no adulto (ver Figura 3). O som da voz é produzido na glote, sendo imediatamente acrescido de ressonância na própria supraglote (BEHLAU, 2001; ANDREWS, 2009).

No que diz respeito à atividade muscular, os músculos laríngeos podem ser classificados regionalmente como músculos intrínsecos e músculos extrínsecos. São denominados intrínsecos os músculos que têm origem e inserção na própria laringe, ao passo que os músculos extrínsecos apresentam apenas uma das inserções na laringe e outra fora dela, como no tórax, mandíbula ou no crânio (BEHLAU, 2001; ANDREWS, 2009).

De acordo com Behlau (2001), a musculatura intrínseca possui relação direta com a função fonatória, ela é reponsável por aproximar (aduzir), afastar (abduzir) e promover tensão das pregas vocais, controlando as funções laríngeas de respiração, esfíncter de proteção e fonação. Já a musculatura extrínseca, não interfere de modo direto na fonação, mas sua influência indireta é extremamente importante, modificando o posicionamento laríngeo, a ponto de repercutir na frequência fundamental da voz, ou de modo simplificado, na taxa de vibração das pregas vocais. A função básica da musculatura extrínseca é a de manter a laringe no pescoço, garantindo a sua estabilidade para que a musculatura intrínseca possa trabalhar efetivamente. Através da elevação ou abaixamento da laringe altera-se o ângulo entre as cartilagens e a tensão entre elas.

Na laringe encontra-se a estrutura mais nobre envolvida no processo de fonação, as pregas vocais; pequenas estruturas musculares que se aproximam para obstruir a corrente de ar laríngea ou se separam para liberar essa corrente de ar (KENT E READ, 2007), como ilustrado por meio da Figura 3.

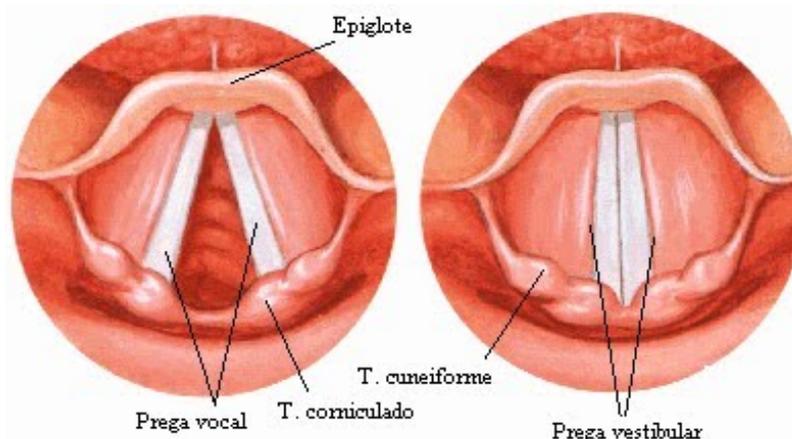


FIGURA 3: Visualização da glote em posição de repouso e em coaptação (SOBOTTA, J.; REINHARDS, P.; REINHARDS, P; 2006).

As pregas vocais são duas dobras de músculo e mucosa que se estendem horizontalmente na laringe, fixando-se anteriormente na face interna da cartilagem tireoidea, formando a comissura anterior, a região de convergência de ambas as pregas vocais. Posteriormente, cada prega vocal conecta-se à cartilagem aritenoidea, cobrindo-a de mucosa. O corpo da prega vocal é composto basicamente pelo músculo vocal e, do ponto de vista mecânico, quando se contrai, funciona como um feixe elástico muito rígido (BEHLAU, 2001).

O som é resultado da vibração das pregas vocais, que estão alternadamente juntas e separadas, em contato uma com a outra de forma basicamente periódica. A taxa de vibração das pregas vocais determina essencialmente a percepção do tom vocal de um falante (*vocal pitch*) (KENT E READ, 2007).

Como referido por Behlau (2001), existe uma variação considerável na frequência fundamental das vozes entre indivíduos de diferentes idades e sexos, ou seja, na frequência básica com que as pregas vocais vibram em um intervalo de segundo (F0). Quanto menor o tamanho da prega vocal, maior será a frequência fundamental da fala do indivíduo. Assim sendo, à proporção que o comprimento natural das pregas vocais aumenta, a frequência fundamental diminui.

O som fundamental produzido pelas pregas vocais é transportado por uma coluna de ar vibratória, atravessa o trato vocal e é transformado e enriquecido pelos sistemas ressonador e articulatório. O trato vocal, ou tubo vocal, começa nas pregas vocais (1) e abrange a parte superior da traquéia (2) e cavidades faríngea (3), oral (4) e nasal (5) (ANDREWS, 2009). Ver Figura 4.

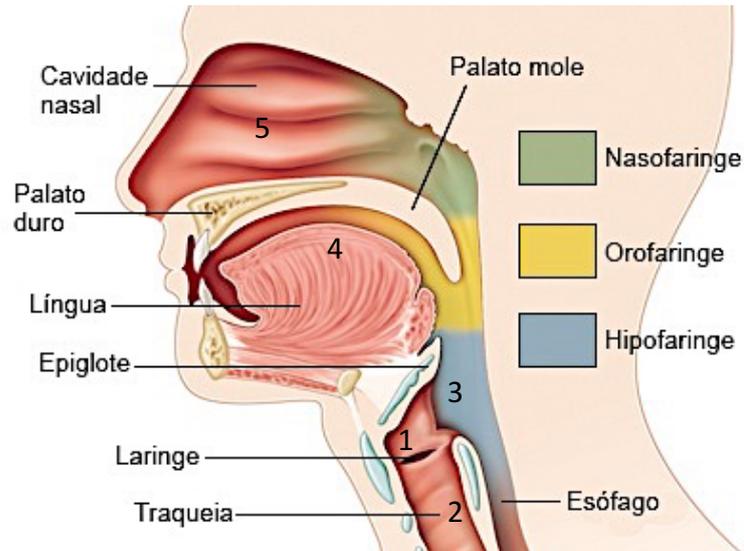


FIGURA 4: Representação do trato vocal, cavidades e articuladores. Fonte: Internet.

A qualidade da voz é resultado de uma complexa interação entre o modo de vibração das pregas vocais, ressonância supraglótica e frequências de vibração subjacentes (frequência fundamental, estrutura harmônica e a proporção harmônico ruído - HNR<sup>1</sup> (HIRSON E DUCKWORTH, 1993).

Para Hollien (1990) um bom conhecimento tanto da fisiologia como da acústica envolvidas na produção da fala/voz é necessário quando se busca analisar dados de natureza forense; saber algo sobre a estrutura da fala é imprescindível quando se quer entender o que pode estar "adequado" ou "alterado" quando um sujeito produz um enunciado particular ou como a fala pode ser decodificada mesmo quando distorcida. O autor refere a importância em se compreender como as mensagens são organizadas (ou codificadas), em um elo entre a produção da fala, transmissão (acústica) e percepção.

## 1.2. Registros vocais: diferentes tipos de fonação

<sup>1</sup> Índice que relaciona o componente harmônico versus o componente de ruído da onda acústica, correspondente ao termo do inglês "Harmonic-to-noise ratio" (HNR).

O termo “registro vocal” tem origem no funcionamento de instrumentos musicais, especialmente no órgão, no qual o conceito de registro relaciona-se a um grupo de tubos controlados por um mesmo fole ou pedal (BEHLAU, 2001).

No que diz respeito à voz humana, o registro refere-se aos diversos modos de emitir os sons da tessitura<sup>2</sup>. Assim, as frequências de um registro apresentam qualidade vocal quase idêntica, com a mesma base fisiológica, perceptiva-auditiva e acústica, ou seja, sons de um mesmo registro apresentam um caráter uniforme de emissão que permite distingui-los de sons pertencentes a outros registros (BEHLAU, 2001).

Segundo Laver (1980), o termo registro é particularmente utilizado na literatura do canto para se referir a modos particulares de vibração das pregas vocais. A nomenclatura também empregada, como muita frequência, para se referir ao nível da gama de pitch, em detrimento do modo com que as pregas vocais vibram. Para o autor, o termo é ambíguo e nem sempre é claro qual o significado associado ao seu uso.

Neste sentido, Hollien (1974) realiza uma descrição dos diferentes registros com base nas variações da gama de frequência fundamental e correlatos fisiológicos. De acordo com o autor os principais registros vocais são: basal, modal e elevado, com zonas de transição ou passagem entre eles. Desta forma, cada registro apresenta diferentes correlatos físicos, acústicos e perceptivos, como apresentados a seguir.

i) Registro basal: dispõe das frequências mais graves da tessitura, variando de 10 Hz a 70 Hz. Perceptualmente é possível ouvir pulsos de vibração durante a emissão, o que se assemelha a um ranger de porta, por isso, esse tipo de registro é também denominado de pulsátil. A intensidade nesse registro é baixa, sendo possível observar pregas vocais encurtadas e mais robustas durante a emissão. Esses correlatos físicos, fisiológicos e perceptivos estão comumente relacionados à ocorrência de laringalização.

ii) Registro modal: é o registro que mais utilizamos na fala habitual e também o de maior extensão, podendo ser dividido em modal grave, médio e agudo. As frequências desse registro estão entre 80Hz a 560Hz. No registro grave encontramos a laringe baixa, as pregas vocais espessas e com grande massa em vibração, o que

---

<sup>2</sup> A tessitura da voz falada abrange a gama de notas ou frequências que utilizamos na fala habitual sem que se gere fadiga vocal. Diz respeito à região de conforto na escala de frequência fundamental (BEHLAU, 2001).

facilita a emissão de tons graves, sendo, portanto, o registro principal da voz falada masculina. A subcategoria modal média representa uma fase intermediária entre grave e elevado. No registro elevado a laringe encontra-se em posição alta, com pregas vocais alongadas. Essa configuração laríngea permite a emissão de tons agudos, sendo um registro típico de vozes femininas (BEHLAU, 2001).

iii) Registro elevado: de modo semelhante ao registro basal, quase nunca ocorre na fala habitual. Nesse registro estão as frequências mais agudas que podemos emitir, de 160 Hz a 800 Hz. Apresenta duas subcategorias, uma maior e mais importante – o subregistro de falsete – e outra menor e de ocorrência rara – o sub-registro de flauta. A maior parte das notas do registro elevado constitui o sub-registro de falsete (BEHLAU, 2001).

No caso do falsete, a massa das pregas vocais que intervém na fonação é drasticamente reduzida. As pregas vocais estão em sua tensão longitudinal máxima, enquanto que as tensões medial e adutora são moderadas. Como consequência da grande redução da espessura das pregas vocais e da alta taxa de tensão longitudinal, a vibração é muito rápida e, portanto, a frequência fundamental sobe, resultando em uma voz de maior *pitch*<sup>3</sup> (Hollien, 1974).

De acordo com Berry (2001), o contraste entre a fonação modal e a não-modal está relacionado com a forma de vibração e o modo de coaptação das pregas vocais. Na fonação modal verifica-se a coaptação completa e a vibração regular das pregas vocais; já na fonação não-modal o mesmo não ocorre, refletindo uma configuração glotal mais complexa.

Andrews (2009) afirma que o registro modal é considerado o mais comum e o predominante durante a conversação. Segundo o autor, pode-se adotar duas grandes categorias para classificação dos tipos fonatórios: modal e não-modal. Esta última, abrange todas as variações de padrão fonatório não-modal, como o falsete e o *vocal fry*<sup>4</sup> (ou som basal).

---

<sup>3</sup> O *pitch* é a sensação psicofísica da frequência fundamental, correspondente à percepção de grave e agudo, e não deve ser confundido com a medida de frequência em si (BEHLAU, 2001). Por não apresentar tradução direta para o português, optou-se por empregar o verbete na língua original, o inglês.

<sup>4</sup> Correspondente ao registro laringalizado. Por vezes, o termo "laringalização" pode assumir outras nomenclaturas, como por exemplo "vocal fry", "glotalização", "creak", "creaky" e "som basal". Embora estes termos apresentem-se enquanto sinônimos em muitos trabalhos, nesta pesquisa iremos adotar o termo "laringalização" para fins de padronização.

Para Segundo, Alves e Trinidad, (2013) a voz modal também é o registro mais comum no fluxo da fala. A vibração das pregas vocais é periódica neste tipo de registro e a ondulação é larga. A pressão - longitudinal, adutiva e medial - aplicada pelas pregas vocais é moderada nesta emissão.

Segundo Ladefoged e Gordon (2001), as diferenças no tipo de fonação sinalizam importantes informações linguísticas em diferentes línguas. No que diz respeito aos estados da glote, os pesquisadores afirmam que os falantes podem controlar o modo como produzem os sons durante a fala, não apenas limitando-se a vibrações regulares das pregas vocais e variações de *pitch*; mas, também, produzindo outros tipos de padrões fonatórios, como o áspero, sussurrado, laringalizado e soproso.

Existem variações controláveis na atividade da glote, e não apenas possibilidades idiossincráticas ou ocorrências patológicas involuntárias. O que parece ser uma qualidade de voz patológica incontrolável para um determinado sujeito, pode ser uma parte necessária do conjunto de contrastes fonológicos de uma língua (LADEFOGED E GORDON, 2001).

Por exemplo, alguns falantes do inglês americano podem apresentar uma qualidade de voz muito soproso, o que será considerado patológico nos falantes dessa língua, enquanto que falantes da língua Gujarati<sup>5</sup> precisam de uma qualidade de voz semelhante para distinguir a palavra /b̄ar/ significado "do lado de fora" da palavra /bar/ significado "doze" (LADEFOGED, 1971). Como observado, uma qualidade vocal específica pode caracterizar um distúrbio de voz em uma determinada língua e um contraste linguístico em outra (LADEFOGED, 1983).

Ladefoged (1971) então sugere a possível existência de uma série contínua de tipos de fonação, definida em termos da abertura entre as cartilagens aritenóides, variando desde a produção afônica (sem vibração das pregas vocais), a uma qualidade vocal soproso, passando por modal, laringalizada, até a produção de uma oclusiva glotal<sup>6</sup>. Essa série contínua é descrita esquematicamente na Figura 5.

---

<sup>5</sup> Gujarati é uma língua indo-ariana, um dos 22 idiomas oficiais da Índia, e língua minoritária do Paquistão.

<sup>6</sup> A oclusiva glotal é uma articulação transiente de natureza consonantal, que tem como correlato acústico um som, essencialmente, brusco, semelhante ao som que as pessoas usam em expressões que possuem duas vogais seguidas como, por exemplo, em /ãʔã/ (expressão de negação). O ponto articulatorio é glotal, porém o modo de vibração é discutido, uma vez que diferentes autores trazem informações divergentes: para alguns o modo é vozeado, para outros é não vozeado, e, ainda, há os que defendem que o modo de vibração depende da consoante alvo (LIMA-GREGIO, 2011).



embora o haúça<sup>7</sup> e algumas outras línguas tchádicas façam tal contraste para fonemas oclusivos. As oclusivas glotais nessas línguas são implosivas e envolvem o abaixamento laríngeo, bem como uma qualidade de voz laringalizada (LADEFOGED E GORDON, 2001).

Alofones não modais segmentalmente condicionados em vogais são extremamente comuns na vizinhança de consoantes que são produzidas com fonação modal. Oclusivas surdas finais em certas variedades do inglês também desencadeiam uma fase de laringalização curta na extremidade da vogal imediatamente anterior. Em muitas línguas alofones não modais ocorrem em vogais que precedem consoantes oclusivas - sujeito a certas restrições prosódicas em algumas línguas (LADEFOGED E GORDON, 2001; REDI & SHATTUCK-HUFNAGEL, 2001).

De acordo com Laver (1980), a classificação inicial proposta por Ladefoged (1971) deixa alguns tipos de fonação sem uma identificação contrastiva com o modo neutro de vibração das pregas vocais, como é o caso do modo de fonação falsete. Segundo o autor, a diferença entre o tipo de fonação modal e os demais tipos de ajustes fonatórios são mais numerosos e complexos do que as descrições e classificações disponíveis até então.

Para o autor, explicar os diferentes tipos de modo de fonação com base apenas no processo de abertura e fechamento da glote não satisfaz as exigências implicadas nos diferentes tipos de ajustes fonatórios. Uma descrição espacial bidimensional do trato vocal seria suficiente para explicar a maior parte dos fenômenos fonéticos. No entanto, uma descrição tridimensional proporciona uma descrição mais precisa das modificações reais que ocorrem no trato.

De acordo com Laver (1980), na tentativa de caracterizar os diferentes tipos de configuração fonatória, a primeira condição é estabelecer o que seria o modo de fonação neutro em comparação aos demais modos descritos. Assim, o ajuste fonatório neutro corresponde aquele no qual a vibração das pregas vocais verdadeiras é periódica, eficiente e sem nenhum atrito auditivamente perceptível. Desta forma, existem outras configurações fonatórias que não são produzidas apenas pelas pregas vocais verdadeiras, como no caso da voz ventricular, por exemplo, ou no qual o modo de vibração é considerado não periódico, como na voz áspera, ou ineficiente e com algum grau de fricção, a exemplo da voz soprosa e da voz sussurrada.

---

<sup>7</sup> É considerada uma das línguas africanas mais importantes, pela extensão territorial em que é falada e pelos fins sociais a que serve.

O autor então realiza uma descrição de ajustes isolando alguns parâmetros de controle laríngeo que são relevantes para a discussão das diferentes configurações fonéticas. Com base na fisiologia laríngeo, o autor propõe então três parâmetros de tensão muscular que interagem com fatores aerodinâmicos do fluxo aéreo pulmonar e de pressão, são estes: tensão adutiva, compressão medial e tensão longitudinal. Ver Figura 6.

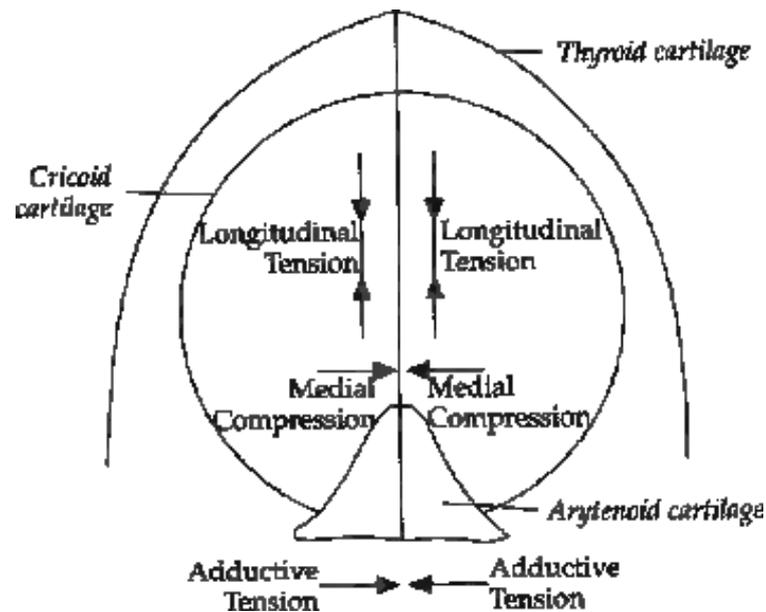


FIGURA 6. Parâmetros de controle laríngeo para a descrição das configurações articulatórias de acordo com Laver (1980).

A tensão adutiva pode ser definida como um processo gerado pela tensão dos músculos interaritenóides, refletindo a coaptação cartilaginosa da glote e do ligamento vocal. Já a compressão medial pode ser gerada pela pressão de compressão das cartilagens aritenóides a partir da contração dos músculos cricoaritenóides laterais, reforçada pela tensão lateral dos músculos interaritenóides. A compressão medial gera o fechamento do ligamento vocal, no entanto, para que a glote cartilaginosa também possa ser coaptada é necessário que ocorra a ativação da musculatura interaritenóidea. A tensão longitudinal, por sua vez, é promovida pela contração do músculo vocal e/ou cricotireoideo. Cada configuração dos diferentes tipos fonatórios apresentam diferentes especificações quanto a manipulação desses três parâmetros (LAVÉ, 1980).

A voz modal é, assim, produzida por uma tensão adutiva moderada e uma moderada tensão longitudinal, em condições nas quais a frequência fundamental

situa-se em uma região mais baixa na gama de F0, tomando como referência uma conversação normal (LAVÉR, 1980).

De acordo com Laver (1980), a descrição dos diferentes tipos fonatórios deve considerar também a possibilidade desses ajustes fonatórios ocorrerem potencialmente em combinação ou de forma isolada. A partir desse critério, o autor propõe três categorias. A primeira categoria é composta pelos modos de fonação modal e falsete. A natureza dessa categoria consiste na possibilidade desses tipos fonatórios ocorrerem separadamente ou em combinação com outros tipos de grupos distintos, mas não em combinação mútua.

A segunda categoria consiste nos tipos de fonação sussurrado e laringalizado. Estes podem ocorrer sozinhos e combinados, gerando uma laringalização sussurrada, por exemplo. Podem ocorrer ainda em combinação com os tipos fonatórios do primeiro grupo, resultando em voz sussurrada, falsete sussurrado, laringalização e sussurro laringalizado; como também, ocorrer em padrões combinados com o primeiro grupo e entre si: voz sussurrada laringalizada e falsete laringalizado sussurrado (LAVÉR, 1980).

A terceira categoria proposta por Laver (1980), consiste em modificações de configuração fonatória que apenas são possíveis de ocorrer em padrões combinados de tipos de fonação e nunca em tipos isolados. São estes: aspereza e soproidade.

O padrão áspero pode combinar-se com o tipos de fonação modal e falsete. O modo soproso pode apenas ser combinado com a voz modal, resultando numa qualidade de voz soproso. Tanto o primeiro como o segundo modo de fonação não são possíveis de serem combinados com a segunda categoria de tipos fonatórios descritos acima, exceto que se faça presente na configuração algum elemento da primeira categoria (LAVÉR, 1980).

Os preceitos apresentados por Ladefoged (1971) e as considerações realizadas por Laver (1980) permitiram com que Esling e Harris (2005) desenvolvessem uma descrição das configurações laríngeas a partir da manipulação de um sistema complexo de válvulas, guiados por uma análise objetiva a partir de correlatos fisiológicos dos estados da glote. O trabalho de Esling e Harris (2005) representa um avanço no que diz respeito à descrição articulatória dos modos de fonação, como será discutido adiante.

### 1.3. A Laringalização nas línguas naturais: correlatos acústicos, articulatórios e perceptuais

Ao longo do fluxo enunciativo, o falante implementa recursos prosódicos capazes de cumprir funções tanto de ordem pragmática como linguística. Essas funções são, geralmente, acompanhadas de mudanças nos parâmetros prosódicos da fala, repercutindo nas medidas de frequência fundamental, intensidade e duração.

A frequência fundamental é representada pelo correlato fisiológico da frequência de vibração das pregas vocais, – parâmetro físico que controla diretamente a sensação do *pitch*, ou em termos perceptuais, a sensação de grave e agudo. Os estudos fonéticos, que tratam da entoação, privilegiam a análise deste parâmetro ao longo dos enunciados (BARBOSA, 2012).

De acordo com Behlau (2001), a frequência fundamental é medida pela velocidade com a qual uma forma de onda se repete em uma unidade de tempo, o que é indicado em ciclos por segundo, na literatura mais antiga, ou em Hz, nos textos mais modernos. Nesta relação, 1 Hz corresponde a 1 c/s (ciclo por segundo).

Apesar da definição acústica e perceptiva de frequência fundamental e *pitch*, respectivamente, estes dois termos são usados indiscriminadamente em grande parte da literatura. Por outro lado, os termos 'entoação' e 'prosódia' também podem ser encontrados em um uso comum, embora, de forma geral, o termo "entoação" esteja conotado às características tonais (F0), enquanto que o termo prosódia, para além das características tonais, envolve tempo (duração) e dinâmica (nível de pressão sonora). Além disso, em um sentido mais amplo, a distribuição tonal local e global podem ser referidas como "prosódia" enquanto que, em sentido estrito, apenas a distribuição tonal global pode ser referida propriamente como "entoação". Assim, as características tonais intrinsecamente lexicais são área da prosódia (BOTINIS; GRANSTRÖM; MÖBIUS, 2001).

O volume, por sua vez, é o correlato perceptivo das variações de intensidade, que é seu parâmetro físico mais importante, enquanto a duração objetiva de unidades do tamanho da sílaba (medida em milissegundos), sendo o parâmetro de controle por excelência da sensação de duração (BARBOSA, 2012).

Tais parâmetros se relacionam com o acento, ritmo e entoação da fala (CRISTÓFARO SILVA, 2011), e são responsáveis por moldar a enunciação,

imprimindo ao que se fala um modo de falar, modo este dirigido intencionalmente ou não ao falante (BARBOSA, 2012).

Um determinado ajuste fonético pode contribuir para o estabelecimento de contraste numa língua ou apenas exercer funções de caráter prosódico. É o que ocorre, por exemplo, com a laringalização, fenômeno que se caracteriza em seu mecanismo de produção como um tipo de fonação não-modal, podendo estar relacionado a uma função prosódica específica, ao estilo do falante ou até mesmo estabelecer contraste em um sistema fonológico específico (REDI & SHATTUCK-HUFNAGEL, 2001; TELLES, 2013; LADEFOGED E GORDON, 2001).

Segundo Esling e Harris (2005) e Edmondson e Esling (2006), os estados das pregas vocais podem ser vistos a partir do funcionamento de seis diferentes válvulas. As possíveis combinações em seus mecanismos de funcionamento geram possibilidades variadas de ajustes fonatórios, dentre os quais inclui-se o *creak voice* ou o registro laringalizado.

Nestes estudos, a atividade laríngea foi monitorada a partir do exame de nasofibroscopia, com capturas de imagens em 30 quadros por segundo, o que possibilitou analisar a dinâmica laríngea e o engajamento das diferentes válvulas com bastante objetividade. Embora as válvulas sejam descritas de forma estática, suas combinações geram várias possibilidades de ajustes fonatórios. Abaixo segue uma descrição sintetizada dessas válvulas e seus mecanismos de atuação. Nas Figuras 7 e 8 é possível identificá-las.

- a) Válvula 1: adução e abdução glotal. Movimento lateral-para-centro (correspondente às funções de abertura e fechamento glótico);
- b) Válvula 2: cobertura e amortecimento parciais da adução glótica pelas pregas ventriculares e vibração das pregas vocais. Movimento lateral-para-centro (ajuste fonético denominado de incursão ventricular presente na laringalização);
- c) Válvula 3: compressão esfíntérica das aritenoides e avanço e elevação das pregas ariepiglóticas por meio do complexo muscular tiroaritenóideo. Movimento ântero-posterior (ajuste fonético presente na laringalização);
- d) Válvula 4: retração da língua e da epiglote, movimento para trás e para baixo, resultando, em casos extremos, no completo fechamento na parede da faringe, por meio do músculo hioglosso (constrição epigloto-faríngea);

- e) Válvula 5: elevação laríngea pelo grupo muscular suprahióideo a partir da contração dos músculos digástrico anterior e posterior, estilohióideo, genihióideo e hioglosso (correspondente à elevação laríngea);
- f) Válvula 6: constrição interna das paredes da faringe devido a ação esfintérea dos constritores superior/médio/inferior (correspondente ao estreitamento faríngeo).

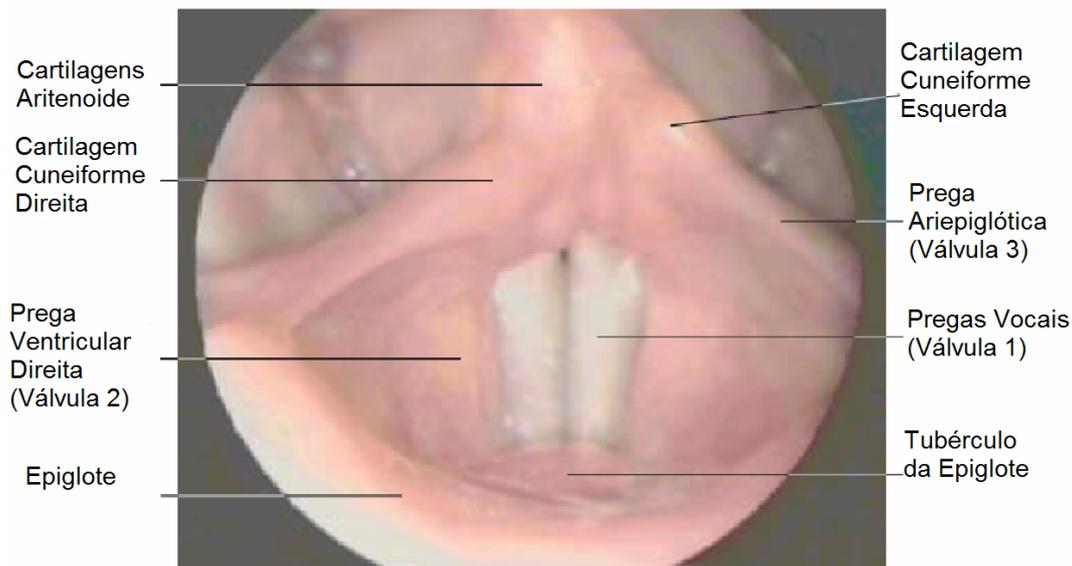


FIGURA 7: Representação do modo de fonação modal. Observa-se a coaptação completa das pregas vocais e a não adução (constrição) das pregas arieplóticas (válvula 3). Exemplo retirado do trabalho de Edmondson & Esling (2006).

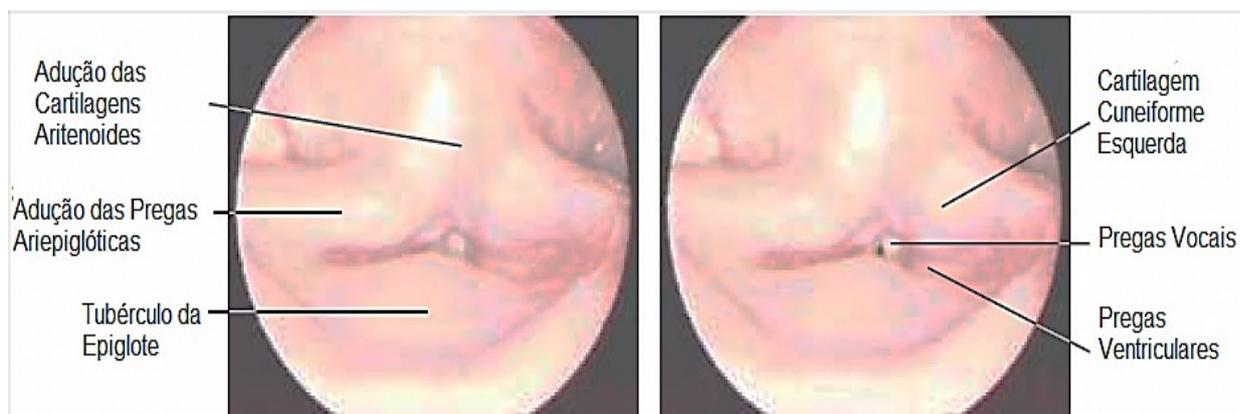


FIGURA 8: Representação do estado da glote durante a laringalização (creak voice). Observa-se a adução (constrição) das pregas arieplóticas, ou da terceira válvula; bem como, a adução das pregas aritenoides, repercutindo na movimentação ântero-posterior das pregas vocais. Exemplo retirado do trabalho de Edmondson & Esling (2006).

De forma sintetizada, fazendo a correspondência destas válvulas com os modos de fonação, têm-se na válvula 1 o registro modal, na válvula 2 a oclusiva glotal moderada, na válvula 2 ou na válvula 3 a laringalização, e na válvula 3 a oclusiva epiglotal <sup>8</sup>(EDMONDSON E ESLING, 2006). Ver Figura 9.

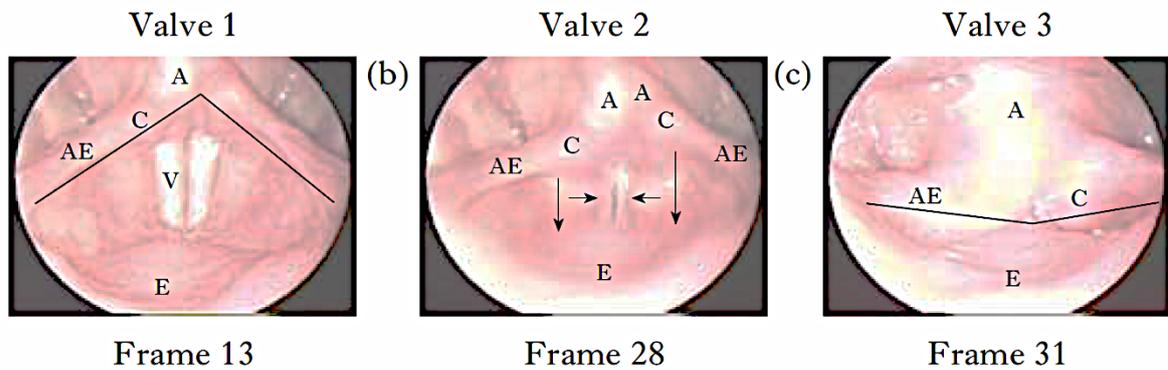


FIGURA 6: Representação dos estados da glote durante a ativação da válvula 1 (registro modal), válvula 2 (registro laringalizado) e válvula 3 (constricção laríngea) registrados por Edmondson & Esling (2006). A = aritenóides; AE = pregas ariepiglóticas; C = cartilagem cuneiforme; E = epiglote; V = pregas vocais (glote).

Para Edmondson e Esling (2006), o abaixamento da frequência fundamental somado ao relaxamento suficiente da glote para a vibração lenta e ondulante são consequência da constricção da terceira válvula, que implica na compressão esfíntérica das aritenoides e no avanço e elevação das pregas ariepiglóticas, por meio do complexo muscular tiroaritenóideo, refletindo numa movimentação ântero-posterior das pregas vocais. Essa descrição articulatória do mecanismo fonatório caracteriza a laringalização.

Ainda de acordo com Edmondson e Esling (2006), o engajamento isolado da válvula caracteriza o chamado 'esfíncter laríngeo', porém, muito frequentemente, o fechamento desta válvula é auxiliado pelas válvulas 1 e 2.

Hanson et al., (2001) e Lindblom (2009) observaram na laringalização um padrão vibratório irregular das pregas vocais, resultado da constricção laríngea, assim como a percepção de uma qualidade vocal soprosa, decorrente de um discreto *gap* na coaptação destas durante a fonação. Segundo Hanson et al., (2001), esse tipo de

<sup>8</sup> Neste segmento, a epiglote e a língua agem como articuladores passivos e a parede da faringe como articulador ativo. Também na oclusiva epiglotal em sua forma padrão (normal ou não-extrema) a epiglote é um articulador passivo. Quando as válvulas 1, 2, 3, 4 e 5 operam em sequência e em conjunto para fechar todo aspecto do tubo epilaríngeo, reduzindo o volume da parte baixa da faringe, ocorre a forma mais extrema de oclusiva epiglotal (LIMA-GREGIO, 2011).

registro não ocorre somente em falantes com distúrbios vocais, mas, está presente de forma muito frequente em sujeitos sem comprometimentos.

A laringalização está tipicamente associada a uma adução intensa das pregas vocais, porém com uma abertura ao longo de sua extensão suficiente para permitir a fonação. O resultado acústico dessa configuração laríngea é uma série de pulsos glotais irregularmente espaçados que dão a impressão auditiva de uma série de *taps* rápidos, como de uma vara ao ser deslizada ao longo de uma grade (LADEFOGED, 1971; LADEFOGED E GORDON, 2001).

A nomenclatura "fonação irregular" é utilizada como um termo guarda-chuva para um padrão de fonação que difere da gama normal de vibração, quase periódica, das pregas vocais. A fonação irregular é caracterizada por uma variação incomum no tempo ou amplitude ao longo de períodos adjacentes de frequência fundamental, ou um grande espaçamento entre os pulsos glotais. A fonação "soprosa" não está incluída nesta definição. Termos comumente usados para descrever fonação irregular incluem "laringalização", "voz crepitante" e "*vocal fry*" (SLIFKA, 2007).

Os correlatos acústicos da laringalização nas línguas naturais vêm sendo abordados por diversos autores, em sua maioria, de língua inglesa. Os estudos evidenciam o abaixamento da frequência fundamental, pulsos glotais irregulares e a queda de intensidade na transição do registro modal para o laringalizado (DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996; HANSON & CHUANG, 2001; REDI & SHATTUCK-HUFNAGEL, 2001; ESLING et al, 2005; LIMA-GREGIO, 2011).

De acordo com Segundo, Alves e Trinidad (2013) a fonação *creak* [laringalizada] constitui o registro mais baixo que a voz humana é capaz de produzir e corresponde ao registro basal. Nesta configuração, as dobras das pregas vocais estão relaxadas e encurtadas longitudinalmente, de modo que sua massa aumenta e vibra com uma frequência fundamental menor que a voz modal, geralmente abaixo de 70-80 Hz.

Dilley, Shattuck-Hufnagel & Ostendort (1996), Redi e Shattuck-Hufnagel (2001) e Belotel-Grenié e Grenié (2003) salientam que as características acústicas envolvidas na laringalização não são aleatórias, antes indicam funções prosódicas específicas, como será discutido na próxima seção.

Segundo Ishi (2004), dentre todos os parâmetros, a relação entre os dois primeiros picos de autocorrelação (NACR) verificou-se ser o principal parâmetro na discriminação entre os registros modal e o laringalizado. Já Garellek (2012) considera

a diminuição nos valores de H1-H2<sup>9</sup>, amplitude dos dois primeiros harmônicos da fala, como mais importante para a identificação do fenômeno. Para Lima-Gregio (2011) a diminuição da intensidade entre os trechos modal e laringalizado é o parâmetro fonético-acústico mais importante na caracterização do fenômeno.

Watt e Burns (2012) constataram em testes de percepção um alto nível de concordância entre ouvintes treinados e ouvintes inexperientes na identificação do registro laringalizado. De acordo com os autores, a laringalização está comumente relacionada com baixas taxas de elocução.

#### 1. 4. Laringalização: Motivações prosódicas

A fonação não-modal, especialmente a laringalização, é comumente empregada na fala habitual como um marcador de limites prosódicos, quer tanto no início e em final de fronteiras, como, por exemplo, no sueco, inglês (KREIMAN, 1982; DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996), finlandês, tcheco e no servo-croata (KREIMAN, 1982; DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996; LADEFOGED e GORDON, 2001).

Em muitas línguas, palavras iniciadas por vogal frequentemente apresentam laringalização, o que é comum no início de unidades prosódicas maiores em detrimento de unidades hierarquicamente inferiores, e mais frequente em sílabas acentuadas quando comparadas às não acentuadas (DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996).

Os ambientes fonológicos sensíveis a este tipo de registro são apresentados em consenso pela literatura: fronteiras prosódicas de frase entoacionais mediais e finais acompanhadas de pausa silenciosa são locais onde o fenômeno é recorrente (DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996; REDI & SHATTUCK-HUFNAGEL, 2001).

Neste sentido, Dilley, Shattuck-Hufnagel & Ostendorf (1996) observaram em um estudo com radialistas falantes do inglês americano, que frases entoacionais iniciadas por vogal propiciam a ocorrência da laringalização, especialmente, quando o item lexical é acentuado. Essa particularidade se deve à mudança abrupta no valor

---

<sup>9</sup> A amplitude do primeiro harmônico (H1) em relação ao segundo (H2) é usado como uma indicação do quociente de abertura, ou do inglês "*open quotient*" (OQ), a razão entre a fase aberta do ciclo glótico em relação ao período total (HANSON E CHUANG, 1999).

da frequência fundamental no item acentuado, situação na qual a ocorrência do fenômeno é potencializada.

De acordo com Dilley, Shattuck-Hufnagel & Ostendorf (1996), a presença de uma pausa anterior ou a marcação de limites frasais influenciaram diretamente a ocorrência do fenômeno. Segundo as autoras, a ocorrência da laringalização justificada pelo contexto segmental é um fator significativo apenas na ausência de uma fronteira prosódica.

O estudo constatou ainda que, embora o fenômeno ocorra tanto em fronteiras de frases entoacionais mediais e finais, são nestas últimas que se observa as maiores taxas de laringalização. Nesse sentido, Andrews (2009) afirma que o fenômeno também está relacionado com a diminuição do fluxo aéreo glotal durante a fala, referindo que em finais de frases longas, a maioria dos falantes realizam a transição do registro modal para o laringalizado.

Drugman et al.; (2013) avaliaram a possibilidade de utilizar fatores contextuais para prever a produção da laringalização. Os pesquisadores investigaram como a informação contextual está relacionada ao uso da voz não-modal laringalizada. De acordo com os autores, um conjunto de fatores contextuais ligados à produção da fala parecem ser altamente relevantes na previsão do fenômeno, como por exemplo, o silêncio ou a pausa.

Isso sugere que a laringalização tem entre outros papéis um papel sintático, delimitando grupos de palavras e, conseqüentemente, facilitando a segmentação de frases. Outros falantes e línguas podem utilizar a voz laringalizada de diferentes maneiras, no entanto, desde que existam os fatores contextuais necessários para caracterizar o uso, acredita-se que estas pistas continuam sendo eficazes (DRUGMAN et al.; 2013).

Redi e Shattuck-Hufnagel (2001) examinaram as taxas de laringalização em frases entoacionais finais comparadas a taxas de laringalização em frases entoacionais intermediárias para seis falantes do inglês americano. As taxas de laringalização nos limites de frases entoacionais finais foram consideravelmente maiores quando comparadas às taxas de laringalização em extremidades de frases entoacionais intermediárias para todos os seis sujeitos avaliados.

Ao avaliar indivíduos falantes do PB com e sem distúrbios articulatórios, Lima-Gregio (2011) constatou que falantes sem comprometimentos produziram laringalizações em ambientes linguísticos previsíveis e específicos. Por sua vez,

indivíduos com distúrbios articulatórios, como aqueles decorrentes da fissura lábio-palatina, apresentaram laringalizações em ambientes linguísticos assistemáticos, como forma de compensar suas dificuldades articulatórias.

Os resultados obtidos por meio da análise acústica revelaram que, além da produção de oclusiva glotal para as consoantes oclusivas, havia uma alta ocorrência de laringalização nas vogais em falantes do PB com alterações articulatórias (LIMA-GREGIO, 2011).

A oclusiva glotal /ʔ/não é um segmento contrastivo no PB, o que significa dizer que a sua produção ou omissão não interfere no significado das palavras. A sua realização é estilística no PB, assim como também é no inglês, desempenhando nessas línguas o papel de alofone.

Garellek (2012) refere que, no Inglês americano, a oclusiva glotal pode ocorrer em três posições fonológicas. Em primeiro lugar, elas ocorrem opcionalmente antes de palavras iniciadas por vogal (glotalização de palavra inicial), de tal forma que “apple” em inglês pode ser pronunciada com ou sem a oclusiva glotal em lugar de /p/, [æpɪ] e [æʔɪ]. Em um segundo contexto, a oclusiva glotal também pode ocorrer durante o reforço glotal opcional (ESLING ET AL., 2005), no qual oclusivas na posição de coda – na maioria das vezes *unreleased* - são produzidas com adução simultânea das pregas vocais. Por exemplo, “cat” pode ser pronunciado com a oclusiva glotal na posição da coda [k<sup>h</sup>æʔt̚]. Por último, a oclusiva glotal pode ocorrer como um alofone de /t/ pós-tônica antes de nasais silábicas, de tal forma que “button” geralmente é pronunciado [ˈbʌʔn̩].

Segundo Wise (1958), no inglês o fenômeno é capaz de contribuir na desambiguação de expressões, como por exemplo, nos casos de “some mice” [səmˈmaɪs] e “some ice” [səmˈaɪs], nos quais, o uso da laringalização no segundo exemplo, auxilia na diferenciação de sentido.

Ainda de acordo com a autora, outro uso muito comum da oclusiva glotal é em separar vogais sucessivas entre as quais não existe glide natural para ligar o hiato, como em “Ana asks” [ˈænə ˈæsk̚s]. A laringalização na vogal inicial de “asks” desempenha esta função. Segundo a autora, a ocorrência da laringalização para separar segmentos vocálicos no inglês é ordenada e severamente dependente das propriedades fonéticas das vogais envolvidas.

Garellek (2011) sugere que a laringalização é perceptivelmente útil, a medida em que os ouvintes podem utilizá-la para interpretar itens lexicais confusos com mais precisão. Por itens lexicais perceptualmente confusos o autor se refere àqueles com uma frequência fundamental relativa baixa em relação aos itens da vizinhança fonológica, e não a itens lexicais essencialmente confusos na língua.

Em um estudo de natureza perceptual, o autor buscou verificar se palavras consideradas mais "confusas", a partir do critério referido, tendem a apresentar maior índice de laringalização do que palavras com maior precisão perceptual. As palavras-alvo foram 20 palavras monossílabas do inglês com estrutura CVC, incluindo as vogais /a, æ/. As obstruintes /p, t, k, b, d, g/ no onset e /p, t/ na coda. Estes tipos de codas foram escolhidos porque, de acordo com o autor, observa-se no inglês a ocorrência frequente da laringalização na vogal anterior na presença desses segmentos.

Este estudo forneceu evidências de que a laringalização em vogais do inglês, antecedendo oclusivas na posição de coda, é mais frequente em palavras ditas "confusas". Palavras com baixa frequência fundamental relativa apresentaram menores valores de H1 - H2, bem como, menor proporção harmônico-ruído quando comparadas a palavras de frequências relativas mais elevadas.

No teste de percepção desenvolvido por Garellek (2011), verificou-se que a laringalização pode auxiliar o ouvinte a resgatar pistas segmentais perdidas pela não precisão na produção, neste caso, as palavras ditas "imprecisas" apresentaram maiores índices de laringalização.

De acordo com Garellek (2011) o fato de a laringalização poder facilitar o reconhecimento da coda na língua inglesa sugere que os falantes aumentem o índice de laringalização vocálica em palavras com menor precisão articulatória com o intuito de auxiliar o ouvinte na percepção destas unidades.

Segundo o autor a ocorrência de oclusivas glotais, antes de vogais em início de palavra, é em grande parte dependente da prosódia. Esses segmentos são mais comuns e mais propensos de serem realizados quando a vogal seguinte é proeminente e posicionada no início da frase.

Ainda de acordo com o autor, embora o domínio frasal e a proeminência tenham influências distintas sobre as vogais no início de palavras, a proeminência parece estar associada ao movimento glotal, independentemente do domínio frasal

em que a palavra possa estar localizada. Por outro lado, o domínio prosódico está associado principalmente com a força do gesto glotal, de tal modo que uma oclusiva glotal é mais suscetível de ser realizada com maior constrição em início de frases e de enunciados do que em frases mediais.

Segundo Garellek (2012), todos os segmentos vozeados em frase-inicial tendem a ter sonoridade irregular. Assim, a presença de irregularidade acústica não pode ser tomada como indicação de um gesto glotal. Presume-se que a indicação mais clara da presença de um segmento laringalizado é a diminuição nos valores de H1-H2, o que não é observado em todas as vogais irregulares, mas geralmente apenas nas proeminentes.

Slifka (2006) atribui uma razão fisiológica para a ocorrência da laringalização em final de oração, especialmente no final de um grupo de respiração ou enunciado. O autor refere que o fenômeno pode ser desencadeado por baixa pressão subglotal, que também é reduzido em início de grupos de respiração ou início de frases. Assim, a menos que a rigidez das pregas vocais e/ou adução aumente, a baixa pressão subglótica no início e final de enunciados poderia induzir a ocorrência de uma fonação irregular.

### 1.5. Relação entre taxa de laringalização e sexo

Embora muitos trabalhos tenham analisado a influência do sexo na taxa de laringalização, ainda não há consenso na literatura quanto a esse aspecto (REDI & SHATTUCK-HUFNAGEL, 2001).

Dilley, Shattuck-Hufnagel & Ostendorf (1996) observaram que mulheres leitoras de notícias de rádio no inglês americano em seu estudo realizaram laringalização de vogais em início de palavra com taxas mais elevadas (40%, 44% e 38%) do que em falantes do sexo masculino (24% e 13%).

Nesta mesma perspectiva, Redi e Shattuck-Hufnagel (2001) buscaram avaliar a ocorrência do fenômeno laríngeo em locutores profissionais, apresentadores de rádio, e locutores não profissionais. O material de coleta utilizado nesta pesquisa consistiu na leitura de um texto com repertório segmental controlado, proveniente do corpus Labnews, composto por 6 locutores profissionais, três homens e três mulheres, e a leitura de sentenças referentes ao corpus ABC, formado por 4 falantes não profissionais, dois homens e duas mulheres.

As autoras observaram que locutores profissionais pertencentes ao grupo do sexo feminino laringalizaram com uma frequência significativamente maior do que os falantes do sexo masculino. No entanto, entre os oradores não-profissionais, os resultados foram controversos. Um sujeito do sexo masculino laringalizou com uma taxa que foi significativamente maior do que qualquer outro falante, apresentando um comportamento desviante em relação ao grupo.

## 1.6. Implicações forenses

Partindo, inicialmente, de uma questão metodológica, de forma a esclarecer as terminologias utilizadas nos estudos de fonética forense, é necessário a descrição das nomenclaturas empregadas. Neste sentido, Hollien (1990) refere dois processos distintos envolvidos na tarefa de reconhecimento de falante, a identificação e a verificação de locutor.

A identificação do falante, especificamente, é definida como o processo de identificação de um locutor desconhecido a partir de amostras de sua fala/voz. Para realizar esta tarefa precisa-se ter alguma familiaridade com as características da fala deste sujeito (uma condição rara) ou ser capaz de relacionar a voz desconhecida com a voz de algum indivíduo específico ou de locutores que integram um grupo de suspeitos. Este processo é um tanto complexo. É particularmente árduo porque a tarefa deve ser realizada mesmo com a presença de distorções no sistema (telefone, filtros de passa banda, ruído, etc.) ou complexidades que dizem respeito ao falante (nível de estresse, disfarce vocal, estado emocional, etc.). O discurso aqui nunca é transparente e o assunto é geralmente pouco definido, especialmente se houver um crime em questão. Muitas vezes, uma associação não é possível; quer o desconhecido não corresponda com nenhum falante entre os suspeitos ou as gravações de voz/fala desconhecidas são simplesmente incipientes para a análise (HOLLIEN, 1990).

Por outro lado, quando o que está em questão é a verificação do falante, o sujeito quer ser reconhecido. Neste contexto, equipamentos de alta qualidade são utilizados, e muitas amostras de referência da fala do locutor estão disponíveis e podem ser continuamente acessadas. A verificação de locutor é um dos vários métodos utilizados para autenticar a identidade vocal do sujeito, quando este não

pode ser identificado por imagem. Este procedimento também é (ou pode ser) usado em operações bancárias por telefone, comunicação com pessoas em cápsulas espaciais ou com sujeitos em locais remotos. Este é o procedimento utilizado na tentativa de identificar amigos, conhecidos ou familiares (HOLLIEN, 1990).

A fonética forense é a aplicação do conhecimento fonético geral a problemas legais, como por exemplo, para a identificação ou caracterização de um falante. Outras possíveis tarefas que um foneticista forense pode realizar incluem o delineamento de um conjunto de vozes, a elaboração do perfil vocal de um locutor ou atividades relacionadas ao aprimoramento da análise de fala, enunciados, etc. Nesse caso, o perito deve comparar uma ou várias amostras de fala de um locutor desconhecido (o ofensor) com uma ou várias amostras de origem conhecida (os suspeitos) (SEGUNDO, ALVES E TRINIDAD, 2013).

De acordo com Morrison (2013) o principal objetivo desta comparação é responder à questão: "quão semelhantes são as propriedades observadas nas amostras de fala conhecida e questionada, sob a hipótese de que a amostra de fala questionada tenha a mesma origem que a amostra conhecida em contraposição à hipótese de que estas tenham uma origem diferente?".

Segundo Rose (2012), além da identificação forense de um locutor, a fonética forense inclui áreas como a descrição do perfil de falantes (na ausência de um suspeito, dizendo algo sobre o sotaque ou características socioeconômicas interpretados por meio da voz/fala do (s) infractor (es)); a identificação de conteúdo (determinando o que foi dito quando as gravações são de má qualidade, até mesmo quando a voz é patológica ou a fala apresenta algum sotaque estrangeiro); e a autenticação de gravações (determinando se uma fita foi adulterada ou não).

Sabe-se que a singularidade fonatória de um falante inclui uma variedade de formas de se pronunciar segmentos vocálicos e consonantais, características individuais de ritmo e durações relativas, de padrões entoacionais e gama de frequência fundamental, tão como, variações na qualidade da voz. No entanto, contrastes entre os sons da fala são limitados pelo âmbito de aplicação das variações no sistema fonológico, enquanto que a qualidade de voz ou timbre pode apresentar altos graus de variação (HIRSON E DUCKWORTH, 1993).

Neste sentido, o fenômeno da laringalização tem sido referido como uma pista em potencial para a identificação de falantes, uma vez que se observa consistência

quanto às particularidades de sua ocorrência entre diferentes locutores (FRENCH, 1994).

Redi e Shattuck-Hufnagel (2001) observaram no inglês americano ampla variação na taxa de laringalização entre indivíduos sem comprometimentos articulatórios durante a leitura, evidenciando na maioria dos casos diferenças estatisticamente significativas.

Em estudo de natureza forense, realizado com falantes do inglês americano, Moosmüller (2001) observou que, ao realizar disfarce vocal, mulheres tenderam a agravar suas vozes de forma a se aproximar de características vocais masculinas, o que implicou na diminuição da frequência fundamental e na maior ocorrência de laringalizações. Nessa perspectiva, Künzel (2000) afirma que existe uma estreita relação entre a frequência fundamental e o tipo de disfarce vocal adotado pelos sujeitos. O autor constatou que indivíduos com frequência fundamental acima da média tenderam a elevá-la ainda mais na execução do disfarce, enquanto aqueles com frequência fundamental abaixo da média registrada, agravaram mais a voz, o que resultou em altas taxas de laringalizações.

Segundo Mathur, Choudhary e Vyas (2016) o disfarce vocal pode causar modificações extremas nos parâmetros acústicos e fonéticos da fala, como na qualidade vocal, no fluxo discursivo, no grau de fonação, no padrão entoacional, na taxa e na intensidade dinâmica da fala. Igualmente, pode desencadear variações nos valores de frequência fundamental da voz.

No experimento visando verificar qual o tipo de informação linguística é preservada quando o sujeito tenta imitar a fala do outro, Cole e Shattuck-Hufnagel (2011) observaram que os locutores raramente não realizam um *pitch accent*<sup>10</sup> e/ou uma fronteira de sintagma entoacional na fala imitada. Todavia, as mudanças na natureza das pistas fonéticas são frequentes, assim como na duração das pausas e na ocorrência de períodos glotais irregulares, associados a fronteiras e *pitch accents* no inglês americano. Os resultados podem sugerir que, mesmo em situações de imitação, seja pouco provável que os falantes laringalizem sob a mesma taxa.

---

<sup>10</sup> Os *pitch accents* são eventos locais que assinalam proeminência no nível do enunciado, e estão associados na teoria Métrica Autossegmental à altura da curva entoacional nos pontos em que esses eventos ocorrem. A literatura sobre o assunto, predominantemente em inglês, destaca a diferença entre *pitch accent* e *stress*, ou seja, entre proeminência melódica e acento lexical. Diferentemente da proeminência, o acento é uma propriedade lexical de sílabas individuais, assinalada no nível da palavra (LUCENTE, 2012).

A necessidade de um maior controle de variáveis, como o sexo, idade, grau de escolaridade, de forma a garantir uma amostra relativamente homogênea, objetiva recriar condições mais próximas a de uma análise forense das características da fala, motivou o desenho metodológico deste trabalho.

É comum que estes demarcadores sociais sofram pouca variação num contexto de análise forense, seja porque o indivíduo é comparado consigo mesmo ou com alguém que lhe assemelha. Acredita-se que a opção por uma população homogênea permita minimizar este aspecto.

Buscou-se, desta forma, verificar de que forma a ocorrência da laringalização poderia indicar um comportamento singular daquele que a produz, à medida em que uma maior ou menor taxa de ocorrência do fenômeno caracterizaria um perfil do locutor ou um estilo de fala próprio.

## 2. MÉTODO

### 2.1 Participantes

Para este trabalho foram formados dois grupos com cinco participantes cada, um grupo do sexo masculino (M) e outro grupo do sexo feminino (F). Os indivíduos foram selecionados por demanda espontânea e possuíam algum grau de familiaridade com o pesquisador, participando do seu círculo de convívio.

Um dos critérios para a composição dos grupos era que os sujeitos recrutados não apresentassem idades muito díspares, de forma a assegurar uma relativa homogeneidade na composição, evitando-se que uma alta variação nas faixas etárias pudesse comprometer a tarefa de comparação entre sujeitos e conseqüentemente a consistência dos resultados. Buscou-se evitar também que variações dialetais pudessem interferir no comportamento linguístico e assim inviabilizar as comparações entre os indivíduos.

Tanto os sujeitos do sexo masculino, M1, M2, M3, M4, M5, quanto os sujeitos do sexo feminino, F1, F2, F3, F4, F5, residiam na cidade de Maceió-AL e/ou em uma cidade circunvizinha, Campo Alegre-AL, localizada a aproximadamente 84 km da capital. Estes apresentavam idades que variaram entre 20 e 26 anos, todos com formação acadêmica universitária ou em formação.

Os seguintes critérios de inclusão foram estabelecidos:

- Idade entre 20 e 30 anos;
- Brasileiros, falantes do mesmo dialeto, residentes em Maceió-AL e em cidades circunvizinhas;
- Ensino médio fundamental completo ou superior;

Critérios de exclusão:

- Ter residido em outro país que não o Brasil por um período superior a 12 meses, nos últimos cinco anos;
- Histórico de patologia de fala ou de voz;
- Queixa auditiva e/ou obstrução nasal;

- Histórico de qualquer cirurgia recente (último mês antes da coleta);

Os indivíduos que não preencheram os critérios de inclusão e exclusão citados acima foram conseqüentemente afastados da pesquisa.

A busca por uma população mais homogênea parte da necessidade de reproduzir condições mais próximas a um contexto de pesquisa forense, em que os sujeitos podem compartilhar similaridades no que diz respeito aos demarcadores sociais e linguísticos.

De acordo com dados do Ministério da Justiça do governo federal, o perfil socioeconômico dos detentos mostra que 55% têm entre 18 e 29 anos, 61,6% são negros e 75,08% têm até o ensino fundamental completo. Em 2004, a taxa brasileira era de 135 presos por 100 mil habitantes<sup>11</sup>. Se considerada apenas a taxa de encarceramento feminino, saltou de 13,58 em 2005 para 32,25 detentas por 100 mil habitantes em 2014, o que afirma a importância de se realizar estudos na área de fonética forense que incluam não apenas sujeitos do sexo masculino, como é bastante comum, mas considere também a população feminina, enquanto um grupo em potencial para análise.

## 2.2. Composição do Corpus

O corpus desta pesquisa consistiu na gravação da recontagem de uma história intitulada "*Pear Film*", um curta metragem de cerca de seis minutos produzido pela Universidade da Califórnia em Berkeley (1975), o qual já foi utilizado por diversos trabalhos da área.

O vídeo contém uma sequência de atos, marcados pela passagem do tempo e mudanças de cenário. Os atores interagem entre si através de gestos e manipulação dos elementos cênicos que compõem o enredo. A ausência de diálogo permite que os telespectadores interpretem as ações e nomeiem os elementos apresentados de acordo com suas próprias percepções.

As gravações dos dez participantes foram realizadas utilizando-se um microfone condensador (beyerdynamic), visando uma captura mais ideal do sinal

---

<sup>11</sup> Dados disponibilizados pelo Ministério da Justiça em 2014, por meio do Relatório descritivo e analítico produzido através do Termo de Parceria nº 817052/2015, firmado entre o Departamento Penitenciário Nacional, a Secretaria Nacional de Segurança Pública e o Fórum Brasileiro de Segurança Pública.

acústico, em placa de áudio externa (USBPre 2 - Sound Devices) a fim de evitar que o ruído interno do computador interferisse na qualidade da gravação. A sala escolhida para as gravações ofereceu condições acústicas necessárias para a análise dos parâmetros de interesse. A opção por não gravar os participantes em cabine acústica, partiu da intenção de buscar garantir um ambiente de coleta menos formal e o mais natural possível. Apenas um pesquisador esteve responsável por instruir e acompanhar os procedimentos da coleta desde o início do processo, minimizando o efeito provocado pela presença dos pesquisadores no comportamento dos sujeitos durante a coleta.

Os sujeitos realizaram a recontagem da história com diferentes durações totais de narração. As gravações foram editadas, de forma que o tempo de análise da fala de cada sujeito fosse o mais isométrico quanto possível, respeitando-se os limites dos constituintes prosódicos. As gravações editadas apresentaram uma média de 1min e 27s, e desvio padrão de 0,05s.

O principal critério para edição dos áudios consistiu em evitar inserir na análise ambos os inícios e terminos das gravações, selecionando trechos centrais do tempo total da gravação, de forma que possíveis variações na taxa de elocução, tanto no início como no final das narrações, influenciassem as medidas de laringalização.

## 2.3. Análise dos dados

### 2.3.1. Análise da laringalização

Embora o fenômeno da laringalização tenha sido analisado tomando como referência unidades prosódicas, a exemplo dos grupos acentuais<sup>12</sup>, a identificação do fenômeno se deu a partir da análise de oitiva de cada segmento produzido no fluxo da fala, segmentos vocálicos e consonantais, adequadamente identificados no *textgrid* do *software* Praat, com confirmação através da análise da forma de onda e do espectrograma<sup>13</sup>(ver Figura 10), a partir de metodologia já utilizada em trabalhos

---

<sup>12</sup> De acordo com Barbosa (2006) os grupos acentuais são unidades delimitadas por dois acentos frasais consecutivos. Os acentos frasais não se referem a aspectos sintáticos, mas, sim a aspectos fonéticos, pois são definidos como proeminências no domínio da produção da fala, assinaladas por picos locais de duração ao longo dos enunciados (Barbosa, 2006).

<sup>13</sup> A análise espectrográfica acústica mede a acústica da onda sonora produzida pela voz (fala), detectando sutilezas no sinal acústico. O registro dessa análise é o espectrograma, um gráfico tridimensional. Atualmente a análise espectrográfica é amplamente utilizada em estudos linguísticos e na identificação de falante (BEHLAU, 2001; KENT E READ, 2007).

como o de Lima-Gregio (2011) Dilley, Shattuck-Hufnagel & Ostendorf, (1996) e Redi e Shattuck-Hufnagel (2001).

Um linguista, com experiência em fonética e um fonoaudiólogo realizaram o julgamento auditivo dos trechos laringalizados. Para análise acústica foi utilizado o *software* Praat, que permitiu a segmentação e análise do conteúdo linguístico.

O diferencial metodológico deste trabalho, assim como o de Lima-Gregio (2011) em relação aos demais estudos na área, diz respeito ao uso de um *script* computacional específico que avalia o fenômeno da laringalização extraindo parâmetros que são relevantes para a sua análise, como frequência fundamental, intensidade, *jitter* (índice de perturbação da frequência fundamental), *shimmer* (índice de perturbação da intensidade), proporção harmônico ruído e a relação entre os dois primeiros harmônicos H1-H2.

O jitter indica a variabilidade da frequência fundamental a curto prazo, medida entre ciclos glóticos vizinhos. Expressa o quanto um período é diferente do anterior ou de seu sucessor imediato, não levando em conta as alterações voluntárias da frequência fundamental. Se o sistema fonatório fosse um mecanismo perfeitamente estável não haveria diferenças nos períodos fundamentais, e, desta forma, o jitter seria zero (BEHLAU, 2001).

Já o conceito de shimmer pode ser explicado por analogia ao seu "companheiro de perturbação", o jitter. O sinal eletrônico não apresenta variações em sua amplitude, porém o sinal sonoro humano possui variações mínimas no controle da saída de ar e na intensidade da emissão, por mais estável e sustentada que seja uma vogal. Assim sendo, de modo análogo à perturbação da frequência fundamental, o shimmer indica a variabilidade da amplitude da onda sonora a curto prazo, representando uma medida de estabilidade fonatória. Assim, a medida de shimmer representa as alterações irregulares na amplitude dos ciclos glóticos, de um ciclo a outro (BEHLAU, 2001).

A proporção harmônico-ruído, conhecida em inglês como Harmonic-to-noise ratio (HNR), contrasta o sinal regular das pregas vocais com o sinal irregular destas e do trato vocal, gerando um índice que relaciona o componente harmônico versus o componente de ruído da onda acústica. Quanto mais elevada a frequência fundamental da voz, maior o componente harmônico da onda acústica para indivíduos normais. O componente ruído é tão maior quanto menor os harmônicos da emissão ou a medida de HNR (BEHLAU, 2001).

De acordo com Behlau (2001), a frequência fundamental tem uma distribuição média de 80 Hz a 250 Hz, nos adultos jovens, sendo que nos homens a faixa de frequências varia de 80 Hz a 150 Hz, e de 150 Hz a 250 Hz nas mulheres.

As medidas relativas de jitter, por sua vez, são expressas em porcentagem e o valor limite de normalidade é de até 0.5%. No caso do shimmer, o valor limite em porcentagem é de até 3%. Já para a medida de proporção harmônico ruído (HNR), Rodrigues (1993 apud Behlau, 2001) verificou valores médios de 9.4 dB para mulheres e 8.6 dB para homens. De acordo com Behlau (2001), quanto mais elevada a frequência fundamental maior o componente harmônico da onda acústica para indivíduos sem patologia vocal.

Segundo Andrews (2009), os valores de intensidade médios para a fala modal encadeada vão de 75 dB a 80 dB, com um desvio padrão de 5,4 dB. No que diz respeito a medida de H1-H2, Queiroz (2012) verificou em seu estudo uma média de -1.46 dB, com desvio padrão de 5.036 para o sexo masculino, enquanto que para o sexo feminino a média foi de 0.69 dB, e desvio padrão de 7.009 dB.

A análise dos parâmetros acústicos extraídos permitiu verificar de que forma estes elementos se comportam na transição do registro modal para o laringalizado, de modo a forecer evidências para a caracterização fonético-acústica do fenômeno no PB, como também permitir a correlação com resultados obtidos por outros trabalhos.

Os dados de fala foram segmentados e organizados em sete camadas no software Praat, como descrito a seguir (Ver Figura 10).

1. CV (para marcação dos segmentos vocálicos e consonantais);
2. CVglottals (para a classificação dos segmentos: os trechos com laringalização foram marcados com “/l” em contraposição o “/m” designou os trechos caracterizados pelo registro modal<sup>14</sup>;
3. H1H2 (para marcação dos trechos em que serão extraídas a medida H1 - H2);
4. Laring (identificação de laringalizações em fronteira de grupo acentual – LF ou fora de fronteira – LNF);
5. Transcrição ortográfica;

---

<sup>14</sup> Os segmentos não vozeados foram classificados com o rótulo “/m”, referente a segmento modal, a fim de atender à formalidades exigidas pelo script “GlottalizationAnalysis”.

6. “VowelOnset” (para marcar os intervalos VV<sup>15</sup>, ou seja, intervalos entre os consecutivos onsets vocálicos);

7. StressGroups (segmentação de grupos acentuais).

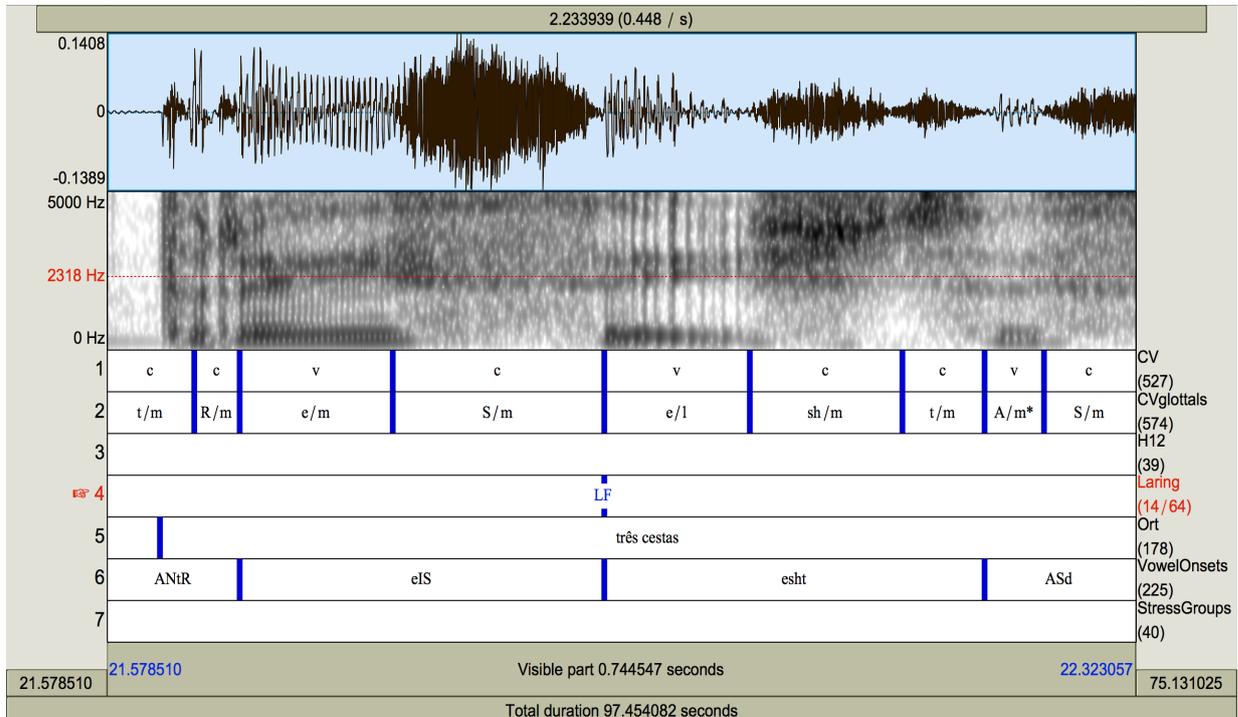


FIGURA 7: Disposição das camadas de segmentação no software Praat, de acordo com nível de análise (segmental e prosódico).

Nos trechos em que houve dúvida quanto à presença ou ausência da laringalização, optou-se pela classificação "modal", a fim de evitar que trechos dúbios comprometessem o cálculo das médias dos parâmetros extraídos.

As medidas de duração da laringalização, frequência fundamental, *jitter*, *shimmer*, intensidade e H1 - H2 foram extraídas automaticamente. A duração da laringalização diz respeito ao tempo de produção da vogal (em milissegundos-ms), compreendido entre o pico do primeiro e o pico do último pulso instável. A frequência fundamental foi calculada por meio do próprio algoritmo do *software* Praat.

O *jitter*, por sua vez, foi calculado por meio da técnica *Cross Correlation*, a partir detecção dos pulsos glotais. Por sua vez, a queda de intensidade foi calculada a partir da média na porção modal em relação à média na porção laringalizada.

<sup>15</sup> Unidade constituída por uma vogal e todos os segmentos consonantais que a seguem, independentemente de fronteira silábica, até o onset da vogal seguinte, vogal essa que determina o início da próxima unidade VV (BARBOSA, 2006).

O shimmer foi calculado pela diferença absoluta média entre as amplitudes dos períodos consecutivos, dividido pela amplitude média.

Para o cálculo de HNR, o algoritmo realiza uma detecção da periodicidade acústica do sinal com base em um método preciso de autocorrelação.

Já a medida de H1 - H2 é foi calculada tomando-se automaticamente, separadamente, a amplitude do primeiro harmônico (H1) e à do segundo (H2) em dB e as subtraindo. Essa amplitude é tomada extraindo-se o espectro de um trecho do sinal de fala, dentro da janela especificada de 25 ms centrada no meio exato da vogal.

A taxa de laringalização vocálica foi calculada por meio da divisão do número de vogais laringalizadas pelo total de vogais que cada indivíduo produziu, fórmula também reproduzida para análise da laringalização em consoantes e para a taxa de laringalização total, tendo-se realizado a divisão de consoantes laringalizadas pelo total de consoantes (intra-sujeito) e total de segmentos laringalizados pelo total de segmentos (intra-sujeito).

Posteriormente, executou-se a comparação das taxas entre-sujeitos a fim de verificar qual medida apresentaria maior variação entre diferentes locutores. Avaliou-se, ainda, o grau de incidência da laringalização em diferentes tipos de segmentos no PB: vogais e consoantes.

A análise estatística, teste paramétrico de proporção, foi realizada para a comparação intra-sujeitos e entre-sujeitos. Na comparação intra-sujeitos, foi observada a distribuição da frequência de laringalização em segmentos vocálicos e consonantais, que se refere à frequência bruta do número de laringalizações em contextos de fronteira e não fronteira. Esta análise se dá por meio da marcação dos contextos glotais na camada "CVglottals" do TextGrid, permitindo avaliar os ambientes de maior ocorrência do fenômeno.

A análise acústica permitiu avaliar quais parâmetros foram mais influenciados pela ocorrência do fenômeno, a saber: a queda de  $f_0$  (em Hz), porcentagem de diferença de jitter e *shimmer*, proporção harmônico-ruído (HNR), queda da intensidade (dB) e repercussão na medida de H1 - H2, para efeito de uma caracterização fonético-acústica do fenômeno.

Na comparação entre-sujeitos, foi avaliado o comportamento dos grupos no que diz respeito às taxas de laringalização de vogais, consoantes e laringalização total, de forma a verificar qual dessas medidas forneceria evidências mais consistentes para a tarefa de comparação de locutor, com base na gama de variação.

Um teste de proporção foi realizado para os dois grupos, homens e mulheres, com o objetivo de avaliar se as laringalizações coincidiam com as marcações de fronteiras de grupos acentuais.

### 2.3.2. Scripts de análise

Os *scripts* são textos que compreendem um menu de comando e ações de comando. Desta forma, quando se aplica um *script* a um programa, como o Praat, seus comandos são executados automaticamente (BOERSMA E WEENINK, 2014).

Dois *scripts* desenvolvidos por Barbosa (2008, 2004) foram utilizados para gerar automaticamente as medidas acústicas associadas à laringalização e segmentar a fala em unidades de grupos acentuais, a partir da segmentação manual prévia: o “GlottalizationAnalysis” e o “SGdetector” (Anexo A e B).

O *script* “GlottalizationAnalysis” permite computar as modificações acústicas nos segmentos vocálicos laringalizados. Ele gera um arquivo de texto contendo para cada vogal: i) a duração em milissegundos (ms); ii) a porcentagem de duração do creaky voice (laringalização) relativa à duração total da vogal; iii) a porcentagem de queda de frequência fundamental da parte modal à parte laringalizada; iv) a porcentagem de diferença de *jitter*, da parte modal à parte laringalizada; v) a porcentagem de diferença de *shimmer*, da parte modal à parte laringalizada; vi) a porcentagem de diferença de H1 - H2, da parte modal à parte laringalizada; vii) a porcentagem de diferença de HNR, da parte modal à parte laringalizada; viii) e a queda de intensidade (dB) da parte modal à parte laringalizada. O arquivo gerado pelo *script* indica ainda o tipo da vogal (modal ou laringalizada) e a posição da laringalização, levando em conta se está à direita, à esquerda ou ambos os lados da vogal.

O segundo *script* “SGdetector” detecta fronteiras prosódicas. A saída desse *script* são dois arquivos de texto (SG.txt e Dur.txt) contendo as seguintes informações:

1. SG.txt: i) a duração do grupo acentual, isto é, a duração entre dois picos locais de duração normalizada consecutivos; ii) o número de unidades VV no grupo acentual;

2. Dur.txt: para cada unidade VV: i) a duração da unidade VV (ms); ii) um número para indicar se a unidade VV está em posição de pré-fronteira (1) ou não (0); iii) valores brutos de z-score; iv) valores normalizados de z-score.

A segmentação em unidades VV permite que o *script* identifique os acentos frasais. Os picos locais de duração que assinalam os acentos frasais são determinados pelos picos locais de duração normalizada e suavizada das unidades V-V.

Com base nos preceitos do modelo dinâmico da fala (BARBOSA, 2006), a execução de tarefas requer como entrada um controle prosódico, que age nos domínios prosódicos de frase fonológica, frase entoacional e no enunciado fonológico, modificando o grau de coarticulação dos gestos articulatórios, exigindo, dessa forma, um elemento que organize temporalmente os gestos das vogais e consoantes.

Barbosa (2007) então propõe a unidade VV como a unidade mínima da programação rítmica da fala, a partir da qual se organiza o grupo acentual. Essa unidade é do tamanho da sílaba, contemplando a sequência de intervalos que correspondem ao onset da vogal até o onset da vogal imediatamente seguinte, incluindo todas as consoantes e glides entre ambas. O uso desta unidade fonética fundamenta-se em trabalhos que mostraram a importância da transição CV para a produção e percepção da fala (POMPINO-MARSCHALL, 1991; BARBOSA, 1996).

Devido à necessidade de manter estável a sequência de transição CV (consoante - vogal), há uma tendência de preservação da duração das unidades entre as transições (unidades VV), com o intuito de manter a estrutura rítmica e entoacional do enunciado, essas unidades resistem à variação mais do que a sílaba, mantendo os demais parâmetros constantes.

Segundo Barbosa (2006) as consequências de diversos mecanismos perturbatórios mostram a pertinência da duração abstrata de uma unidade delimitada por dois onsets de vogal consecutivos como parâmetro de ordem. A homogeneidade duracional da unidade VV torna-se o principal argumento a seu favor.

A decisão por eleger o grupo acentual como unidade básica para a análise prosódica neste trabalho, parte da necessidade de estabelecer parâmetros físicos como critério de segmentação, a partir de um modelo matemático computacional que leva em conta características acústicas no processo de segmentação da fala, a exemplo da duração.

De acordo com Barbosa (2006) para delimitar os grupos acentuais a partir da localização dos acentos frasais, três tipos de técnica podem ser empregados a depender do interesse da pesquisa. A primeira técnica é de base perceptiva e pode ser feita pela escuta do corpus de pesquisa por uma bateria de ouvintes, aos quais se solicita marcar as palavras que lhes parecem proeminentes. No entanto, além de ser um tipo de marcação passível de variação entre os sujeitos, a proeminência de outiva não leva em conta apenas a percepção da duração, mas de todos os parâmetros fonético-acústicos susceptíveis de assinalar algum tipo de saliência perceptiva, como

um *pitch accent*, por exemplo.

Segundo o autor, a dificuldade dessa técnica, no caso do texto lido, é definir a partir de qual porcentagem se pode considerar que há uma proeminência correspondente a um acento frasal, e, portanto, a uma fronteira de grupo acentual.

A segunda técnica parte de picos de duração de unidades VV ao longo dos enunciados e seleciona dentre as palavras em que ocorreram esses picos, aquelas que foram percebidas por pelo menos 20 % dos ouvintes (escolha também arbitrária, mas pelo menos guiada por uma seleção inicial fundamentada em critérios de produção). Essa técnica permite incluir palavras em final de enunciado, às vezes não assinaladas pela primeira técnica por normalmente terem valores baixos de frequência fundamental e serem pouco proeminentes. Os picos de duração das unidades VV nessa condição estão associados ao alongamento característico de fronteira final e, portanto, deveriam ser legitimamente considerados fronteiras de grupo acentual - acentos frasais (BARBOSA, 2006).

A terceira técnica é fundamentada apenas em dados de duração produzida e é perfeitamente adequada para o estudo da produção da organização temporal do enunciado. Para tanto usa-se um procedimento semiautomático que parte da segmentação e etiquetagem manuais de unidades VV ao longo dos enunciados. A partir dessa etapa, três etapas automáticas são efetuadas. Essas etapas partem de um procedimento de normalização básico em estatística, que consiste em obter um valor, chamado de z-score, que especifica o afastamento do valor medido em relação a uma média do mesmo, em unidades de desvio-padrão. Essa normalização é fundamental em estudos de duração, pois minimiza os efeitos de duração intrínseca (um segmento pode durar mais simplesmente porque é, em média, intrinsecamente mais longo) e de variabilidade da duração (um segmento pode durar mais simplesmente porque é mais elástico, variável). O sujeito tem internalizado o conhecimento desses dois componentes intrínsecos, mesmo de outros sujeitos, e está apto a interpretar desvios em relação ao mesmo. Uma evidência disso é que os valores de z - score estão mais correlacionados com a percepção de proeminência do que os valores de duração bruta de unidades VV, embora não sejam completamente isomórficos a essa percepção, visto que a mesma depende de outros fatores além da duração (BARBOSA, 2006).

As três etapas acima referidas foram implementadas no *script SGdetector*. De acordo com Barbosa (2006) a opção por essa técnica não é motivada apenas por uma

questão de praticidade, uma vez que dispensa a bateria de testes de outiva, mas porque permite um estudo da variação fina da duração de unidades VV ao longo dos grupos acentuais, mesmo em casos em que um pico de duração não é percebido como proeminente. A produção desse pico é necessária para a naturalidade da enunciação, já que pequenos picos de duração são sistematicamente produzidos pelos sujeitos ao longo dos enunciados.

### 3. RESULTADOS

Neste trabalho, foram analisados 3.133 segmentos consonantais e vocálicos, e a ocorrência de 477 laringalizações. A partir destes valores, foram calculadas as taxas de laringalização individuais de acordo com a incidência do fenômeno em segmentos vocálicos (Figura 11) e consoantais (Figura 12), separadamente, assim como, a taxa de laringalização total considerando os dois tipos de segmentos.

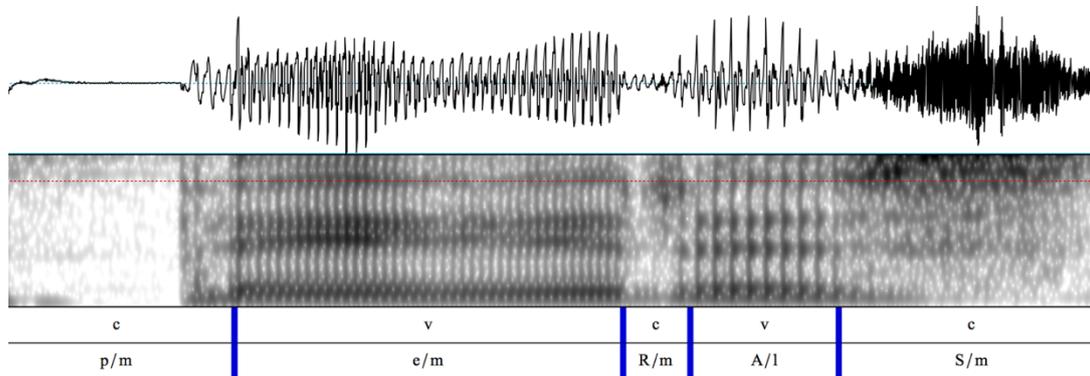


FIGURA 11: Exemplo de laringalização vocálica durante o trecho final da palavra “pêras”. Na primeira camada os rótulos classificam os segmentos em vogal ou consoante, na segunda identifica-se o segmento com o rótulo modal (m) ou laringalizado (l). Observa-se na vogal ‘a’ a irregularidade no sinal acústico, reforçada pela ocorrência de pulsos glotais espaçados no espectrograma.

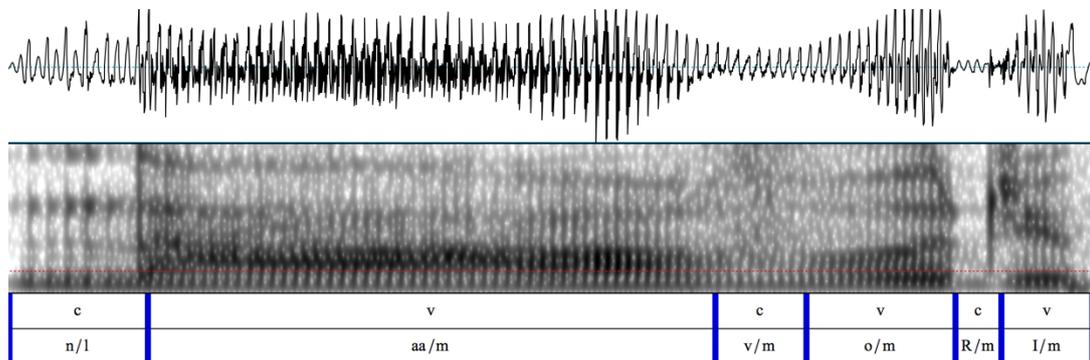


FIGURA 12: Exemplo de laringalização consonantal no primeiro segmento da frase “na árvore”. Na primeira camada os rótulos classificam os segmentos em vogal ou consoante, na segunda identifica-se o segmento com o rótulo modal (m) ou laringalizado (l). Observa-se na vogal ‘a’ a irregularidade no sinal acústico, reforçada pela ocorrência de pulsos glotais espaçados no espectrograma.

### 3.1. Taxas de laringalização vocálica e consonantal

A partir dos valores apresentados na Figura 13, é possível observar uma maior incidência do fenômeno em segmentos vocálicos quando comparado a segmentos consonantais. Pode-se verificar que a maior taxa de laringalização em segmento vocálico (LV) registrada, sujeito F4 (24%), é consideravelmente maior em relação à maior taxa registrada para segmentos consonantais (LC), sujeitos F1 e F4 (4%). Todos os sujeitos apresentaram maiores taxas de ocorrência do fenômeno em vogais em detrimento de consoantes. Essa diferença, entre a proporção de laringalização em segmentos vocálicos e consonantais foi estatisticamente significativa para alfa = 5%, com valor de  $p=0.006$ .

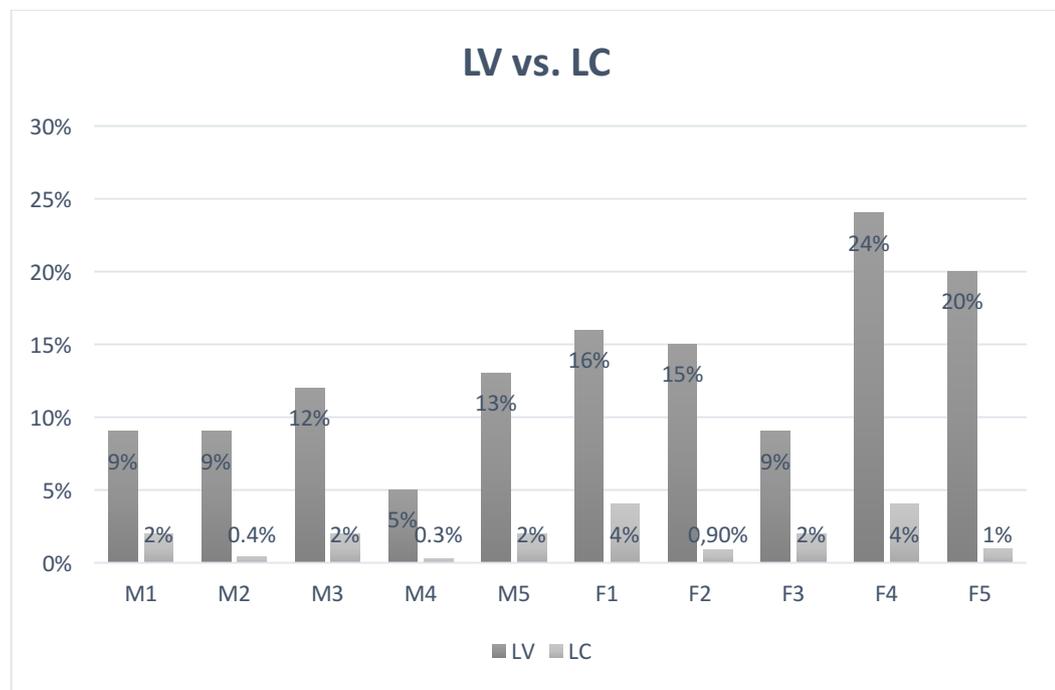


FIGURA 8: Relação entre taxa da laringalização em vogais e taxa de laringalização em consoantes. Diferenças estatisticamente significativas na comparação das proporções  $p=0.006$ .

### 3.2. Variação nas taxas de laringalização entre-sujeitos

No que diz respeito ao grau de variação nas taxas apresentadas pelos indivíduos, constatou-se uma maior variação na produção do fenômeno para segmentos vocálicos ( $p=1.255 \cdot 10^{-11}$ ). É possível verificar na Figura 14, mais

detalhadamente, uma variação de 5% a 24% na incidência da laringalização em vogais, enquanto que para consoantes (Figura 15), a taxa variou de 0.3% a 4%, uma gama proporcionalmente mais reduzida ( $p=0.0003$ ).

No que diz respeito frequência total de laringalização (LTOTAL) apresentada pelos indivíduos, observou-se uma variação de 3% a 14% nas taxas de ocorrência do fenómeno ( $p=4.289 \cdot 10^{-13}$ ). Ver Figura 16.

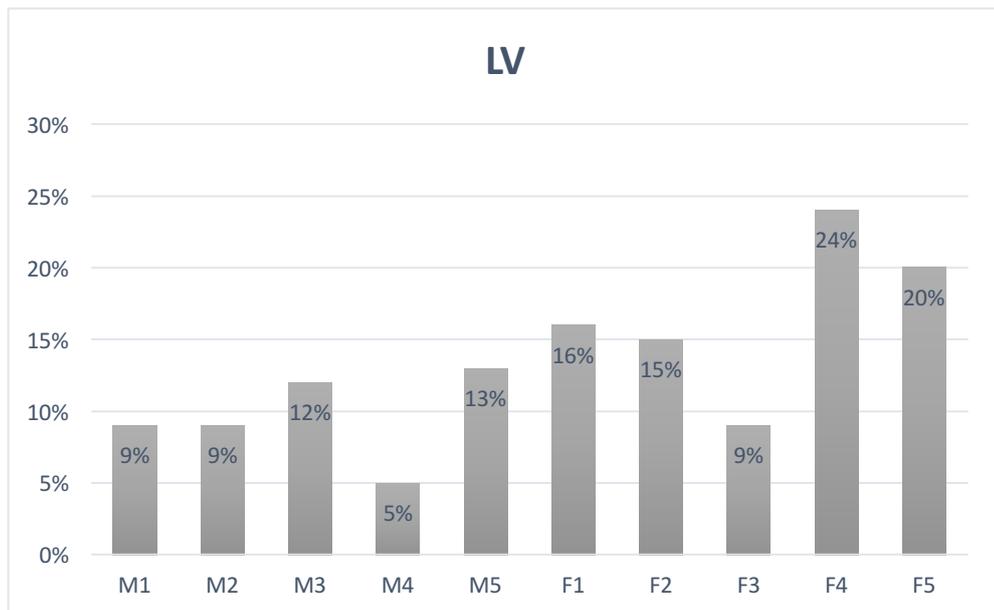


FIGURA 9: Taxas individuais de laringalização em vogais. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 1.255 \cdot 10^{-11}$ ), diferença estatisticamente significativa na comparação entre os grupos F- feminino e M- masculino ( $p=5.539 \cdot 10^{-8}$ ).

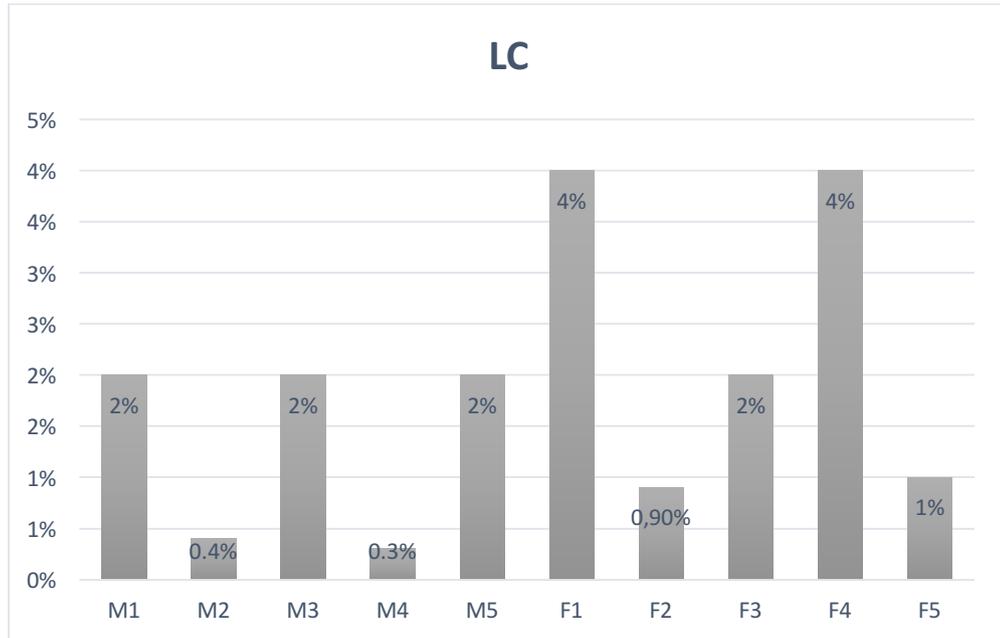


FIGURA 10: Taxas individuais de laringalização em consoantes. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 0.0003147$ ), diferença estatisticamente significativa na comparação entre os grupos F- feminino e M- masculino ( $p = 0.01269$ ).

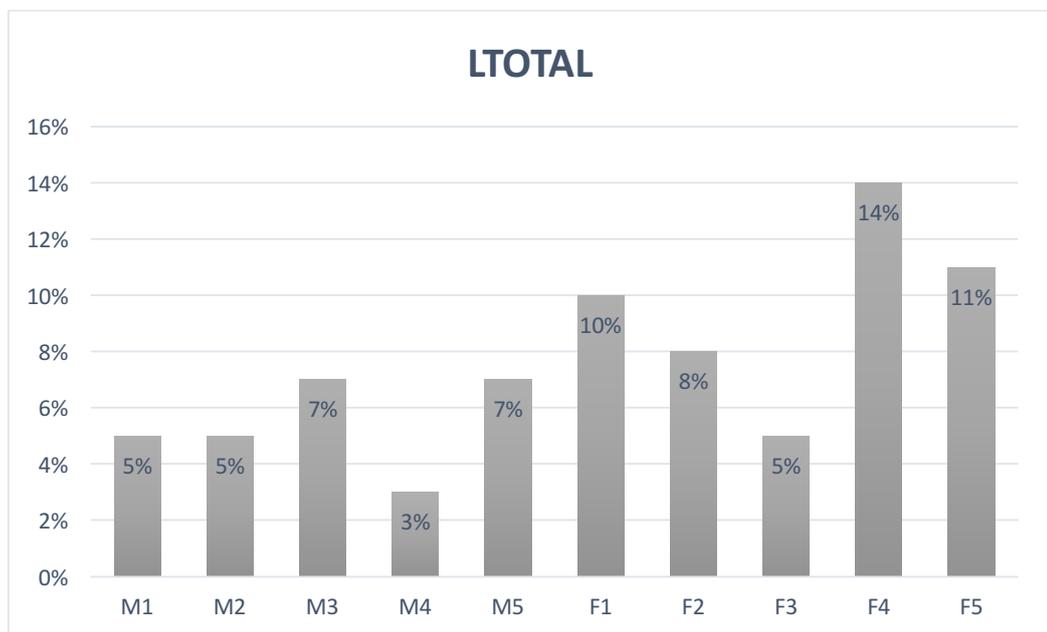


FIGURA 11: Taxas individuais de laringalização total. Diferenças estatisticamente significativas na comparação das taxas entre os sujeitos ( $p = 4.289.10^{-13}$ ), diferença estatisticamente significativa na comparação entre os grupos F- feminino e M- masculino ( $p = 5.225.10^{-9}$ ).

Apesar da variabilidade nas taxas de laringalização ter se apresentado mais consistente nas taxas de laringalização vocálica e laringalização total, pode-se afirmar, no entanto, que todas as três medidas analisadas foram capazes de estabelecer distinção entre diferentes sujeitos no que diz respeito à frequência de ocorrência do fenômeno, revelando valores de  $p < 0.05$ .

As Tabelas 1, 2 e 3 apresentam em modelo de combinações quais combinações entre sujeitos apresentaram significância estatística no que diz respeito a diferenças nas taxas de laringalização. Por meio deste modelo, é possível observar em quais pares a ocorrência do fenômeno pode caracterizar uma pista efetiva na diferenciação dos estilos de locução.

As combinações foram avaliadas levando-se em conta três diferentes medidas, laringalização vocálica, consonantal e laringalização total. Os sujeitos foram comparados entre si, de acordo com as seguintes combinações: homens/homens, homens/mulheres, mulheres/mulheres.

A tabela 1 apresenta as combinações que foram significativas ou marginalmente significativas para a taxa de laringalização vocálica de acordo com a correção de Bonferroni, assumindo valor de  $p = 0.001$ . Nesta condição, 10 combinações entre indivíduos distintos mostraram-se significativas, três combinações intra-sexo (F3 / F4; F3 / F5; M5 / M4) e sete combinações entre sexos.

A correção de Bonferroni é um método estatístico apropriado quando mais de um teste está sendo realizado simultaneamente. A realização de todas as comparações possíveis por pares em um corpus representa perda de poder estatístico ou de consistência estatística. Nestas situações, a correção de Bonferroni (Teste de comparação múltipla com base na estatística de Student t) ajusta o nível de significância observado para o fato de que múltiplas comparações estão sendo feitas. Neste procedimento, dividi-se o valor crítico, digamos 0,05, pelo número de comparações que estão sendo realizadas, garantindo uma maior consistência em procedimentos de comparação.

**TABELA 1: Combinações significativas ou marginalmente significativas taxa de laringalização vocálica.**

<b>combinação</b>	<b>significância</b>	<b>X<sup>2</sup></b>	<b>p-valor</b>
<b>F3 / F4</b>	SIGNIFICATIVO	20.721	5.314. 10 <sup>-6</sup>
<b>F3 / F5</b>	SIGNIFICATIVO	15.074	0.0001
<b>M3 / F4</b>	SIGNIFICATIVO	12.462	0.0004
<b>F4 / M5</b>	MARGINALMENTE SIGNIFICATIVO	10.117	0.001
<b>F4 / M1</b>	SIGNIFICATIVO	20.5	5.963. 10 <sup>-6</sup>
<b>F4 / M4</b>	SIGNIFICATIVO	38.008	7.046. 10 <sup>-10</sup>
<b>F4 / M2</b>	SIGNIFICATIVO	18.857	1.409. 10 <sup>-5</sup>
<b>F2 / M4</b>	SIGNIFICATIVO	17.608	2.714. 10 <sup>-5</sup>
<b>M5 / M4</b>	SIGNIFICATIVO	10.628	0.001
<b>M4 / F1</b>	SIGNIFICATIVO	17.401	3.026. 10 <sup>-5</sup>

A taxa de laringalização vocálica mostrou-se mais consistente na elicitación de combinações estatisticamente significativas, em comparação à taxa consonantal. Na Tabela 2 pode-se observar que a incidência do fenômeno em segmentos consonantais foi consideravelmente menos efetiva na diferenciação de sujeitos. Apenas duas combinações entre-sexos foram significativas ou marginalmente significativas.

**TABELA 2: Combinações significativas ou marginalmente significativas taxa de laringalização consonantal.**

<b>combinação</b>	<b>significância</b>	<b>X<sup>2</sup></b>	<b>p-valor</b>
<b>F1 / M4</b>	SIGNIFICATIVO	11.104	0.0008
<b>F4 / M4</b>	MARGINALMENTE SIGNIFICATIVO	9.6937	0.0018

Por sua vez, as combinações referentes à comparação de taxas de laringalização total mostraram-se ainda mais produtivas, quando comparadas às obtidas por meio de consoantes e vogais. Para esta medida (LTOTAL), 16 combinações apresentaram-se estatisticamente significativas ou marginalmente significativas. Cinco combinações intra-sexo e onze combinações entre sexos elicitaram significância.

É possível verificar que, tanto para a medida de laringalização vocálica como laringalização total, o fenômeno revelou uma maior distinção intra-sexo no grupo feminino. Neste grupo, foram constatadas duas combinações significativamente distintas para a primeira medida (LV) e três para a segunda (LTOTAL), enquanto que no grupo masculino, verificou-se apenas uma combinação significativa para a primeira medida e duas para a segunda.

**TABELA 3: Combinações significativas ou marginalmente significativas taxa de laringalização total.**

<b>COMBINAÇÃO</b>	<b>SIGNIFICÂNCIA</b>	<b>X<sup>2</sup></b>	<b>P-valor</b>
<b>F4 / F2</b>	MARGINALMENTE SIGNIFICATIVO	10.425	0.001
<b>F4 / F3</b>	SIGNIFICATIVO	20.891	4.862. 10 <sup>-6</sup>
<b>F4 / M3</b>	SIGNIFICATIVO	14.537	0.0001
<b>F4 / M1</b>	SIGNIFICATIVO	22.524	2.075. 10 <sup>-6</sup>
<b>F4 / M5</b>	SIGNIFICATIVO	12.693	0.0003
<b>F4 / M4</b>	SIGNIFICATIVO	46.871	7.58. 10 <sup>-12</sup>
<b>F4 / M2</b>	SIGNIFICATIVO	23.336	1.36. 10 <sup>-6</sup>
<b>F2 / M4</b>	SIGNIFICATIVO	17.21	3.347. 10 <sup>-5</sup>
<b>F1 / M4</b>	SIGNIFICATIVO	27.891	1.283. 10 <sup>-7</sup>
<b>F1 / M2</b>	SIGNIFICATIVO	10.708	0.001067
<b>F5 / F3</b>	SIGNIFICATIVO	10.638	0.001
<b>F5 / M1</b>	SIGNIFICATIVO	11.528	0.0006
<b>F5 / M4</b>	SIGNIFICATIVO	31.539	1.954. 10 <sup>-8</sup>
<b>F5 / M2</b>	SIGNIFICATIVO	13.101	0.0002
<b>M3 / M4</b>	MARGINALMENTE SIGNIFICATIVO	10.223	0.001
<b>M5 / M4</b>	SIGNIFICATIVO	13.087	0.0002

### 3.3. Laringalização e sexo

No que diz respeito a ocorrência da laringalização nos grupos, feminino e masculino, é possível verificar que as maiores taxas do fenômeno foram registradas por falantes do sexo feminino. Observa-se na Figura 14 que, as quatro maiores taxas de laringalização vocálica registradas foram produzidas por mulheres (sujeitos F4, F5, F1 e F2).

Na análise da laringalização em consoantes o mesmo pode ser constatado (Figura 15), a duas maiores taxas foram produzidas por mulheres (F1 e F4), assim como na taxa de laringalização total, em que se constata as quatro maiores taxas referentes à falantes do sexo feminino (Figura 16).

Dentre todos os sujeitos, F4 foi quem produziu as maiores taxas de laringalização vocálica (24%) e consonantal (4%), juntamente com F1 (16%) e (4%) respectivamente. Do mesmo modo, F4 foi quem produziu a maior taxa de laringalização total (14%).

No geral, a análise estatística revelou que as mulheres laringalizaram mais neste corpus do PB quando comparadas a falantes do sexo masculino.

### 3.4. Ocorrência da laringalização em relação ao grupo acentual

Os resultados revelam uma tendência de que o fenômeno da laringalização esteja associado a regiões de fronteiras prosódicas no PB, mais especificamente, a fronteiras de grupos acentuais, unidade prosódica imediatamente acima da sílaba e inferior a unidade entoacional (BOTINIS; GRANSTRÖM; MÖBIUS, 2001).

Para todas as ocorrências do fenômeno, uma classificação geral foi utilizada a fim de caracterizar o contexto de sua incidência. A ocorrência da laringalização foi classificada em duas categorias: “1.” Em fronteira de grupo acentual, “2.” Não fronteira.

Pode-se verificar por meio da Figura 17 a proporção de ocorrência da laringalização nestes contextos. Nota-se que, para todos os sujeitos, a laringalização em fronteiras de grupos acentuais foi numericamente superior a ocorrência do fenômeno em regiões onde não há fronteira.

A análise estatística revelou que para 9 dos 10 sujeitos avaliados, esta diferença elicitou significância estatística (Tabela 4). Apenas no sujeito M2, a diferença numérica entre a incidência de laringalização em fronteira e não fronteira não evidenciou significância. Neste corpus, 83% das laringalizações estiveram alinhadas a fronteiras de grupos acentuais, 17% não ocorreram em região de fronteira prosódica.

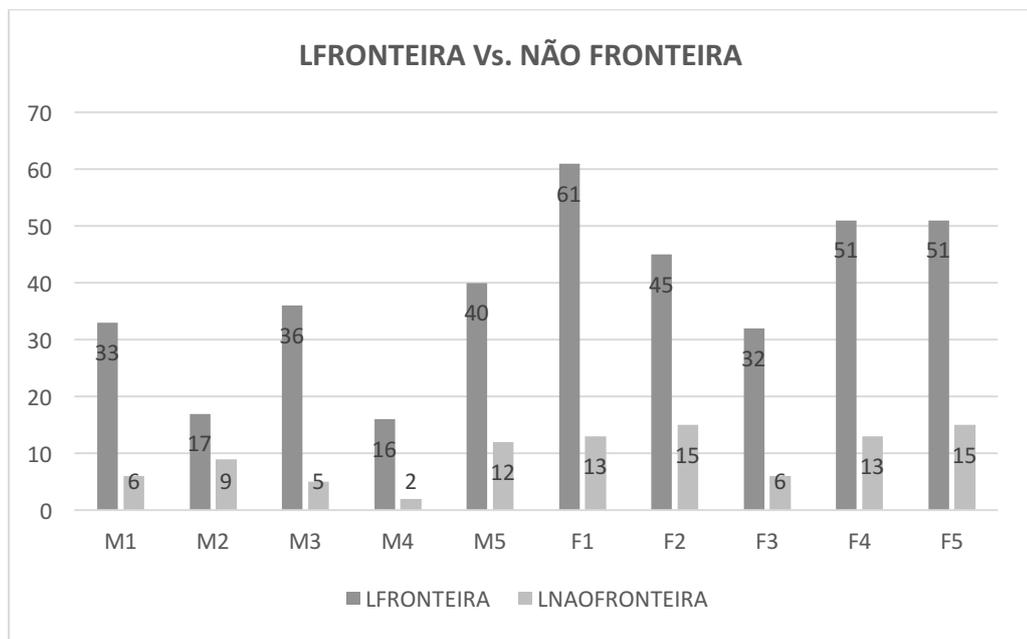


FIGURA 12: Proporções da ocorrência de laringalização em fronteiras de grupos acentuais vs. não fronteira. Valores individuais.

**TABELA 4: Proporções da ocorrência de laringalização em fronteiras de grupos acentuais vs. não fronteira.**

SUJEITO	L Fronteira	L Não fronteira	X <sup>2</sup>	p-valor
M1	33	6	12.121	0.0004
M2	17	9	0	1
M3	36	5	17.361	3.091.10 <sup>-5</sup>
M4	16	2	7.5625	0.0059
M5	40	12	5.625	0.01771
F1	61	13	18.951	1.341.10 <sup>-5</sup>
F2	45	15	4.3556	0.0368
F3	32	6	11.281	0.0007
F4	51	13	11.294	0.0007
F5	51	15	7.8431	0.005

### 3.5. Parâmetros fonético-acústicos envolvidos na laringalização

A partir da seleção dos trechos vocálicos nos quais houve confirmação da ocorrência de laringalização, foi possível executar a extração automática dos parâmetros fonético-acústicos relacionados ao fenômeno, a exemplo da duração mediana da laringalização em relação à duração total do segmento, correspondente aos valores centrais da medida, a porcentagem da queda de frequência fundamental do trecho modal para o laringalizado, a diferença de intensidade do trecho modal para o laringalizado, a porcentagem de diferença de *jitter*, *shimmer* e a proporção harmônico ruído (HNR) nos trechos modal e laringalizado, a partir do cálculo das médias entre os diferentes registros. Na Tabela 5 é possível observar os resultados obtidos.

Neste corpus, verificou-se a tendência geral de que a laringalização ocupe apenas uma porção da duração total do segmento vocálico. A mediana foi de 71%, o que significa dizer que a extensão da laringalização dura, em geral, 71% da duração total de todo o segmento.

No que diz respeito aos parâmetros fonético-acústicos mais influenciados pela ocorrência do fenômeno, observa-se que a intensidade foi o parâmetro mais modificado na passagem do trecho modal para o laringalizado, revelando uma queda de 13% na transição.

As porcentagens de queda de frequência fundamental e diferença de jitter dos trechos modais para os trechos laringalizados também sofreram modificações, no entanto, observa-se por meio da inspeção dos valores que a queda de intensidade demonstrou-se ser o parâmetro mais importante como evidência acústica da ocorrência do fenômeno.

**TABELA 5: Modificações nos parâmetros fonético-acústicos envolvidos na laringalização.**

<b>% Queda de F0</b>	<b>%Queda int (dB)</b>	<b>dif % Jitter</b>	<b>dif % Shimmer</b>	<b>dif % HNR</b>	<b>H1H2 Homens</b>	<b>H1H2 Mulheres</b>
3%	13%	1%	0%	0%	2.5 dB dp: 3.4	0.39 dB dp:3.7

No cálculo das diferenças de medianas de H1 - H2 entre os trechos modais e laringalizados, apenas as vogais /a/, /ε/, /ɔ/ foram incluídas na análise, para evitar que o efeito de formante afete a amplitude dos harmônicos. Assim, ao analisar vogais com valores de F1 baixo, a exemplo das vogais altas e médio-altas, a amplitude de F1 faz aumentar possivelmente o segundo harmônico, H2, o que envia saber se o aumento é por efeito da qualidade de voz ou de formante. Neste sentido, as vogais citadas, com valores de F1 altos, alteram harmônicos bem mais afastados de H2, dando mais confiabilidade ao cálculo da medida de H1 - H2.

É possível verificar, no presente trabalho, que o valor médio da relação entre os dois primeiros harmônicos apresentou-se relativamente mais reduzido no sexo masculino quando comparado ao grupo feminino, evidenciando uma diminuição de

2.5 dB, no primeiro grupo. Verifica-se, no entanto, que houve mudança nos valores de H1 - H2 entre os trechos modal e laringalizado para ambos os grupos, caracterizando valores de H1 - H2 mais reduzidos nos trechos com laringalização.

## 4. DISCUSSÃO

A análise dos dados permitiu uma compreensão mais objetiva acerca do fenômeno da laringalização no PB, revelando particularidades ainda não relatadas para esta língua. Neste capítulo, as hipóteses previamente levantadas serão confrontadas com os dados obtidos, ao mesmo tempo em que outros trabalhos servirão como base para discussão levantada.

### 4.1. Laringalização vocálica e consonantal

Neste trabalho, os segmentos vocálicos foram os segmentos que apresentaram os maiores índices de laringalização. A laringalização vocálica não apenas demonstrou-se mais frequente, como também foi a medida que apresentou maior variação entre os indivíduos.

Estudos referem funções linguísticas diferenciadas para segmentos vocálicos e consonantais. Toro et al., (2008) defende que as consoantes estão geralmente relacionadas com processamento de informações lexicais das palavras, a partir da diferenciação dos elementos, enquanto que as vogais assumem, de modo geral, o papel de auxiliar no processamento de informações prosódicas destinadas à organização da sintaxe, à pauta prosódica e à estruturação discursiva.

Selkirk (1984) propõe uma sequência de sonoridade para as línguas, organizando os segmentos em uma escala hierárquica de sonoridade. Nesta hierarquia, a autora faz referência aos segmentos vocálicos, tanto vogais altas, médias como baixas, enquanto os elementos mais sonoros na cadeia da fala. Para ela, as línguas são sensíveis à esta escala, podendo utilizá-la para aplicar restrições até mesmo sobre quais elementos e a sequência que eles poderão ocupar numa sílaba.

Partindo do pressuposto de que a laringalização é uma função da produção de voz na laringe, os seus efeitos são mais evidentes em segmentos de fala que envolvem vibração das pregas vocais, os chamados elementos "sonoros". Por outro lado, os segmentos da fala em que não ocorre vibração das pregas vocais, espera-se que esses elementos "surdos" permanecem relativamente inalterados (HIRSON E DUCKWORTH, 1993).

Embora o intuito deste trabalho não tenha sido avaliar quais as motivações segmentais para a ocorrência da laringalização, é evidente que a presença de

sonoridade no segmento é uma condição imprescindível para que o fenômeno se realize. Os altos valores nas taxas de laringalização vocálica podem sugerir uma preferência do fenômeno em ocorrer naqueles segmentos da fala com maiores índices de sonoridade.

#### 4.2. Implicações para a comparação de locutor

Hollien (1990) propõe variáveis linguísticas que devem receber atenção numa tarefa de comparação entre falantes, a exemplo da frequência fundamental geral, máxima, mínima e média. Segundo o autor muitos indivíduos apresentam padrões frequência fundamental habituais que podem ajudar o ouvinte a identificá-los por meio do padrão fonatório. Outras variáveis que devem ser investigadas de acordo com o autor são: a intensidade, a articulação - o foco neste caso é especialmente nas produções idiossincráticas de segmentos vocálicos e consonantais, bem como nas proporções de formantes das vogais e processos de coarticulação; a qualidade vocal, a prosódia - destacando a importância da análise da entoação e das durações dos elementos; e a análise minuciosa das características da fala.

Esta última categoria não deve ser considerada menos importante, antes, diz respeito a elementos específicos da fala como (i) dialeto, (ii) uso incomum do acento ou variações linguísticas (iii) padrões linguísticos idiossincráticos, (iv) alterações de fala e (v) pronúncias idiossincráticas. Esses recursos vão além das considerações simples de articulação ou prosódia referidos acima, e podem fornecer pistas muito poderosas no reconhecimento de fala (HOLLIEN, 1990).

Neste contexto, a análise da laringalização pode contribuir como uma pista prosódica importante na tarefa de comparação de locutor, justificada pela gama de variação de suas medidas em diferentes locutores, afirmando o seu uso como estilístico no PB e, portanto, revelador de um comportamento linguístico indiosincrático.

De acordo com Hirson e Duckworth (1993), em geral, um índice eficaz para a identificação de falante depende de: (a) baixa variação do fenômeno intra-falante, e (b) a alta variação inter-falante, situação que garante uma maior consistência na comparação de locutores.

No presente trabalho, a análise estatística permitiu verificar com consistência quais das medidas envolvidas na laringalização demonstraram-se mais adequadas

para a tarefa de comparação de locutor. Os dados revelaram de forma sistemática que as pistas acústicas evidenciadas em vogais foram significativamente mais sólidas para a aplicação forense no PB.

Em um contexto real de comparação de estilos de locução, o objetivo que fundamenta a análise é sempre verificar, com algum grau de certeza, o quanto os registros de fala que o avaliador dispõe são semelhantes ou diferentes entre si (MORRISON, 2013).

Nesse sentido, as diferenças intra-sexo são, em certa medida, muito importantes para o uso forense. Exceto que se esteja lidando com um caso de disfarce vocal, a capacidade que uma medida tem de realizar distinção entre falantes do mesmo sexo parece imprescindível, uma vez que a tendência geral é de que falantes do mesmo sexo sejam comparados numa investigação.

Os resultados obtidos sugerem as medidas de laringalização vocálica e laringalização total como mais consistentes para o uso forense. Estas medidas foram capazes de eliciar um maior número de combinações significativas em comparação com a taxa consonantal. Por si só, as pistas consonantais foram pouco expressivas no desempenho desta tarefa, no entanto, quando associadas às vogais, foram capazes de acentuar as diferenças e ampliar o repertório de combinações significativas intra-sexo.

No que diz respeito ao número de combinações estatisticamente significativas nos grupos, foi possível verificar um maior poder de distinção de pares no grupo feminino, tanto para a taxa de laringalização vocálica como para a taxa de laringalização total. No grupo feminino, talvez pela maior gama de variações nas taxas, a medida demonstrou-se mais efetiva na capacidade de distinção de estilos locucionais.

De fato, a análise de uma única característica ou aspecto da fala possui pouco valor pericial se analisado de forma isolada, no entanto, a análise do conjunto de parâmetros possíveis pode muito dizer sobre as preferências daquele que os utiliza (HOLLIEN, 1990). A análise da laringalização, neste sentido, vem contribuir para o refinamento na análise dos parâmetros da fala em contextos forense.

De acordo com Segundo, Alves e Trinidad (2013) a estratégia mais comum utilizada por suspeitos para disfarçar suas vozes implica em mudanças no padrão laríngeo ou no tipo de fonação, no entanto, segundo os autores, é difícil manter este tipo de disfarce vocal por longos períodos de tempo.

Segundo, Alves e Trinidad (2013) realizaram a análise do primeiro corpus de vozes disfarçadas na língua espanhola. O corpus "CIVIL", composto por 60 falantes de espanhol, com idades entre 20 e 35 anos, os quais foram submetidos à execução de três tarefas: conversa espontânea, leitura de palavras e leitura de texto, fazendo uso de três diferentes tipos de fonação: modal, falsete e laringalizada.

Ao avaliar as características que permanecem apesar das tentativas do falante de disfarçar sua voz, em resultados preliminares os autores puderam constatar que a tarefa de reconhecimento de falante é significativamente mais fácil quando os sujeitos tentam disfarçar suas vozes por meio do registro falsete do que quando laringalizam. Este resultado, no entanto, foi mais significativo para as vozes femininas.

Hirson e Duckworth (1993), por sua vez, buscaram avaliar em seu experimento se a laringalização constitui um disfarce vocal eficaz na tentativa de ocultar a identidade de fala. A eficácia da estratégia de disfarce foi medida pela redução na precisão com a qual foneticistas foram capazes de associar um trecho lido com fonação predominantemente laringalizada de um locutor com a leitura modal do mesmo, em um corpus com registro de 10 vozes.

Ficou claro neste experimento paradigmático que, associar as gravações laringalizadas dos falantes às suas vozes modais foi uma tarefa muito menos precisa do que associar a gravação modal desses sujeitos às suas amostras de fala modal correspondentes.

Constatou-se que os ouvintes são bastante melhores na identificação de um par quando ele realmente ocorre, do que na tarefa de excluir uma amostra que não corresponde com qualquer uma das amostras do teste (gravação A ou B). No geral, ouvintes apresentaram um pior resultado na tarefa de discriminação quando o elemento A ou B envolveu uma voz laringalizada. Isto sugere que a identificação de fala de um sujeito somente por meios perceptuais pode conduzir a falsos positivos, e que este efeito é ainda maior quando uma das gravações comparada é laringalizada (HIRSON E DUCKWORTH, 1993).

Embora grande parte dos estudos tenham-se voltado para a análise da laringalização enquanto um tipo de estratégia de disfarce vocal, pouco é sabido sobre de que forma a presença ou a ausência do fenômeno, em situações espontâneas de fala, pode contribuir para a construção de um perfil locucional e conseqüentemente para a diferenciação de locutor. O presente trabalho buscou, de forma objetiva,

preencher esta lacuna, buscando reproduzir condições mais naturais de coleta e análise do fenômeno.

#### 4.3. Variável Sexo e Laringalização

Os resultados foram consistentes quanto à influência do sexo sobre as taxas de laringalização. No corpus utilizado nesta dissertação verificou-se que mulheres laringalizaram significativamente mais do que falantes do sexo masculino, para todas as medidas avaliadas, de acordo com o teste paramétrico de proporção ( $p = 5.539. 10^{-8}$ ). Este achado, ainda não verificado anteriormente para o PB, concorda e contrapõe resultados de pesquisas realizadas em outros países.

Henton e Bladon (1988) realizaram o primeiro estudo quantitativo em larga escala voltado para o uso sociocultural da laringalização, a partir da análise de dois dialetos do Inglês Britânico. Os autores observaram que falantes pertencentes ao grupo masculino, de ambas as variedades estudadas, laringalizavam com maior frequência quando comparados a falantes do sexo feminino. Com base em seus experimentos, foi possível verificar que a voz laringalizada estivesse associada a um marcador de robustez da fala masculina e desta forma empregado por homens para afirmar a masculinidade.

Pittam (1987) chegou a uma conclusão semelhante em um teste perceptual no qual os ouvintes foram submetidos à escuta de várias qualidades de voz, incluindo a voz laringalizada, em enunciados produzidos por falantes do inglês australiano. Os ouvintes foram instruídos a avaliar a qualidade de voz nas dimensões de status e aspectos sociais. Os resultados obtidos indicaram que a utilização da fonação laringalizada esteve associada a um alto status social, sugerindo o fenômeno como mais característico no sexo masculino, talvez pelos baixos valores de frequência fundamental que o acompanham.

Yuasa (2010), por sua vez, constatou que falantes do sexo feminino no inglês americano empregaram esta qualidade de voz com maior frequência do que os homens e mulheres japonesas falantes do inglês. Seus testes de percepção indicaram que os ouvintes percebem a voz crepitante de mulheres americanas como soando fundamentalmente mais "educado", "profissional", "elevado" e "urbano", assim como também mais "hesitantes", "não-agressivas" e "informais" do que quando produzem o registro modal. De acordo com a autora, é possível que os significados socioculturais

indexados a esta qualidade de voz possam ter sido reinterpretados como uma nova tendência de comportamento verbal feminino nos últimos anos.

Dilley, Shattuck-Hufnagel e Ostendorf (1996), Redi e Shattuck-Hufnagel (2001), verificaram a tendência de falantes profissionais do sexo feminino em produzir maiores taxas de laringalização no inglês americano em comparação a locutores profissionais do sexo masculino.

Segundo Yuasa (2010), em decorrência dos baixos valores de frequência fundamental que acompanham a laringalização, o uso com frequência do fenômeno por apresentadoras nos Estados Unidos pode, certamente, configurar uma estratégia empregada para projetar uma imagem de autoridade, que é mais característica da fala masculina.

De acordo com Tajfel (1974) as motivações socioculturais para o uso da laringalização por mulheres pode variar dependendo dos tipos de valores que cada sociedade aspira ou em decorrência de aspectos culturais intrínsecos. Em uma sociedade como a dos Estados Unidos, por exemplo, as mulheres podem intuitivamente reconhecer que o sucesso pessoal e profissional pode depender de uma assimilação ao grupo dominante (masculino). Elas podem, desta forma, recorrer a mudanças no seu comportamento através da apropriação de condutas desse grupo.

Ingle (2005), ao analisar as peculiaridades do inglês falado no noroeste dos Estados Unidos, constatou que as mulheres daquela região apresentam uma tendência a utilizar a voz laringalizada mais frequentemente do que homens. A partir da análise do conteúdo de fala de 14 participantes, pareados em idade, sexo e status social, a autora refere que o uso frequente da voz laringalizada pode ocorrer como um possível recurso dialetal de mulheres daquela região em particular.

Henton e Bladon (1988) especulam que os falantes do sexo feminino podem não precisar contrair a estrutura das pregas vocais com tanta intensidade para atingir o estado de agrupamento ou constrição, como fazem os homens, por exemplo, porque o seu menor comprimento de prega vocal naturalmente contribui para a deformação e espessura necessária à produção do registro laringalizado.

Segundo Shattuck-Hufnagel (2001) os fatores que contribuem para as diferenças no sexo referentes à taxa de laringalização podem ser anatômicos, sociolinguísticos ou talvez uma combinação dessas variáveis.

#### 4. 4. Sobre a relação entre laringalização e prosódia

Existem diferentes maneiras de se segmentar um discurso. De acordo com Kreiman (1982) sentenças e parágrafos são primariamente unidades semânticas. Apesar disso, não há razões, *a priori*, para que tais unidades não ocorram em um discurso; pelo contrário, essas unidades se realizam na fala e têm formas e características próprias, as quais os falantes produzem de forma eficiente e os ouvintes são capazes de identificar.

Existem diferentes formas de tais fenômenos serem concebidos em um diálogo. Podem consistir basicamente em pistas de uma unidade terminal, pistas essas que sinalizam o fato de uma sentença ou parágrafo estar definitivamente acabado e uma nova unidade estar prestes a começar; podem consistir em pistas sinalizadoras do início de uma nova unidade; ou ainda consistir em pistas que compreendem fronteiras e indicam que uma unidade terminou e que outra já se iniciou (KREIMAN, 1982).

Segundo o autor, quando se trata do fluxo da fala, não faz sentido assumir que existam apenas pistas de unidade terminal. É possível que as pistas forneçam informações sobre fronteiras após o fato, no entanto, o conteúdo semântico pode, presumivelmente, fornecer informações para os interlocutores sobre ambos os fins e começos das unidades discursivas.

Oliveira, Cruz e Silva (2012) ao analisar o papel que as pistas prosódicas exercem no processo de reconhecimento da estrutura discursiva de narrativas espontâneas, verificaram que a estrutura subjacente a um dado gênero discursivo oral pode ser identificada com um alto grau de confiabilidade, mesmo quando, apenas as características prosódicas do discurso são preservadas, desprezando informações lexicais, sintáticas e semânticas. De acordo com os pesquisadores, isso sugere que a prosódia, além de atuar significativamente na estruturação do discurso, sob a perspectiva da produção, tem um relevante papel também para a percepção da estrutura discursiva.

*Phrasing* (processo também conhecido como agrupamento) está associado com a segmentação do enunciado em unidades prosódicas variáveis. Teorias prosódicas e estudos fonológicos referem várias categorias e unidades que abrangem desde a unidade silábica ao enunciado como um todo (NESPOR E VOGEL, 1986). Para uma descrição e análise entoacional, no entanto, a sílaba, o grupo acentual e a

frase entoacional são, na sua maioria, unidades mais reconhecidas (BOTINIS; GRANSTRÖM; MÖBIUS, 2001).

O grupo acentual (conhecido também pelos sinônimos pé, unidade tonal, palavra prosódica, etc.) é a unidade prosódica imediatamente acima da sílaba. Refere-se a um intervalo que parte da sílaba acentuada até, mas não incluindo, a próxima sílaba tônica. Uma sequência prosódica de sílabas acentuadas e não acentuadas; é assim que se estruturam os grupos acentuais, os quais podem estar correlacionados a gestos tonais distintos, onde *onsets* e *offsets* tonais estão alinhados com o início e o fim de seus correspondentes grupos acentuais. A distribuição e a relação entre os grupos acentuais dentro de um enunciado determinam o ritmo, que é uma característica básica das línguas com diferentes estruturas prosódicas (BOTINIS; GRANSTRÖM; MÖBIUS, 2001).

Entre o grupo acentual e a frase entoacional, reconhecidas para a descrição e análise de diferentes línguas naturais, uma unidade intermediária tem sido sugerida, a frase entoacional intermediária, a partir da qual uma frase entoacional pode ser decomposta em várias frases entoacionais intermediárias (BOTINIS; GRANSTRÖM; MÖBIUS, 2001).

De acordo com Botinis; Granström; Möbius, (2001) a frase entoacional (também referida como unidade entoacional, sentença prosódica, grupo respiratório, etc.) refere-se a estruturas de entoação coerentes e sem rupturas prosódicas importantes. Uma frase entoacional pode consistir em uma única sílaba até frases sintáticas; cláusulas, sentenças e um enunciado maior podem assim serem pronunciados com uma ou várias frases entoacionais, que podem não ter qualquer correspondência direta com a sintaxe ou unidades semânticas. Conseqüentemente, a correlação dos limites de frases entoacionais e limites sintáticos é casual ao invés de causal. Por outro lado, as frases entoacionais podem estar associadas a diferentes tipos de sentença, a saber, declarativas, perguntas, comandos, etc. Frases entoacionais podem também estar relacionadas a unidades de informação, às quais podem ser classificadas como unidades de informação mais ou menos autônomas.

Parece claro que a ocorrência e distribuição da laringalização no PB leva em conta a estrutura prosódica da língua, como ocorre no inglês americano, por exemplo.

Dilley, Shattuck-Hufnagel & Ostendorf (1996); Redi e Shattuck-Hufnagel (2001) verificaram o alinhamento entre a ocorrência da laringalização e fronteiras prosódicas

de frases entoacionais mediais e finais no inglês. As autoras afirmam, no entanto, que são nestas últimas em que se observa maiores índices de laringalização.

Do mesmo modo, Lima-Gregio (2011) verificou no PB, em estudo com falantes com e sem patologia de fala, o alinhamento, neste segundo grupo, entre o fenômeno e fronteiras de unidades prosódicas de grupo acentual. Quando avaliados os sujeitos com alteração de fala, a exemplo da fissura labiopalatina, a pesquisadora relatou uma menor concordância no que diz respeito ao contexto de incidência do fenômeno.

Em seu trabalho, Lima-Gregio (2011) também adotou uma metodologia semiautomática de segmentação e extração dos parâmetros acústicos da laringalização, elegendo o grupo acentual enquanto unidade prosódica mínima para análise, mediante a aplicação do *script* SGdetector, o mesmo utilizado neste trabalho.

A autora sugere que a força prosódica nas fronteiras pode favorecer a laringalização, em decorrência de uma tensão mais extrema e variável nas pregas vocais em torno da fronteira. Este achado sustenta, portanto, a hipótese de que o acento frasal em fronteira atrai as laringalizações (LIMA-GREGIO, 2011).

Em uma abordagem dinâmica<sup>16</sup> da análise da fala, Lima-Gregio (2011) refere que haveria uma falha no componente supervisor, quando o sujeito produz laringalização em regiões fora de fronteira. Ou seja, numa produção normal, o componente supervisor age inibindo a laringalização nas regiões não fronteiriças, a fim de garantir a função prosódica de demarcação nas regiões de fronteiras. Quando essa inibição não ocorre ou ocorre no contexto equivocado é sinal de que o mecanismo de supervisão não está restringindo adequadamente os erros gerados pelo modelo dinâmico da tarefa.

Lucente (2012), por sua vez, em uma análise entoacional do PB, transcrito por meio do sistema DaTo, observou uma consistente coincidência entre fronteiras de grupos acentuais, contornos entoacionais e segmentos discursivos. Tais coincidências entre entoação, ritmo e discurso, ou entre aspectos prosódicos e discursivos sugere a hipótese de que no modelo dinâmico do ritmo da fala - MDR, proposto por Barbosa (2006), paralelamente ao oscilador silábico, que determina a

---

<sup>16</sup> A teoria dos sistemas dinâmicos proposta por Kelso (1995) defende a tese de que toda resposta de um comportamento humano, cognitiva ou motora, desenvolve-se e se estabiliza a partir de um padrão dinâmico auto-organizável. A ideia de auto-organização está ligada à de sinergia funcional, que é resultante da coordenação entre os movimentos que interagem entre si visando à realização de uma tarefa. Ou seja, os padrões dos movimentos interagem entre si com o intuito de um ajuste que melhor atenda às restrições internas e externas para o desempenho da tarefa (LIMA-GREGIO, 2011).

proeminência baseando-se apenas na duração dos segmentos, exista um oscilador glotal, o qual determina proeminências e fronteiras entoacionais de acordo com a frequência de vibração de pregas vocais na glote, sinalizados pelo movimento de subida ou descida da frequência fundamental.

Ainda segundo Lucente (2012), o alinhamento entre estruturas discursiva (DSs), rítmica (SGs) e entoacional (contornos dinâmicos) indica um sistema dinâmico de produção da fala e da entoação que tem o grupo acentual como atrator.

Esse alinhamento sugerido por Lucente (2012) entre grupos acentuais e unidades discursivas, poderia fundamentar a hipótese de que a laringalização no PB possa exercer a função prosódica de sinalizar fronteiras, a partir do seu alinhamento com grupos acentuais e justificada pelos baixos valores de frequência fundamental comum ao fenômeno, fornecendo pistas ao ouvinte sobre a organização discursiva.

Há relatos na literatura de outros fatores prosódicos que também podem estar relacionados a ocorrência do fenômeno. Pesquisadores sinalizam o papel do *pitch accent* e das pausas na incidência da laringalização pré e pós fronteira (DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996; REDI E SHATTUCK-HUFNAGEL, 2001), como também das hesitações como um tipo especial de fronteira prosódica (LUCENTE, 2012). Nos casos em que a ocorrência laringalização não esteve associada à presença de fronteira prosódica, observou-se de forma frequente a presença de hesitações e reformulações no fluxo de fala acompanhadas do gesto glotal. Sugere-se a realização de trabalhos futuros que analisem a influência desses fatores no processo de segmentação da fala e na produção do fenômeno.

#### 4.5. Correlatos acústicos e fisiológicos da laringalização e fronteiras prosódicas

Diversos estudos relatam que fronteiras prosódicas são geralmente assinaladas por meio de uma série de pistas acústicas, incluindo mudanças na frequência fundamental, fonação irregular, duração do segmento e pausas. Os falantes geralmente dão pausas nos limites estruturais tanto no discurso lido como na fala espontânea (KREIMAN, 1982; SLIFKA, 2006; SLIFKA, 2007).

A laringalização, caracterizada por um padrão de fonação irregular, tem sido associada a regiões de fronteiras prosódicas em diversas línguas naturais, como verificado no atual trabalho e referido por estudos da área. Esse alinhamento entre a

incidência do fenômeno e regiões de fronteira tem implicações diretas relacionadas a fisiologia e aos parâmetros acústicos envolvidos.

Segundo Ladefoged e Gordon (2001), na maioria das vezes os segmentos que são descritos como sendo realizados com fonação não-modal não sustentam o trecho não-modal ao longo de todo o segmento. Em vez disso, realiza-se como parte de um segmento sonoro e/ou transborda para um segmento adjacente. Uma característica temporal especialmente comum da fonação não-modal é o seu alongamento em contextos em que outras propriedades perceptualmente importantes são minimizados.

Neste trabalho, a ocorrência do fenômeno laríngeo apresentou uma duração de 71% da duração total da vogal, esteve relacionada a diminuição da intensidade do sinal sonoro, redução nas medidas de H1-H2 e da frequência fundamental, bem como, repercussões sobre o índice de perturbação da frequência fundamental (*jitter*).

De acordo com Queiroz (2012) as medidas acústicas referentes a frequência fundamental, sua primeira derivada, o declínio espectral e o espectro de longo termo, no domínio de longo termo, além de diferença de intensidade dos dois primeiros harmônicos (H1-H2), no domínio de curto termo, foram capazes de diferenciar grupos de ajustes fonatórios modal e não modal de qualidade vocal nas esferas fonatória e de tensão. No grupo de ajuste fonatórios, as medidas de frequência fundamental e de sua primeira derivada permitiram diferenciar o ajuste neutro (modal), dos demais (não-modal). No caso dos ajustes de tensão e de sua associação do trato vocal (posição vertical de laringe), as medidas de declínio espectral e de espectro de longo termo foram importantes. As medidas de H1-H2 demonstraram-se sensíveis a ajustes de tensão e a alguns ajustes fonatórios não modais.

Segundo Hanson e Chuang (1999), do conjunto de harmônicos produzidos pela vibração de pregas vocais o primeiro harmônico tem sua amplitude associada ao grau de excursão lateral das pregas vocais durante a produção vocal; assim, altos valores deste harmônico indicam a descontinuidade do contato das pregas vocais. Nesta relação, quanto maior for o valor da intensidade do primeiro harmônico (H1), em relação ao segundo (H2), menor é o coeficiente de contato entre as pregas vocais.

De acordo com Szakay e Torgersen (2015), na análise da inclinação espectral calculada a partir da diferença entre as amplitudes do primeiro e segundo harmônicos (H1 e H2), valores mais elevados da medida indicariam um padrão de fonação soproso, em decorrência do menor coeficiente de contato entre as pregas vocais

(OQ), enquanto que valores reduzidos poderiam indicar uma qualidade de voz mais laringalizada, a partir de um maior coeficiente de contato.

Szakay e Torgersen (2015) analisaram em um estudo a qualidade de voz em falantes do inglês Britânico (dialeto falado em Londres), a partir da análise de parâmetros acústicos como a estrutura harmônica (H1-H2) e a frequência fundamental. Os autores compararam o discurso de 28 falantes do Inner London Hackney, de falantes ingleses e não ingleses, com o discurso de 14 oradores anglo-holandeses de Havering. Os resultados revelaram que a qualidade da voz e *pitch* diferiram significativamente entre as duas localizações, onde os discursos dos falantes da Inner London apresentaram em média um *pitch* mais baixo e uma qualidade de voz mais ofegante em comparação a outra localidade. Segundo os pesquisadores, este dado sugere que a medida H1-H2 não é necessariamente dependente dos valores globais de  $f_0$ .

No presente trabalho, no entanto, foi possível constatar uma tendência de menores valores de H1-H2 nos trechos de laringalização em falantes do grupo masculino, no qual se observou uma menor média na medida de frequência fundamental (132 Hz) em comparação ao grupo feminino (205 Hz). Assim, menores valores nas medidas de H1-H2 poderiam indicar uma produção mais característica do fenômeno, em decorrência de uma maior constrição das pregas vocais durante a sua realização.

Os resultados do presente trabalho mostram que, a diferença de intensidade do registro modal para o laringalizado foi o parâmetro mais modificado pela incidência da laringalização, registrando uma queda superior a 7dB entre as médias dos dois diferentes tipos fonatórios. Este achado pode sugerir a inspeção desta medida como uma evidência acústica relevante na identificação do fenômeno. Os resultados obtidos concordam e corroboram os achados da literatura pesquisada (DILLEY, SHATTUCK-HUFNAGEL & OSTENDORT, 1996; HANSON E CHUANG, 2001; REDI E SHATTUCK-HUFNAGEL, 2001; ESLING ET AL., 2005; GORDON E LADEFOGED, 2001, LIMA-GREGIO, 2011).

Dilley, Shattuck-Hufnagel & Ostendort (1996) e Redi e Shattuck-Hufnagel (2001) constataram na análise do fenômeno laríngeo que a diminuição na amplitude do sinal acústico esteve frequentemente associado à ocorrência da laringalização. Lima-Gregio (2012) por sua vez, evidenciou em seu trabalho uma queda de até 2 dB na transição do trecho modal para o trecho laringalizado, embora sem elicitar

significância estatística, foi parâmetro mais consistente na análise acústica inter e intra sujeitos.

Slifka (2006) buscou avaliar a ocorrência da fonação irregular e seus correlatos fisiológicos a partir da análise das extremidades dos enunciados no inglês americano. Nesse trabalho, foram avaliados 4 sujeitos, com idades ente 21 e 28 anos. Um conjunto de dados foram colhidos enquanto os sujeitos realizavam uma tarefa de leitura, consistindo no registro do sinal acústico, do fluxo de ar, dos diâmetros da caixa torácica e do abdômen e análise da pressão esofágica, a partir de uma metodologia que integralizou dados acústicos a respostas fisiológicas.

De acordo com Slifka (2006), a maioria das ocorrências da vibração irregular no final de enunciados mostram um aumento aparente na área mediana da glote durante alguns ou todos os ciclos de vibração das pregas vocais, o que leva ao acoplamento acústico entre os espaços supraglótico e infraglótico, refletindo o aumento das perdas de energia acústica.

Na pesquisa referida, para a maioria dos casos de fonação irregular nas extremidades dos enunciados, a pressão subglótica apresentou-se diminuída quando comparada aos demais contextos - geralmente menor que 5 cm H<sub>2</sub>O<sup>17</sup>. Observou-se nestas condições, que a área de repouso da glote está aumentada ou se desenvolvendo para um estado neutro (SLIFKA, 2006).

A diminuição do fluxo aéreo glótico, parece estar relacionado com as modificações nos parâmetros acústicos observados na ocorrência da laringalização em fronteira, no que diz respeito ao decréscimo na medida de frequência fundamental, alteração nos valores de jitter (indicando um padrão irregular de vibração das pregas vocais) e, sobretudo, a diminuição nos valores da medida de intensidade.

Para explicar a forma e o funcionamento da produção vocal Van Den Berg (1954) estabelece a Teoria Mioelástica-aerodinâmica, modelo de produção vocal mais aceito até o momento, o qual explica o processo de modulação vocal por meio da associação de três aspectos, o controle do estiramento, espessura e rigidez das pregas vocais, associado à energia cinética fornecida pelo fluxo aéreo subglótico,

---

<sup>17</sup> A pressão do ar nos pulmões, denominada de pressão alveolar ( $P_{ALV}$ ), é a soma da pressão pleural ( $P_{PL}$ ) e da pressão estática de recuo dos pulmões ( $P_L$ ), como representado pela equação:  $P_{ALV} = P_{PL} + P_L$ . Esta medida permite calcular as variações no fluxo aéreo em repouso e durante a fonação. O grau de constrição entre as pregas vocais e cavidade oral auxiliam na redução e no aumento da pressão alveolar (SLIFKA, 2006).

combinando forças de duas naturezas, a elasticidade dos músculos laríngeos e as forças físicas aerodinâmicas da respiração.

De acordo com Andrews (2009), pode-se sintetizar a função de cada um desses aspectos da seguinte forma: i) Forças musculares (mio): aduzem as pregas vocais, criando compressão medial na linha média; ii) Aerodinâmica: o ar exalado dos pulmões cria suficiente pressão subglótica para superar a compressão medial e deslocar as pregas vocais, criando um fluxo de ar que passa pelas pregas vocais; iii) Elástica: as pregas vocais recuam para sua posição de repouso, fechando-se na linha média antes de a pressão do ar forçá-las a se abrirem novamente.

Desta forma, a Teoria Mioelástica-aerodinâmica da fonação postula que a voz é produzida pelas contribuições combinadas da ação muscular, elasticidade do tecido e efeitos aerodinâmicos. As pregas vocais fecham-se no início da fonação pela ação dos músculos cricoaritenóideo lateral e interaritenóideo, causando a compressão medial das pregas vocais na linha média. Então, a pressão do ar subglótico aumenta e vence a resistência das pregas fechadas, e o ar é liberado para dentro do trato vocal, que vibra em toda a extensão do trato, onde as ondas sonoras são modificadas durante a transmissão. Então, as pregas vocais começam a fechar-se novamente, constringem a via aérea (e o ar que passa pelo espaço estreitado) e se juntam mais ainda por causa da ação sugadora, conhecida como efeito de Bernoulli (ANDREWS, 2009).

Para Andrews (2009), o controle da intensidade ocorre por meio da interação entre a pressão subglótica, a adução das pregas vocais e o equilíbrio entre harmônico-formante. Segundo o autor, quando qualquer um desses fatores aumenta ou diminui, também aumenta ou diminui a percepção da intensidade vocal, ou em termos perceptuais "*loudness*".

De acordo com Behlau (2001), a Teoria Mioelástica-aerodinâmica desvincula a frequência fundamental da intensidade, conseguindo lidar adequadamente com esses dois parâmetros. A produção da frequência fundamental é realizada pela contração dos músculos tensores, enquanto que a intensidade é produzida pela contração dos músculos adutores, por meio do controle das variações do fluxo aéreo.

Muitos parâmetros que controlam a fonação passam por um processo de transição no mesmo período de tempo em que os falantes fazem a passagem da postura de fala para a função respiratória, com o propósito de concluir a exalação (SLIFKA, 2006). Neste sentido, a fonação irregular ocorre a partir da combinação

particular dos mesmos parâmetros também envolvidos no controle fonético responsável pela finalização de enunciados - por vezes dando origem a fonação regular, por vezes dando origem a fonação irregular. Nas extremidades dos enunciados que terminam com uma respiração, esta última condição pode ser mais provável do que em outras partes do enunciado, uma vez que muitos dos parâmetros que controlam a fonação estão mudando no mesmo intervalo de tempo que a transição do evento de fonação para a respiração (SLIFKA, 2007).

Este argumento pode explicar a razão pela qual a fonação irregular é tão frequente em finais de unidades prosódicas maiores, como no término de frases entoacionais finais, quando comparadas às frases entoacionais mediais, como verificado por Dilley, Shattuck-Hufnagel & Ostendorf (1996) e Redi e Shattuck-Hufnagel (2001), no inglês americano. É possível que a influência de dois mecanismos distintos, um de base fisiológica e o outro prosódica, reforcem a ocorrência deste tipo de fonação alinhada ao término de unidades prosódicas maiores, como as frases entoacionais finais.

De acordo com Slifka (2007), uma vez que a fonação irregular não é lexicalmente contrastiva para as vogais no inglês americano, o falante não precisa evitar este tipo de fonação no final de uma enunciação (nem obrigatoriamente produzi-lo). Sabe-se, no entanto, que a ocorrência do fenômeno não se limita a unidades prosódicas superiores, como aquelas de frases entoacionais, por exemplo; o fenômeno tem sido amplamente referido como recorrente em unidades prosódicas que não necessariamente coincidem com grupos de respiração, como é o caso do grupo acentual e das frases entoacionais intermediárias (REDI & SHATTUCK-HUFNAGEL, 2001; LIMA-GREGIO, 2011), limitando o fator fisiológico enquanto condição única para que ocorra.

Regiões de silêncio no fluxo da fala são comumente produzidas por meio da compressão das pregas vocais ou através do distanciamento destas. Dado que a fonação irregular normalmente surge a partir de ambas estas ações, propõe-se que exista um papel de preferência para fonação irregular dentro dos sistemas fonológicos, assim, a fonação irregular associada ao silêncio na produção da fala funcionaria como uma pista para caracterizar um contraste segmental ou algum tipo de estrutura prosódica (SLIFKA, 2007).

Segundo Slifka (2007) a diminuição da periodicidade encontrada em inícios de frases talvez possa ser entendida como um processo de categorização fonológica do

enunciado, de forma a agrupá-lo em níveis mais baixos na hierarquia prosódica. Isto é, embora as restrições respiratórias na sonoridade só estejam presentes nos inícios e finais de enunciados, a vocalização irregular pode ser produzida por falantes nos inícios de todas as frases, mesmo em enunciados na posição medial, na qual não há diminuição do fluxo aéreo glótico pelo efeito da fisiologia vocal.

Deste modo, de acordo com a autora, o fenômeno pode ser perceptualmente motivado: a irregularidade fonatória serviria como uma pista perceptual referente às fronteiras frasais, motivada prosodicamente, como tem sido postulado para a ocorrência da laringalização na grande maioria dos estudos (SLIFKA, 2007).

## 5. CONCLUSÃO

As análises desenvolvidas neste trabalho possibilitaram responder as proposições inicialmente realizadas acerca da laringalização no PB. A partir da metodologia aplicada, foi possível entender também de que forma a ocorrência do fenômeno pode contribuir na diferenciação dos estilos de elocução e, conseqüentemente, na tarefa de comparação de locutores.

De modo geral, verificou-se que há uma variação consistente no que diz respeito as taxas de laringalização entre diferentes locutores. Dentre as medidas avaliadas, a taxa de laringalização vocálica e laringalização total foram as medidas que permitiram maior diferenciação de pares. Este achado aponta para uma maior consistência no que diz respeito à incidência da laringalização em segmentos vocálicos em detrimento dos segmentos consonantais.

Neste estudo, foi possível constatar comportamentos distintos entre diferentes locutores no que diz respeito a influência da variável sexo sobre as taxas de laringalização. O grupo feminino demonstrou uma tendência de apresentar maiores taxas de laringalização em comparação ao grupo masculino, corroborando a grande maioria dos estudos sobre laringalização.

No que diz respeito ao domínio prosódico, o alinhamento entre fronteiras de unidades prosódicas, a exemplo do grupo acentual, e a ocorrência da laringalização, revelou a importância da estrutura prosódica na motivação e restrição do fenômeno. Acredita-se que laringalização possa estar relacionada a algum tipo de organização discursiva, mediante o alinhamento referido e as modificações acústicas que acompanham o fenômeno.

As medidas de frequência fundamental, *jitter*, H1-H2 e intensidade foram os parâmetros acústicos mais influenciados pela ocorrência da laringalização, tendo sido esta última a medida mais modificada. Estes achados possibilitaram a caracterização fonético-acústica do fenômeno.

Os resultados obtidos contribuem para o avanço na ciência da fala e para o aperfeiçoamento dos estudos fonéticos de natureza forense, a partir da análise de aspectos prosódicos e de características indiossincráticas da fala. Neste sentido, a análise da laringalização configura um fenômeno da produção oral capaz de revelar diferentes estilos de fala, mediante variações nas taxas relacionadas, ao mesmo

tempo em que contribui para o entendimento da organização prosódica do português brasileiro.

## 6. REFERÊNCIAS

- ANDREWS, M. L. Manual de tratamemto da voz: da pediatria à geriatria. 7ª edição. São Paulo: Cengage Learning, 2009.
- BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos da Linguagem, Rev. Est. Ling., Belo Horizonte*, v. 20, n. 1, p. 11-27, jan./jun. 2012.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas: Pontes/Fapesp. 2006.
- BARBOSA, P. A.; MADUREIRA, S. **Manual de fonética acústica experimental**. São Paulo: Cortez, 2015.
- BARBOSA, P. A. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration. *Proceedings of the first ESCA Tutorial and Research Workshop on speech production modeling*. p. 85-88, 1996.
- BARBOSA, P. A. From syntax to acoustic duration: a dynamical model of speech rhythm production, *Speech Communication*, v. 49, p. 725-742, 2007.
- BEHLAU, M. *Voz o livro do especialista I*. São Paulo: Revinter, 2001.
- BELOTEL-GRENIÉ, A. & GRENIÉ, M. The Creaky Voice Phonation And The Organisation Of Chinese Discourse. *Proceedings of the International Symposium on Tonal Aspects of Languages*, 2004.
- BERRY, D. A. Mechanisms of modal and nonmodal phonation. *Journal of Phonetics*, 29, p. 431-450, 2001.
- BOERSMA, Paul & WEENINK, David (2014). Praat: doing phonetics by computer [Computer program]. Version 5.4, retrieved 4 October 2014 from <http://www.praat.org/>
- BOTINIS, ANTONIS, BJORN GRANSTRÖM, AND BERND MÖBIUS. "Developments and paradigms in intonation research." *Speech Communication* 33.4 : 263-296. Print, 2001.
- COLE, J.; SHATTUCK-HUFNAGEL, S. "The phonology and phonetics of perceived prosody: What do listeners imitate?". In: *Proceedings of the Interspeech 2011*. Florence, Italy, 21-31 August, p. 969-972, 2011.
- CRISTÓFARO, S. T. *Dicionário de fonética e fonologia*. ed. Contexto. São Paulo, 2011.
- DILLEY, L.; SHATTUCK-HUFNAGEL, S.; OSTENDORF, M. Glottalization of word-initial vowel as a function of prosodic structure. *Journal of Phonetics*, v. 24, p. 423-444, 1996.

Drugman, T., Kane, J., Raitio, T. & Gobl, C. Prediction of creaky voice from contextual factors. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 7967-7971, 2013.

EDMONDSON, J. A.; ESLING, J. H. The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies. *Phonology*. v. 23, p. 157-191, 2006.

ESLING, J. H.; FRASER, K. E.; HARRIS, J. G. Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics*, v. 33, p. 383-410, 2005.

ESLING, J. H.; HARRIS, J. G. States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In HARDCASTLE, W. J; BECK, J. (ed). *A festschrift for John Laver*. Mahwah, NJ: Lawrence Erlbaum Associates. p. 347-383. 2005.

FRENCH, P. An overview of forensic phonetics with particular reference to speaker identification. *Forensic Linguistics* 1: p. 169-181, 1994.

GARELLEK, M., 2011. Lexical effects on english vowel laryngealization. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, 2011.

GARELLEK, M., Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics. *UCLA Working Papers in Phonetics*, (110), p.1-23, 2012.

Gordon, M. & Ladefoged, P. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383-406, 2001.

GRØNNUM, N., Laryngealization or Pitch Accent – the Case of Danish Stød. *Proceedings of the 7th International Conference on Speech Prosody (SP2014)*, p.804-808, 2014.

HANSON HM, CHUANG ES. Glottal characteristics of male speakers: acoustic correlates and comparison with female data. *J Acoust Soc Am*. 1999.

HANSON, H. M., STEVENS, K. N., ST, G., HILL, M., CHEN, M. Y., & SLIFKA, J. Towards models of phonation Hong-Kwang Je4 Kuo. *Journal of Phonetics*, p. 451-480, 2001.

HANSON, H. M.; CHUANG, E. S. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *J. Acoust. Soc. Am*. v. 106, n. 2, p.1064-77, 2001.

HENTON, CAROLINE; ANTHONY BLADON. "Creak as a Sociophonetic Marker." in *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*, ed. Larry M. Hyman and Charles N. Li, 3–29. London: routledge, 1988.

HIRSON, A. Y DUCKWORTH, M. Glottal fry and voice disguise: A case study in forensic phonetics. *Journal of Biomedical Engineering*, 15, 193-208, 1993.

HOLLIEN, H. On vocal registers, *Journal of Phonetics*, 2, 125-143, 1974.

HOLLIEN, H. F. *The acoustics of crime: The new science of forensic phonetics*. New York: Plenum, 1990.

Ingle, Jennifer. *Pacific Northwest Vowels: A Seattle Neighborhood Dialecto Study*. Paper presented at the 149th Meeting of the Acoustical Society of America, Vancouver, B.C., May 16–20, 2005.

ISHI, CARLOS TOSHINORI. "Analysis of autocorrelation-based parameters in Creaky voice." *Acoustical Science and Technology* 25.4: 299-302. Print, 2004.

KENT, R.; READ, C. *The acoustic analysis of speech*. San Diego: Singular, 2007.

KREIMAN, J., Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics*, 10(March), p.163-175, 1982.

KÜNZEL, H. Effects of voice diguise on speaking fundamental frequency. *Forensic Linguistics*, p 149-179, 2000.

LADEFOGED, P. *Preliminaries to linguistic phonetics*. Chicago: University of Chicago, 1971.

Ladefoged, P. The linguistic use of different phonation types. In *vocal fold physiology: Contemporary research and clinical issues* (D. Bless & J. Abbs, editors), pp. 351-360. San Diego: College Hill Press, 1983.

LIMA-GREGIO AM. *Oclusiva glotal e laringalização em sujeitos com fissura palatina: um estudo segundo abordagem dinamicista [Tese]*. Campinas (SP): Universidade Estadual de Campinas; 2011.

LINDBLOM, BJÖRN. F0 lowering, creaky voice, and glottal stop: Jan Gauf-fin's account of how the larynx works in speech. *Proceedings, FONETIK 2009, Dept. of Linguistics, Stockholm University*, 2009.

LUCENTE, L. *Aspectos dinâmicos da entoação e da fala no português brasileiro. Tese (Doutorado em Linguística) - Unicamp, Campinas, SP, 2012.*

MATHUR S, CHOUDHARY BSK, VYAS CJM. Effect of Disguise on Fundamental Frequency of Voice. *J Forensic Res* 7:327. 2016. doi:10.4172/2157-7145.1000327.

MOOSMÜLLER, S. The influence of creaky voice on formant frequency changes. *Forensic Linguistics*, 8: p100-112, 2001.

MORRISON, G. Vowel inherent spectral change in forensic voice comparison. In G.Morrison & Assmann, P.(Eds.) *Vowel inherent spectral change* (pp. 263-283). Heidelberg, Germany: Springer-Verlag, 2013.

NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.

OLIVEIRA JR., MIGUEL; CRUZ, REGINA; SILVA, EBSON WILKERSON. **A relação entre a prosódia e a estrutura de narrativas es- pontâneas: um estudo perceptual**. *Revista Diadorim / Revista de Estudos Linguísticos e Literários do Programa de Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro*. Volume 12, Dezembro 2012. [<http://www.revistadiadorim.letas.ufrj.br>]

PITTAM, JEFFERY. "Listeners' Evaluations of Voice Quality in Australian English Speakers." *Language and Speech* 30: 99–113, 1987.

POMPINO-MARSCHALL, B. The syllable as a prosodic unit and the so-called p-centre effect. *Forschungsberichte der Institut für Phonetik und Sprachliche Kommunikation der Universität München*, v. 29, p. 66-124, 1991.

QUEIROZ, R. M. *Qualidade vocal: análise acústica de ajustes fonatório e de tensão laringe [Dissertação de mestrado]*. Pontifícia Universidade Católica de São Paulo - PUC, 2012.

REDI, L.; SHATTUCK-HUFNAGEL, S. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, v. 29, p. 407-29, 2001.

ROSE, F. *Forensic Speaker Identification*. London: Taylor & Francis, 2002.

SAN SEGUNDO, E., ALVES, H., & FERNÁNDEZ TRINIDAD, M. (2013). CIVIL Corpus: Voice quality for speaker forensic comparison. *Procedia - Social and Behavioral Sciences*. doi:10.1016/j.sbspro, 2013.

SELKIRK, E. on the major class features and syllable theory. In: ARONOFF; OEHRLE (org). *Language Sound Structurn* Cambridge, Mass: MIT Press, p. 107-136, 1984.

SLIFKA, J. Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice*, 20:171–186, 2006.

SLIFKA, J. Irregular phonation and its preferred role as cue to silence in phonological systems. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS 16)*, pages 229–232, Saarbrucken. Stevens, 2007.

SOBOTTA, J.; REINHARDS, P.; REINHARDS, P. *Atlas de Anatomia Humana: cabeça, pescoço e extremidade superior*. 22 ed. v. 1. Rio de Janeiro: Guanabara Koogan, 2006.

SZAKAY, A; TORGERSEN, E. N. An acoustic analysis of voice quality in London English: The effect of gender, ethnicity and f0. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, 2015.

TAJFEL, H. "Social identity and intergroup Behaviour." *Social Science Infor- mation* 13.2: 65–93, 1974.

TELLES, S. Traços laringais em Latundê (Nambikwára do Norte). Boletim do Museu Paraense Emílio Goeldi, Ciências Humanas 8. P. 291-306, 2013.

TORO, J. M. et al. The quest for generalizations over consonants: asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception & Psychophysics* 70 (8), p 1515-1525, 2008.

WATT, D; BURNS, J. Verbal descriptions of voice quality differences among untrained listeners.. In: York Papers in Linguistics Series 2, Vol. 12, 2012.

WISE, C. M. Introduction to phonetics. Englewood Cliffs, N.J.: Prentice-Hall, 1958.

YUASA, I. P. Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85.315-37, 2010.

## ANEXO A

## SCRIPT GLOTTALIZATION ANALYSIS

```

# Glottalisation analysis.psc
# Script implemented by Plinio A. Barbosa (IEL/Unicamp) for obtaining
#
# Do not distribute without the author's previous authorisation
# Credits:
# Parameters' input
# Copyright (C) 2008 Barbosa, P. A. (pabarbosa.unicampbr@gmail.com) #
# This program is free software; you can redistribute it and/or modify
# it under the terms of the GNU General Public License as published by
# the Free Software Foundation; version 2 of the License.
# This program is distributed in the hope that it will be useful,
# but WITHOUT ANY WARRANTY; without even the implied warranty of # MERCHANTABILITY or FITNESS FOR A PARTICULAR
PURPOSE. See the # GNU General Public License for more details.
#
form File acquisition word File LEITURA_KA_DJCFurso1.wav word Fileout LEITURA_KA_DJCFurso1.txt integer VvTier 1 integer
CvTier 2 integer AnnotationTier 3 integer SyntaxTier 4 choice Segmenttype 1 button V button C endform
Read from file... 'file$'
To Pitch (cc)... 0 75 15 no 0.03 0.45 0.01 0.35 0.14 300
To PointProcess select all
filename$ = selected$("Sound") select Sound 'filename$' To Intensity... 100 0.0 yes filegrid$ = filename$ + ".TextGrid" Read from file...
'filegrid$' filedelete 'fileout$'
fileappend 'fileout$' # VV | segment | syntax | sttime | phonation/side| segment dur | % creaky | % f0 drop | int drop (dB) | % dif jitter
'newline$' nintcvtier = Get number of intervals... 'cvTier' nintcvtier = nintcvtier - 1 ctVvant = 0
for i from 2 to nintcvtier select TextGrid 'filename$'
cvID$ = Get label of interval... 'cvTier' 'i'
if (cvID$ = "V") or (cvID$ = "v") or (cvID$ = "v ") or (cvID$ = "V ") cvID$ = "V" endif
if (cvID$ = "C") or (cvID$ = "c") or (cvID$ = "c ") or (cvID$ = "C ") cvID$ = "C" endif
if (cvID$ = segmenttype$)
start = Get starting point... 'cvTier' 'i' end = Get end point... 'cvTier' 'i' sidecreaky$ = "" j = start + 0.020
ctannotationant = Get interval at time... 'annotationTier' 'j' ctVV = Get interval at time... 'vvTier' 'j' vv$ = Get label of interval...
'vvTier' 'ctVV' ctSyntax = Get interval at time... 'syntaxTier' 'j' syntax$ = Get label of interval... 'syntaxTier' 'ctSyntax' loopend =
0 repeat if j <= end
ctannotation = Get interval at time... 'annotationTier' 'j' if ctannotationant <> ctannotation loopend = loopend + 1 if
loopend = 2 ct = ctannot endif
ctannot = ctannotationant endif
ctannotationant = ctannotation endif j = j + 0.020 until j >= end
if (loopend = 1 ) or (loopend = 2)
label$ = Get label of interval... 'annotationTier' 'ctannot'
phontypeant$ = mid$(label$,index(label$,"")+1,length(label$)- index(label$,""))
label$ = Get label of interval... 'annotationTier' 'ctannotation' phontype$ = mid$(label$,index(label$,"")+1,length(label$)-
index(label$,"")) sttimeant = Get starting point... 'annotationTier' 'ctannot' edtimeant = Get end point... 'annotationTier'
'ctannot' sttime = Get starting point... 'annotationTier' 'ctannotation' edtime = Get end point... 'annotationTier'
'ctannotation' durant = edtimeant - sttimeant dur = edtime - sttime if loopend = 2
stt2 = Get starting point... 'annotationTier' 'ct' edt2 = Get end point... 'annotationTier' 'ct' durant2 = edt2 - stt2 dur
= dur + durant2 sidecreaky$ = "/LR"
endif
totaldur = dur + durant select Pitch 'filename$' select PointProcess 'filename$' plus Sound 'filename$' Edit
editor PointProcess 'filename$'
Select... sttimeant edtime Zoom... sttimeant edtime
pause Corrija e depois pressione Continue Select... sttimeant edtimeant jitterant = Get jitter (local) Select... sttime
edtime jitter = Get jitter (local) Close endeditor

```

```

select PointProcess 'filename$'    To PitchTier... 0.02
mf0ant = Get mean (points)... 'sttimeant' 'edtimeant'    mf0 = Get mean (points)... 'sttime' 'edtime'    select Intensity 'filename$'
intant = Get mean... 'sttimeant' 'edtimeant' dB    int = Get mean... 'sttime' 'edtime' dB
if phontypeant$ = "m" and (phontype$ = "l" or phontype$ = "g" or phontype$ = "g*")    pcreaky = dur/totaldur
f0drop = 2*(mf0 - mf0ant)/(mf0 + mf0ant)    intdrop = int - intant    if loopend = 1    sidecreaky$ = "/R"    endif
djitter = 2*(jitter - jitterant)/(jitter + jitterant)    endif
if (phontypeant$ = "l" or phontype$ = "g" or phontype$ = "g*") and phontype$ = "m"    pcreaky = durant/totaldur
f0drop = 2*(mf0ant - mf0)/(mf0 + mf0ant)    intdrop = intant - int    if loopend = 1    sidecreaky$ = "/L"    endif
djitter = 2*(jitterant - jitter)/(jitter + jitterant)    endif    else
totaldur = end - start
label$ = Get label of interval... 'annotationTier' 'ctannotation'    phontype$ = mid$(label$,index(label$,"")+1,length(label$)-
index(label$,""))    if phontype$ = "l" or phontype$ = "g" or phontype$ = "g*"    pcreaky = 1    f0drop = 0    intdrop =
0    djitter = 0    else    pcreaky = 0    f0drop = 0    intdrop = 0    djitter = 0    endif    endif
totaldur = totaldur*1000
segment$ = left$(label$,index(label$,"")-1) if ctVVant = ctVV    nl$ = "||"    elseif ctVVant = 0    nl$ = ""    else    nl$ = newline$    endif
fileappend 'fileout$' 'nl$' 'vv$' | 'segment$' | 'syntax$' | 'start:3' | 'phontype$'sidecreaky$' |
'totaldur:0' | 'pcreaky:2' | 'f0drop:2' | 'intdrop:2' | 'djitter:3' ctVVant = ctVV    endif endfor

```

## ANEXO B

## SCRIPT SG DETECTOR

```

# SGdetector.psc
# Script implemented by Plinio A. Barbosa (IEL/Unicamp) for detecting # stress group boundaries from production
criteria, namely VV durations. Input: previously # segmented
# VV intervals (TextGrid)
# Please, do not distribute without the author's previous authorisation
# The TextGrid and Reference-statistics (zaldo.TableOfReal) files need to be in the
# same directory!!! # Date: May 27th, 2004 form Aquisição dos arquivos
word Arquivo_(TextGrid) LEITURA_KA_DJCFurso1 integer VvTier 1 choice Referencia: 1 button Zaldo
endform
# Lê o arquivo de referencia com as triplas (segmento, média, desvio-padrão) do locutor
# Referencia. A variável nseg contém o número total de segmentos do arquivo de referência
Read from file... 'referencia$'.TableOfReal nseg = Get number of rows
#
# Lê arquivo e TextGrid (desde q tenha o mesmo nome do arquivo de som arq$ = arquivo$ + ".TextGrid" Read
from file... 'arq$' begin = Get starting time end = Get finishing time
nselected = Get number of intervals... 'vvTier' nselected = nselected - 2 arqout$ = arquivo$ + "dur" + ".txt"
filedelete 'arqout$'
arqoutstrgrp$ = arquivo$ + "SG" + ".txt" filedelete 'arqoutstrgrp$'
fileappend 'arqout$' % Segmentos acusticos, duracao (ms) , z, z suav., fronteira
'newline$'
fileappend 'arqoutstrgrp$' % Duração do grupo acentual | Número de sílabas 'newline$' select TextGrid
'arquivo$'
initialtime = Get starting point... 'vvTier' 2 for i from 1 to nselected adv = i + 1
nome$ = Get label of interval... 'vvTier' 'adv' itime = Get starting point... 'vvTier' 'adv' ftime = Get end point...
'vvTier' 'adv' dur = ftime - itime dur = round(dur*1000) call zscorecomp 'nome$' 'dur' dur'i' = dur z'i' = z nome'i'$
= nome$ adv = i + 1
select TextGrid 'arquivo$' endfor smz1 = (2*z1 + z2)/3 deriv1 = smz1 smz2 = (2*z2 + z1)/3 deriv2 = smz2 - smz1
i = 3 if smz1 < smz2 minsmz = smz1 maxsmz = smz2 else minsmz = smz2 maxsmz = smz1
endif
while i <= (nselected-2) del1 = i - 1 del2 = i - 2 adv1 = i + 1 adv2 = i + 2
smz'i' = (5*z'i' + 3*z'del1' + 3*z'adv1' + z'del2' + 1*z'adv2')/13 deriv'i' = smz'i' - smz'del1' if smz'i' <
minsmz minsmz = smz'i' endif if smz'i' > maxsmz maxsmz = smz'i' endif i = i + 1 endwhile tp1 = nselected - 1
tp2 = nselected - 2
smz'tp1' = (3*z'tp1' + z'tp2' + z'nselected')/5 deriv'tp1' = smz'tp1' - smz'tp2' if smz'tp1' < minsmz minsmz =
smz'tp1' endif if smz'tp1' > maxsmz maxsmz = smz'tp1' endif
smz'nselected' = (2*z'nselected' + z'tp1')/3 deriv'nselected' = smz'nselected' - smz'tp1' if smz'nselected' <
minsmz minsmz = smz'nselected' endif
if smz'nselected' > maxsmz maxsmz = smz'nselected' endif
tempfile$ = "temp.TableOfReal" filedelete 'tempfile$'
fileappend 'tempfile$' File type = "ooTextFile short" 'newline$' fileappend 'tempfile$' "TableOfReal" 'newline$'
fileappend 'tempfile$' 'newline$' fileappend 'tempfile$' 2 'newline$' fileappend 'tempfile$' columnLabels
[]: 'newline$' fileappend 'tempfile$' "position" "smoothed z" 'newline$' tpp = nselected + 2
fileappend 'tempfile$' 'tpp' 'newline$' time = initialtime
fileappend 'tempfile$' row[1]: "0" 0.0 0.0 'newline$' boundcount = 0 sdur = 0 sssl = 0 sdurSG = 0 svar = 0
for i from 1 to nselected tempismz = smz'i' tpnome$ = nome'i'$ adv1 = i + 1 btime'i' = 0 time = time +
dur'i'/1000

```

```

fileappend 'tempfile$' row['adv1']: "'tpnome$'" 'time' 'tempzmz' 'newline$' if i <> nselected  adv1 = i + 1
if (deriv'i' >= 0) and (deriv'adv1' < 0)  boundary = 1  boundcount = boundcount + 1
  btime'i' = time  bctime'boundcount' = time  else  boundary = 0  endif  else  del1 = i - 1  if smz'i' >
smz'del1'  boundary = 1
  boundcount = boundcount + 1
  btime'i' = time  bctime'boundcount' = time  else  boundary = 0  endif  endif  tempz = z'i'  tempdur =
dur'i'  sdur = sdur + tempdur  sdurSG = sdurSG + tempdur  ssyl = ssyl + 1
fileappend 'arqout$' 'tpnome$' 'tempdur' 'tempz:2' 'tempzmz:2' 'boundary' 'newline$' if boundary == 1
fileappend 'arqoutstrgrp$' 'sdurSG' 'ssyl' 'newline$'  sdurSG = 0  ssyl = 0  endif  endfor
meandur = sdur/nselected for i from 1 to nselected svar = svar + (dur'i' - meandur)^2  endfor
stddevdur = sqrt(svar/(nselected - 1))
fileappend 'arqout$' Média e desvio-padrão: 'meandur:0' 'stddevdur:0' 'newline$'  tp = i+1
fileappend 'tempfile$' row['tp']: "X" 'end' 0 'newline$'  select all
Remove
Read from file... 'tempfile$'
Draw scatter plot... 1 2 0 0 0.0 0.0 0.0 0.0 12 no + yes select all
Remove Red
for i from 1 to nselected - 1  if btime'i' <> 0  bt = btime'i'
  Draw line... 'bt' 'minsmz' 'bt' 'maxsmz'  endif  endfor
Black
# Write a TextGrid with the stress group boundaries fileout$ = arquivo$ + "2.TextGrid"  filedelete 'fileout$'
fileappend 'fileout$' File type = "ooTextFile short" 'newline$'  fileappend 'fileout$' "TextGrid" 'newline$'  fileappend
'fileout$' 'newline$'  fileappend 'fileout$' 'begin' 'newline$'  fileappend 'fileout$' 'end' 'newline$'  fileappend 'fileout$'
<exists> 'newline$'  fileappend 'fileout$' 1 'newline$'  fileappend 'fileout$' "IntervalTier" 'newline$'  fileappend
'fileout$' "StressGroups" 'newline$'  fileappend 'fileout$' 'begin' 'newline$'  fileappend 'fileout$' 'end' 'newline$'  tmp =
boundcount + 2
fileappend 'fileout$' 'tmp' 'newline$'  fileappend 'fileout$' 0.00 'newline$'  fileappend 'fileout$' 'initialtime' 'newline$'
fileappend 'fileout$' "" 'newline$'  temp = initialtime  for i from 1 to boundcount  fileappend 'fileout$' 'temp'
'newline$'  temp = bctime'i'
fileappend 'fileout$' 'temp' 'newline$'  fileappend 'fileout$' "" 'newline$'  endfor
fileappend 'fileout$' 'temp' 'newline$'  fileappend 'fileout$' 'end' 'newline$'  fileappend 'fileout$' "" 'newline$'
##
arqgrid1$ = arquivo$ + ".TextGrid"  arqgrid2$ = fileout$  Read from file... 'arqgrid1$'  Read from file... 'arqgrid2$'
select all
Merge ##
procedure zscorecomp nome$ dur  sizeunit = length (nome$)  sumofmeans = 0  sumofvar = 0  cpt = 1  while cpt
<= sizeunit  nb = 1  terminate = 0  k = 1
  seg$ = mid$(nome$,cpt,1)  if cpt < sizeunit
    if mid$(nome$,cpt+1,1) == "h" or mid$(nome$,cpt+1,1) == "N"      nb = nb + 1
    seg$ = seg$ + mid$(nome$,cpt+1,1)  endif
    if (cpt+nb <= sizeunit)  tp$ = mid$(nome$,cpt,1)  call isvowel 'tp$'
    if ((mid$(nome$,cpt+nb,1) = "I") or (mid$(nome$,cpt+nb,1) = "U")) and truevowel
      seg$ = seg$ + mid$(nome$,cpt+nb,1)  nb= nb+1  endif  endif  endif  j = 1  select all
  tableID = selected ("TableOfReal")  select 'tableID'
  while (j <= nseg) and not terminate  label$ = Get row label... 'j'
    if seg$ = label$  terminate = 1  mean = Get value... 'j' 1  sd  = Get value... 'j'
2  sumofmeans = mean + sumofmeans  sumofvar= sd*sd + sumofvar  endif  j = j+1  endwhile  if
not terminate
  exit Não encontrou o fone 'seg$'. Corrija o arquivo TableOfReal  endif  cpt= cpt+nb  endwhile

```

```
z = (dur - sumofmeans)/sqrt(sumofvar) endproc
procedure isvowel temp$ truevowel = 0
if temp$ = "i" or temp$ = "e" or temp$ = "a" or temp$ = "o" or temp$ = "u" or temp$ = "I" or temp$ = "E"
...or temp$ = "A" or temp$ = "O" or temp$ = "U" truevowel = 1 endif endproc
```