

ESCOLA DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
DOUTORADO

MÁRCIO OPPLIGER PINTO

**ÍNDICE DE VARIABILIDADE - UM CRITÉRIO DE AVALIAÇÃO DE PARÂMETROS
ACÚSTICOS DE VOGAIS ORAIS E CONSOANTES NASAIS COMO ELEMENTO PARA
A PERÍCIA DE COMPARAÇÃO DE LOCUTORES.**

Porto Alegre
2017

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

MÁRCIO OPPLIGER PINTO

ÍNDICE DE VARIABILIDADE – UM CRITÉRIO DE AVALIAÇÃO DE PARÂMETROS
ACÚSTICOS DE VOGAIS ORAIS E CONSOANTES NASAIS COMO ELEMENTO
PARA A PERÍCIA DE COMPARAÇÃO DE LOCUTORES.

Tese apresentada como requisito para obtenção
do grau de Doutor em Linguística pelo Programa
de Pós-Graduação em Letras da Pontifícia
Universidade Católica do Rio Grande do Sul.

Orientadora: Prof^a Dra. Cláudia Regina Brescancini

Porto Alegre

2017

Ficha Catalográfica

P659 Pinto, Márcio Oppliger

Índice de Variabilidade – um critério de avaliação de parâmetros acústicos de vogais orais e consoantes nasais como elemento para a Perícia de Comparação de Locutores / Márcio Oppliger Pinto . – 2017.

232 f.

Tese (Doutorado) – Programa de Pós-Graduação em Letras, PUCRS.

Orientadora: Profa. Dra. Cláudia Regina Brescancini.

1. Fonética Acústica e Articulatória. 2. Fonética Forense. 3. Variabilidade Individual. 4. Comparação de Locutores. 5. Fonologia Gestual. I. Brescancini, Cláudia Regina. II. Título.

MÁRCIO OPPLIGER PINTO

ÍNDICE DE VARIABILIDADE – UM CRITÉRIO DE AVALIAÇÃO DE PARÂMETROS
ACÚSTICOS DE VOGAIS ORAIS E CONSOANTES NASAIS COMO ELEMENTO
PARA A PERÍCIA DE COMPARAÇÃO DE LOCUTORES.

Tese apresentada como requisito para obtenção
do grau de Doutor em Linguística pelo Programa
de Pós-Graduação em Letras da Pontifícia
Universidade Católica do Rio Grande do Sul.

Aprovada em: 16 de outubro de 2017.

BANCA EXAMINADORA:

Prof^a. Dr^a. Cláudia Regina Brescancini (Presidente)
PUCRS

Prof. Dr. Alexandro Rodrigues Meireles
UFES

Prof. Dr. Ubiratã Kickhöfel Alves
UFRGS

Prof. Dr. Dênis Fernandes
PUCRS

Prof. Dr. Lorí Viali
PUCRS

Porto Alegre
2017

Às mulheres da minha vida,
Milena, Mariá e Waleska,
e aos meus pais, Margarida e Geruzes.
Vocês me permitem compreender o amor.

AGRADECIMENTOS

Antes e acima de qualquer outra menção, devo um profundo agradecimento à Professora Doutora Cláudia Regina Brescancini, modelo de dedicação, comprometimento e seriedade. O convívio durante quase quatorze anos sempre foi marcado por admiração e respeito de minha parte e por uma atenção sempre prestada com zelo às minhas ideias, argumentações e principalmente aos meus defeitos, que não são poucos. Encerro essa etapa de vida lamentando um pouco a “perda” da orientadora, mas alegre por ter a convicção de haver estabelecido uma relação a qual me sinto à vontade para reconhecer como de amizade.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior por haver me concedido bolsa em um primeiro momento do doutorado e ao Instituto-Geral de Perícias do Rio Grande do Sul por haver me contemplado com bolsa em um segundo momento até a conclusão desse processo.

Ao Programa de Pós Graduação em Letras agradeço pela maneira acolhedora que sempre me foi dedicada em todos os momentos em que necessitei de apoio institucional. Uma menção especial deve ser feita às funcionárias administrativas que sempre me atenderam com carinho e dedicação, sejam as atuais, Tati e Alessandra, como a Isabel e, com um sentimento especial, a saudosa Mara, que esteja descansando em paz em algum plano do Universo.

Faço menção também às colegas do Departamento de Criminalística, Peritas Criminais e Fonoaudiólogas Aline Costa Porto e Cíntia Schivinski Gonçalves, que colaboraram demasiado para a promoção da Fonética Forense no Brasil e sempre estiveram muito próximas das minhas angústias acadêmicas. Agradeço aos colegas, atuais e aos que já não estão mais, da Seção de Perícias em Áudio e Imagens (SEPAI), que sofreram de alguma maneira com a minha necessidade de conciliar o trabalho com a pesquisa. Em especial, agradeço à Perita Bárbara Zaffari Cavedon por colaborar como substituta e muito além do que seria esperado para essa atribuição, por ocasião em que assumi a chefia da SEPAI. A sua generosidade, em momento que as tarefas do doutorado me demandavam intensamente, foi fundamental para a conclusão desta tese.

Agradeço também aos colegas do Grupo de Estudos em Processamento de Áudio e Fonética Acústica, o GEPAFA, pelas interessantes discussões e pelo belo trabalho desenvolvido. É uma satisfação colaborar com esse grupo.

Por fim, agradeço à família, que me acolhe nos momentos difíceis e com quem irei celebrar a conquista, e ao Universo, que aparentemente conspira a meu favor. Muito obrigado!

Tristeza mora comigo por causa da solidão.
Eu pareço uma andorinha querendo fazer verão,
Uma gota de água doce querendo ser ribeirão,
Uma semente caída querendo ser plantação.
Mas, olhando pr'o deserto, eu sou apenas um grão
de areia, um grão de areia um grão...
(Abadá capoeira)

RESUMO

A presente pesquisa buscou estabelecer um indicador de variabilidade de parâmetros acústicos utilizados como elementos de confronto no exame pericial de Comparação de Locutores. O Índice de Variabilidade proposto por este estudo considera a relação entre variabilidade intraindividual e interindividual e pretende-se que sirva como referência para determinar a vantagem de um parâmetro acústico sobre outro dentre aqueles utilizados na Comparação de Locutores. Além de ser obtido a partir dos valores dos formantes vocálicos, parâmetros tradicionais na Comparação de Locutores, o Índice de Variabilidade foi obtido com base nos formantes das consoantes nasais e a partir das medidas de espaço acústico vocálico e distâncias euclidianas entre vogais, com o intuito de verificar a aplicabilidade desses parâmetros na Perícia de Comparação de Locutores. Os estímulos que serviram de base para o estudo foram sílabas tônicas pronunciadas em frase veículo e a coleta, realizada repetidamente em dias e turnos diferentes, e visou o controle da variabilidade, favorecendo uma variabilidade mínima dos parâmetros acústicos. Quanto à metodologia, vinte indivíduos participaram da coleta e os coeficientes de variação individual médio do conjunto e geral do conjunto serviram como referentes de variabilidade intraindividual e interindividual respectivamente. Os resultados de variabilidade, apesar de não terem sido bem-sucedidos no que diz respeito às consoantes nasais, sugerem que o Índice de Variabilidade permite estabelecer quais os melhores parâmetros para serem utilizados na Comparação de Locutores. O Índice permitiu ordenar as vogais e também embasou a proposta de que as medidas de espaço acústico e de Distâncias Euclidianas são possíveis parâmetros a serem utilizados na Comparação de Locutores. Diferentemente da Fonética Acústica e Articulatória, que fundamenta a pesquisa, a Fonologia Articulatória compõe o trabalho como aporte teórico acessório em relação a interpretação dos dados, porém entende-se que seja teoria valiosa para explicação dos dados linguísticos que usualmente encontram-se na prática forense. As Distâncias Euclidianas entre os vértices do triângulo vocálico – [a], [i] e [u] – serviram para propor uma análise de como a implementação computacional do modelo da Dinâmica de Tarefas, o TADA, trata vogais de forma contínua e consoantes de forma categórica.

ABSTRACT

The present research aimed to establish an indicator of variability of acoustic parameters used as elements of confrontation in the expert examination of Voice Comparison. The Variability Index proposed here considers the relationship between intraindividual and interindividual variability and is intended to serve as a reference to determine the advantage of one acoustic parameter over another among those used in Voice Comparison. Besides being obtained from the values of the vowel formants, traditional parameters in the Voice Comparison, the Variability Index was obtained from the nasal consonant formants and from the measurements of vocalic acoustic space and Euclidean Distances between vowels, with the aim to verify the applicability of these parameters in Voice Comparison examination. The stimulus that served as the basis for the study were tonic syllables pronounced in a carrier sentence and the data collection, performed repeatedly in different days and shifts, and aimed at the control of variability, favoring a minimum variability of the acoustic parameters. As for the method, twenty subjects participated in the data collection and the individual average coefficients of variation of the set and general average coefficients of variation of the set served as referents of intraindividual and interindividual variability respectively. As the results of variability, although were not achieved with nasal consonants, suggest that the Variability Index allows us to establish the best parameters to be used in the Voice Comparison. The Index allowed to order the vowels and also supported the proposal that the measurements of vocal acoustic space and Euclidean Distances are possible parameters to be used in Voice Comparison. Unlike Acoustic Phonetics and Articulatory, underlying the research, Articulatory Phonology composes the work as an accessory theoretical contribution in relation to the interpretation of the data of the present doctoral dissertation, but it is understood that it is a valuable theory to explain the linguistic data usually found in forensic practice. The Euclidean Distances between the vertices of the vocalic triangle - [a], [i] and [u] - served to propose an analysis of how the computational implementation of the Task Dynamics model, TADA, treats vowels in a continuously way and consonants as a categorical form.

LISTA DE FIGURAS

Figura 2.1 – Circuito de conversa entre dois indivíduos numa perspectiva unidirecional.	29
Figura 2.2 – Circuito da fala sob uma perspectiva multimodal de <i>feedback</i>	31
Figura 2.3– Ilustração da Teoria Fonte-Filtro	34
Figura 2.4 – Espectrograma da frase “Ele dirá MA calmamente”.....	36
Figura 2.5 – Relação fonte/filtro para [ə].	37
Figura 2.6 – Relação LPC/FFT para determinação de formantes	38
Figura 2.7 – Ilustração do conceito de largura de banda	39
Figura 2.8 – Derivação do trato vocal em função de área.....	42
Figura 2.9 – Configurações ilustrativas do trato vocal.....	43
Figura 2.10 – Acústica do trato vocal, seus nodos e antinodos.....	44
Figura 2.11 – Relação de $F_1 \times F_2$ das vogais e os respectivos formatos de tubos ressonadores	45
Figura 2.12 – Posição de língua para a realização de três vogais.	46
Figura 2.13 – Espaço vocálico segundo posição do ponto mais alto do corpo da língua.	47
Figura 2.14 – Trapézio com o alfabeto fonético das vogais.	48
Figura 2.15 – Representações do espaço acústico.....	49
Figura 2.16 – VC com os respectivos valores para F_1 e F_2 (valores na tabela abaixo).50	
Figura 2.17 - Comparação dos sistemas de vogais (cardeais, PE e PB).	57
Figura 2.18 - Comparação das distribuições de vogais (entre os dois gêneros e entre o PB e PE).	58
Figura 2.19 - Comparação dos sistemas de vogais.	59
Figura 2.20 - Comparação dos sistemas de vogais.	62
Figura 2.21 - Comparação de oclusivas nasal e oral com mesmo ponto de articulação.	67
Figura 2.22 – Mecanismo aerodinâmico do trato vocal.	68
Figura 2.23 – Relação da constrição supraglotal e abertura vélica.	71

Figura 2.24 – Comportamento das ressonâncias no início da nasal (“implosion”, no inglês).....	72
Figura 2.25 – Comportamento das ressonâncias na soltura de [m].	72
Figura 2.26 – Comportamento das ressonâncias na soltura de [n].	73
Figura 2.27 – Exemplo de cotejo utilizado em laudo de CL.	90
Figura 2.28 – Exemplo de cotejo utilizado em laudo de CL.	92
Figura 3.1 – Trajetória do lábio inferior para a sequência [abə].	101
Figura 3.2 – Efetores finais, variáveis do trato e articuladores envolvidos.	103
Figura 3.3 – Inventário dos conjuntos articuladores e seus parâmetros.	105
Figura 3.4 – Pauta gestual da palavra “camper”	107
Figura 3.5 – Pauta gestual da palavra “palm”.....	108
Figura 3.6 –Modelo gestual computacional.....	110
Figura 3.7 – Pauta gestual para a palavra <i>span</i> : modelo gestual computacional.....	112
Figura 3.8 – Parcial da Tabela I.2 de “gestos atribuídos a segmentos”: Oclusivas.....	114
Figura 3.9 – Parciais da tabela de “gestos atribuídos a segmentos”: Vogais.....	115
Figura 4.1 – Exemplo de etiquetagem utilizada. Janela de edição do Praat.	122
Figura 4.2 – Exemplo de janela do <i>script</i> utilizado sobre a janela de edição do Praat..	124
.....	124
Figura 5.1 – Comparação de pesquisas do PB.....	131
Figura 5.2 – Dispersão das medidas de consoantes nasais [me, mɛ, mɔ, mu] – Participante EGS	150
Figura 5.3 – Dispersão das medidas de consoantes nasais – Participante MFS	153
Figura 5.4 – Dispersão das medidas de consoantes nasais – Participante AFRB	154

LISTA DE QUADROS

Quadro 2.1 – F1 e F2 para VC, PE e PB.	56
Quadro 2.2 – Visão geral dos trabalhos em acústica de vogais.....	64
Quadro 2.3 – Vogais cardeais e resumo dos valores médios de vogais dos trabalhos referidos.....	65
Quadro 2.4 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [a] com suas respectivas larguras de banda. [Hz]	76
Quadro 2.5 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [i] com suas respectivas larguras de banda. [Hz]	76
Quadro 2.6 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [u] com suas respectivas larguras de banda. [Hz]	76
Quadro 2.7 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [a] com suas respectivas larguras de banda. [Hz]	77
Quadro 2.8 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [i] com suas respectivas larguras de banda. [Hz]	77
Quadro 2.9 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [u] com suas respectivas larguras de banda. [Hz]	77
Quadro 4.1 – Idades e níveis de escolaridade dos participantes.....	118
Quadro 4.2 – Número de sessões de gravação por participante.	125
Quadro 5.1 – Comparação com trabalhos que mediram F1 e F2.	131
Quadro 5.2 – Comparação com trabalhos que incluíram f0 e F3, além de F1 e F2.....	132
Quadro 5.3 – Diferença em percentuais deste estudo para aqueles de Rauber (2006) e Escudero et al. (2009).....	133
Quadro 5.4 – Resultados gerais de média e variabilidade (desvio padrão, mínimo, máximo, amplitude, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas aparece na segunda coluna.	136
Quadro 5.5 – Ordenamento dos sons vocálicos por parâmetro acústico.	137
Quadro 5.6 – Ordenamento por classe do contexto anterior.....	141
Quadro 5.7 – Ordenamento por consoante contexto anterior - nasais.....	142

Quadro 5.8 – Ordenamento por consoante contexto anterior – fricativas desvozeadas	143
Quadro 5.9: Resultados gerais de média e variabilidade do VSA e distâncias euclidianas (desvio padrão, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas aparece na segunda coluna.	146
Quadro 5.10 – Estatística descritiva resumida das consoantes nasais.....	147
Quadro 5.11 – Comparação com os achados de Seara (2000).	148
Quadro 5.12 – Número de participantes por contexto selecionado.....	155
Quadro 5.13 – Resultados gerais para as consoantes nasais, estratificadas em bilabial e alveolar, expressos pela média e variabilidade (desvio padrão, amplitude, mínimo, máximo, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas aparece na segunda coluna.	155
Quadro 5.14 – Resultados gerais para as consoantes nasais, estratificadas por sílaba, expressos pela média e variabilidade (desvio padrão, mínimo, máximo, amplitude, coeficiente de variação interindividual, coeficiente de variação intraindividual e a índice de variabilidade).	156
Quadro 5.15 – Ordenamento das sílabas medidas por IV.	161
Quadro 5.16 – Similaridade na CL – valores de CVintra.	163

LISTA DE ABREVIATURAS E SIGLAS

ATR	Raiz da língua avançada (traço fonológico)
B1	Largura de banda do primeiro formante
B2	Largura de banda do segundo formante
B3	Largura de banda do terceiro formante
B4	Largura de banda do quarto formante
B5	Largura de banda do quinto formante
CL	Comparação de locutores
CV	Coefficiente de variação
CV _{inter}	Coefficiente de variação interindividual
CV _{intra}	Coefficiente de variação intraindividual
DE	Distância euclidiana
DE _{au}	Distância euclidiana entre [a] e [u]
DE _{ia}	Distância euclidiana entre [i] e [a]
DE _{iu}	Distância euclidiana entre [i] e [u]
f ₀	Frequência fundamental
F1	Primeiro formante
F2	Segundo formante
F3	Terceiro formante
FCR	Razão de centralização formântica
FFT	Transformada rápida de Fourier
FN1	Primeiro formante nasal
GLO	Abertura glotal
IPA	Associação Fonética Internacional
IV	Índice de variabilidade
LA	Abertura labial
LnVSA	Área logarítmica de espaço vocálico
LP	Protrusão labial

LPC	Codificação preditiva linear
MRI	Imagens de ressonância magnética
PB	Português brasileiro
PE	Português europeu
TADA	Aplicação da Dinâmica de Tarefa (Task Dynamic Application)
TBCD	Grau de constrição do corpo da língua
TBCL	Local de constrição do corpo da língua
TTCD	Grau de constrição da ponta da língua
TTCL	Local de constrição da ponta da língua
VC	Vogais cardeais
VEL	Abertura vélica
VSA	Área de espaço vocálico

SUMÁRIO

1	INTRODUÇÃO	17
2	FONÉTICA ACÚSTICA E ARTICULATÓRIA	28
2.1	CONSIDERAÇÕES GERAIS: PRODUÇÃO E PERCEPÇÃO	28
2.2	FUNDAMENTOS TEÓRICOS	32
2.3	VOGAIS	40
2.3.1	Espaço acústico como medida aplicada	52
2.3.2	Descrições do espaço vocálico do PB	54
2.4	CONSOANTES NASAIS	66
2.5	APLICAÇÃO FORENSE	79
2.5.1	Introdução	79
2.5.2	As vogais como parâmetros de identificação	81
2.5.3	As nasais como parâmetros de identificação	85
2.5.4	Relação entre a variabilidade inter e intraindividual	89
3	FONOLOGIA ARTICULATÓRIA	96
3.1	A CONTESTAÇÃO DA UNIDADE DE ANÁLISE TRADICIONAL	97
3.2	O GESTO ARTICULATÓRIO E SUA IMPLEMENTAÇÃO	100
3.3	AS VOGAIS NA FONOLOGIA ARTICULATÓRIA	111
4	METODOLOGIA	117
4.1	AMOSTRA E SELEÇÃO DE PARTICIPANTES	117
4.2	COLETA DE DADOS	118
4.2.1	Equipamentos e ambiente de gravação	120
4.3	ANÁLISE ACÚSTICA: INSTRUMENTO E PROCEDIMENTO	121
4.3.1	Preparação, denominação e etiquetagem dos arquivos	121
4.3.2	Script para medição	123
4.4	VARIÁVEIS	124
4.5	RECURSOS PARA A ANÁLISE ESTATÍSTICA	125
5	APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS	128
5.1	VOGAIS	129

5.1.1	Consideração sobre os efeitos do contexto anterior à vogal	140
5.1.2	Espaço acústico.....	145
5.2	CONSOANTES NASAIS	147
5.3	DISCUSSÃO	157
5.3.1	Aplicação forense	158
5.3.2	Tratamento teórico.....	165
5.3.2.1	A comensurabilidade entre Fonética e Fonologia e a prática forense	165
5.3.2.2	A situação peculiar das vogais em relação ao modelo teórico.....	167
6	CONSIDERAÇÕES FINAIS	174
	REFERÊNCIAS.....	178
	APÊNDICE A – Termo de consentimento livre e esclarecido	188
	APÊNDICE B – Termo de compromisso para utilização de dados	190
	APÊNDICE C – Instrumento de coleta	191
	APÊNDICE D - Gráficos de distribuição dos formantes das consoantes nasais Participante ACRR	192
	ANEXO A – Alfabeto ARPABET retirado do manual do TADA*	210
	ANEXO B – Tabela I.2 retirada do manual do TADA	211
	ANEXO C – <i>Script</i> utilizado para a medição dos dados	215

1 INTRODUÇÃO

No universo da criminalística é corriqueira a análise de vestígios materiais que indiquem a autoria de um delito ou a identificação de uma vítima. As possibilidades de exame pelo perito podem recair sobre uma impressão digital encontrada na empunhadura de uma arma ou próxima a uma janela arrombada, um fio de cabelo ou resíduo de sangue, encontrados no porta-malas do automóvel de um suspeito de homicídio, resíduos de epitélio, encontrados sob as unhas de uma vítima morta em um contexto de embate corporal e até mesmo o registro em vídeo de um rosto ou fisionomia de indivíduo envolvido em um assalto a um caixa eletrônico. Todos são exemplos de objeto de exame na área da identificação humana para fins forenses. O que eles têm em comum é o fato de fornecerem características anátomo-biológicas que são constantes em cada ser humano e diferentes entre os seus pares, permitindo que se definam aquelas características que são particulares ao indivíduo sob investigação. Utilizar indícios de elementos anatômicos é estratégia natural e foi a base da identificação na criminalística, quando Alphonse Bertillon, no final do século XIX, desenvolveu o Sistema Antropométrico para a identificação humana. Estratégias de identificação através de características físicas ou comportamentais compõem a área da Identificação Biométrica, sendo que uma dessas estratégias se dá através da análise da fala e da voz¹ dos seres humanos.

A manifestação verbal oral diferencia o ser humano das demais espécies e pode servir como elemento de particularização de um indivíduo a ponto de torná-lo único entre os indivíduos de sua espécie. A atividade de reconhecimento através de elementos da comunicação oral fundamenta-se na ideia de que, ao comunicar-se oralmente, o indivíduo transmite informações que permitem caracterizá-lo com relação à sua formação cultural e social, constituição orgânica, região de origem e morada, condições de saúde e até mesmo emocionais. Reconhecer uma pessoa ao telefone, antes que ela se apresente, e atribuir o nome a uma personalidade, ao ouvir uma entrevista no rádio, são situações corriqueiras. Ao realizar uma dessas tarefas, o indivíduo está realizando o

¹ Para que fique clara a aplicação dos termos “fala” e “voz” durante o texto, o primeiro se referirá à articulação da voz com o objetivo de comunicação verbal e o segundo se referirá ao processo fisiológico de produção sonora no trato vocal, usualmente obtida na laringe através da vibração das pregas vocais (CRYSTAL, 2008, p. 514).

*reconhecimento do falante*². É a noção mais geral e abrangente, que inclui o processo natural que as pessoas realizam no dia a dia ao atribuir um rosto a uma fala ou mesmo ao som da voz, bem como qualquer processo de tomada de decisão baseado em características do sinal³ acústico transmitido.

A respeito da noção de reconhecimento do falante, Nolan (1983) entende que a definição adequada para o conceito leigo de reconhecimento de falante seja "any decision-making process that uses some features of the speech signal to determine if a particular person is the speaker of a given utterance"⁴ (ATAL, 1976, citado por NOLAN, 1983, p. 5). Com o objetivo de delimitar a área de atuação do técnico na tarefa de reconhecimento, Nolan (1983, p. 07; 1997, p. 745) e Rose (2002, p. 92) propõem uma divisão da tarefa em duas classes: reconhecimento de falantes estabelecido através da aplicação de técnicas analíticas para obtenção dos resultados e o reconhecimento de falantes estabelecido naturalmente por oitiva, aos quais se refere como reconhecimento técnico de falantes e reconhecimento leigo de falantes⁵, respectivamente.

A tarefa de reconhecimento técnico de falantes pode ser dividida em duas categorias, de acordo com o objetivo do procedimento: verificação ou identificação. O reconhecimento por verificação de falante busca confirmar a veracidade de uma alegada identidade, o que ocorre normalmente quando um indivíduo deseja acessar um ambiente seguro, como acesso à conta bancária via canal telefônico, por exemplo. É o tipo de tarefa operada por sistemas de verificação automática e não será foco da pesquisa aqui apresentada. O reconhecimento por identificação de falante é a tarefa de se tomar um registro de fala cuja autoria é desconhecida e comparar com um ou mais registros de

² Diferente de **reconhecimento de fala**, processo onde o que se quer definir é o conteúdo do que está sendo dito. É o procedimento utilizado, por exemplo, em programas computacionais de edição de texto a partir da fala, onde se dispensaria a necessidade de teclar o texto. Esse conceito não será abordado no presente trabalho.

³ "Sinal" deve ser interpretado conforme os conceitos da Teoria da Informação de Shannon, ou Teoria Matemática da Comunicação (SHANNON, 1948), que trata dos sistemas de comunicação de forma geral. Assim, um sinal pode ser uma representação de um impulso elétrico, da pressão de água ou do som, conforme o fenômeno em questão. No presente estudo, está relacionado com as noções do processamento digital de sinais acústicos. Quando houver a menção, neste texto, a "sinal de voz", "sinal de fala" ou "sinal acústico", a intenção é expressar que o sinal é uma representação, analógica ou digital, de uma fonte sonora originada no trato vocal.

⁴ Qualquer processo de tomada de decisão que utilize características do sinal de fala para determinar se uma pessoa em particular é o falante de uma dada amostra.

⁵ Tradução do Autor para "technical speaker recognition" e "naive speaker recognition".

origem conhecida para determinar se o desconhecido é ou não da mesma origem de algum dos conhecidos (NOLAN, 1997, p. 745; ROSE, 2002, p. 82; HOLLIEN. 2002, p. 5).

Com relação aos elementos pesquisados em registros da expressão oral de um indivíduo específico, a identificação de falantes busca (1) determinar parâmetros de voz e de fala que sejam indicadores de características anátomo-biológicas e idioletais, que dizem respeito a aspectos próprios do indivíduo, particularizando-o, bem como (2) determinar características dialetais e socioculturais, que dizem respeito a aspectos contextuais e, mesmo não se prestando a definir características particulares do indivíduo, definam seu perfil de inserção em uma comunidade de fala⁶, auxiliando em procedimentos de triagem ou corroborando dados individualizantes.

Apesar de ter alguns elementos análogos a outras áreas da identificação humana, a fala está longe de ser um fenômeno constante em um indivíduo, como o são o DNA e as impressões digitais. Ao comparar-se a fala e a impressão digital, é possível afirmar que uma impressão digital é única, ou seja, varia consideravelmente quando tomado o âmbito interindividual, pois dois indivíduos não apresentarão a mesma digital, e não varia quando considerado o âmbito intraindividual, já que a conformação das papilas de um dedo do indivíduo será sempre a mesma, ao passo que a fala apresentará maior variabilidade de um mesmo parâmetro registrado repetidas vezes. Tal noção, aliada à relativa facilidade, quando comparada a outras técnicas, de obtenção de amostras, tornou a identificação através de impressões digitais a forma mais difundida e tradicional na área da identificação civil e criminal.

A analogia entre os processos de identificação através da fala e das impressões digitais foi apresentada por Kersta (1962), que defendeu o princípio de que a imagem espectrográfica obtida pelo sonógrafo era equivalente à impressão digital, no que diz respeito à sua condição individualizadora, batizando o achado de *voiceprint*, ou impressão vocal⁷, numa tentativa de remeter à expressão *fingerprint*, ou impressão digital.

⁶ Segundo Hymes (1974), a comunidade de fala é antes uma entidade social do que linguística, corroborando o que Labov (1972) já havia afirmado: a comunidade de fala é definida pelo compartilhamento de um conjunto de normas de comportamento e não pela concordância no uso de elementos de linguagem.

⁷ Tradução do Autor.

A técnica de comparação através da impressão vocal proposta por Kersta (1962) sofreu severas críticas, pois essa analogia ao termo *fingerprint* estabelece um paralelo entre os dois vestígios como se a imagem espectrográfica apresentasse uma consistência como vestígio tão evidente e intuitiva quanto a impressão digital, o que oferece uma noção equivocada a respeito dos princípios da identificação de falantes (HOLLIEN, 1990; NOLAN, 1997; KÜNZEL, 1994; BONASTRE ET AL., 2003). Tal equívoco de perspectiva consiste no fato de que a variabilidade da fala respeita a dicotomia "orgânico/adquirido", ou seja, as diferenças entre as vozes podem ser separadas em dois grupos fundamentais: diferenças orgânicas e adquiridas, ou seja, as que podem ser determinadas pelo aparato orgânico e pela aprendizagem; por outro lado, a impressão digital não sofre, ou sofre muito pouco, a ação do "adquirido"⁸ (GLENN E KLEINER, 1968; WOLF, 1972; NOLAN, 1983) e, dessa maneira, a variabilidade individual inerente ao vestígio linguístico deve ser considerada de maneira diversa em relação ao vestígio das papilas digitais.

As críticas ao termo *voiceprint* (HOLLIEN, 1990⁹; NOLAN, 1997; KUNZEL, 1994; BONASTRE ET AL., 2003) também serviram como um alerta à tendência de que os operadores do Direito¹⁰ se deixassem levar pelo óbvio apelo da evolução tecnológica, confiando mais no resultado visual do espectrograma do que na formação e conhecimento do profissional que estava interpretando os dados oferecidos pelo recurso tecnológico. Como um dos críticos mais enfáticos à técnica de *voiceprint identification*, Hollien et al. (2014, p. 3-5) e Hollien (2016, p. 335) chegam a afirmar que a notoriedade que a técnica rapidamente obteve foi um entrave ao desenvolvimento da pesquisa em identificação de falantes, pois causou uma retração das agências oficiais de fomento em função do "sucesso" daquela técnica. Apesar de haver o entendimento da multidisciplinaridade requerida para a atuação na área identificação de falantes, a ênfase na Linguística como formação de base é um ponto de acordo (HOLLIEN, 1990, pg. 335; ROSE, 2002, pg. 332). Antes de abordar-se o surgimento da atuação de linguistas no

⁸ Em inglês, *learned*.

⁹ Nessa obra há um capítulo inteiro dedicado à discussão do "*Voiceprint*" *Problem*, título do capítulo, o que oferece uma noção sobre a polêmica envolvendo o tema, que não é relevante para a presente pesquisa.

¹⁰ Bacharéis em Direito que atuam no âmbito jurídico e tomam as decisões nos procedimentos legais. São os delegados, advogados, juízes, promotores de justiça e defensores públicos.

ambiente forense, porém é necessário um comentário a respeito da Linguística e suas áreas.

A Fonética como uma das áreas de estudo da Linguística é uma noção óbvia. Os termos utilizados para definir a atuação forense, porém, não respeitam exatamente esse critério de pertinência de um conjunto em relação ao outro. Os termos “Fonética Forense” e “Linguística Forense” delimitam campos práticos de atuação que surgiram de forma independente, com o primeiro estando relacionado aos estudos fonéticos e o segundo mais relacionados aos estudos do discurso, falado ou escrito. Segundo esse raciocínio, a tese que aqui se apresenta, dedicada à identificação de falantes, estaria relacionada à área da Fonética Forense e não da Linguística Forense.

McMenamin (2002, p. 70-71), ao descrever com detalhes os campos da Linguística Forense, cita dois campos relativos ao estudo dos sons da linguagem para fins forenses, a Fonética Auditiva e a Fonética Acústica, o que seria equivalente à delimitação da Fonética Forense recém citada, numa tentativa de relacionar os campos de atuação tal como se organiza o campo da própria Linguística. Assim, apesar de historicamente a Fonética apresentar-se como precursora da atividade de pesquisadores no âmbito forense, atualmente está se consolidando o entendimento de que a Linguística, como área mais abrangente do conhecimento, ocupa lugar de destaque nesse universo.

Na área da Fonética Forense, especificamente, a incursão de linguistas é uma realidade mais tradicional nos países europeus, EUA e Austrália, onde, desde a década de 60, conforme Hollien (1990, 2002, 2014 e 2016), se difundiram as solicitações por parte do poder judiciário para a resolução de questões como identificação de falantes em registros de áudio, definição de características dialetais com intenção de elaboração de perfil, determinação do conteúdo efetivo de uma fala cuja inteligibilidade apresentava-se prejudicada ou mesmo para a verificação de hipótese de edição do registro na mídia. Essas ações no âmbito forense podem ser consideradas o que originou o campo da Fonética Forense e a consolidação dessa área foi, em boa parte, influenciada pelo estabelecimento, a partir de 1991, da Associação Internacional de Fonética Forense (FRENCH, 1994, pp. 170, 182-4; NOLAN, 1997, p. 746).

No Brasil, a Fonética Forense percorre um caminho mais tardio e com características peculiares, uma vez que, por força do nosso sistema legal, o Estado tem, a princípio, a obrigação de oferecer, na esfera criminal, o trabalho pericial à sociedade. O marco histórico pode ser considerado o Caso Magri, fato ocorrido em 1991, em que o ministro do Trabalho e da Previdência Social, Antônio Rogério Magri, foi gravado admitindo ter aceitado propina para facilitar liberação de recursos do FGTS para uma empresa. Os peritos do Instituto Nacional de Criminalística, incumbidos de determinar de quem eram as vozes presentes na gravação e sem formação para cumprir tal tarefa, recorreram à Universidade de Campinas para realizar os exames (CALDAS NETTO, 2003). Figueiredo (1994), após participação no caso histórico referido, testou a eficiência de diversos parâmetros acústicos para a identificação de falantes. Além da escassez de trabalhos nacionais na época, não se percebia um trabalho em conjunto entre os órgãos periciais e as universidades¹¹.

Assim, os primeiros trabalhos surgiram dentro dos órgãos oficiais, não necessariamente realizados por linguistas, como é o caso de Braid (1999), Perito Oficial de formação em Engenharia Eletrônica, que aborda os princípios articulatórios, acústicos e fisiológicos da fala, bem como oferece uma noção introdutória ao procedimento de identificação de falantes, além de abordar alguns aspectos legais da interceptação telefônica.

Como Perito Criminal do estado do Rio Grande do Sul desde 2003, envolvido com as perícias na área da Fonética Forense, particularmente com a identificação através da fala e da voz, o autor percebeu a relevância da questão da variabilidade individual inerente à natureza dos dados de fala que serviam como elementos de cotejo. Os pesquisadores na área da identificação analisam e escolhem cuidadosamente os parâmetros indicadores de características físicas do aparato vocal, que tenderiam a variar

¹¹ O cenário nacional foi alterado com a multiplicação do conhecimento científico, já estabelecido na área acadêmica, através da implementação, a partir de 2007, da Capacitação Nacional para Peritos Criminais em Fonética Forense, promovida pela Secretaria Nacional de Segurança Pública e operacionalizada pelo Instituto Nacional de Criminalística. A iniciativa, além de capacitar dezenas de peritos federais e estaduais dos respectivos órgãos periciais, e instrumentalizar os setores desses servidores, uma vez que, para cada participante, uma estação de trabalho era cedida ao setor correspondente, aproximou os órgãos periciais das universidades, pois o corpo docente da capacitação conta com pesquisadores de universidades brasileiras, dentre eles os Doutores Alexsandro Rodrigues Meireles, Cláudia Regina Brescancini (PUCRS), Dermeval da Hora (UFPB), Irene Queiroz Marchesan (CEFAC), Luisa Barzaghi Ficker (PUC-SP), Plínio Barbosa (UNICAMP) e Sandra Madureira (PUC-SP).

menos no indivíduo em contraposição a parâmetros linguísticos adquiridos, os quais tenderiam a variar menos no grupo. Em realidade, segundo Nolan (1997), a plasticidade apresentada pelo trato vocal /h/ confere um amplo escopo de variação¹² controlada. Isso ocorre pois, para se atingir um som específico, a configuração do trato vocal é determinada, por um lado, pela fonologia e, por outro, pelas características orgânicas individuais. Uma das premissas do processo de identificação pela voz e fala é que, durante a realização de um determinado som, o formato do trato vocal de um indivíduo apresenta uma invariabilidade que o particulariza.

Relacionados diretamente com o formato do trato vocal, os dois primeiros formantes das vogais orais, parâmetros consagrados na identificação de falantes, são representativos das dimensões vertical (primeiro formante) e horizontal (segundo formante) da cavidade oral, dimensões que variam principalmente como consequência da posição da língua nessa cavidade. Esse articulador apresenta grande mobilidade, característica que ocasiona variação considerável, o que exige uma série de cuidados durante o processo de identificação, como, por exemplo, o controle dos contextos linguísticos em que ocorre a vogal analisada.

A pesquisa em tela¹³ visa a colaborar com o exame de identificação de falantes realizado no âmbito forense, conforme a definição feita por Nolan (1983), procedimento denominado, no Brasil, *Comparação de Locutores* (doravante, CL)¹⁴. Tal exame pericial compara, quando a qualidade do material de áudio permite, um registro de fala objeto de delito com um registro de fala de referência, cuja autoria é conhecida, preferencialmente obtido por perito. O objetivo desse exame é determinar o grau de semelhança entre os dois registros.

A proposta desta pesquisa visa a contribuir para a área forense de identificação biométrica, particularmente no que se refere à variabilidade individual. Todas as variáveis

¹² O termo “variação” será utilizado neste texto com o sentido amplo, como sinônimo de variabilidade, sem conotação específica própria da área da Sociolinguística.

¹³ O projeto desta tese foi aprovado em 19/07/2014 pela Comissão Nacional de Ética e Pesquisa sob o Certificado de Apresentação para Apreciação Ética – CAAE – de número 32588114.3.0000.5336.

¹⁴ Denominação atualmente adotada pela comunidade pericial brasileira, em substituição aos termos “verificação de falante” e “verificação de locutor”, anteriormente empregados. Foi proposta pelo Perito Criminal Federal Charles Valente, em apresentação oral, por ocasião do IX Seminário Nacional de Fonética Forense, em 22 de agosto de 2012, em Brasília/DF.

de pesquisa e o aporte teórico utilizado vão, de alguma maneira, ser confrontados com esse tema: (a) através de uma coleta de dados configurada de forma a permitir análise e discussão a respeito da variabilidade intra e interindividual, uma vez que serão consideradas elocuições obtidas através de frases-veículo, repetidamente gravadas – são vinte repetições programadas – para cada indivíduo participante do estudo, o que permitirá estabelecer-se a variabilidade intraindividual. A amostra em análise refere-se a um conjunto de indivíduos (vinte sujeitos), que, associado ao número de repetições individuais, resultará num *corpus* de tamanho estimado de 420 gravações, ou cerca de 25.000 ocorrências sonoras, e permitirá estabelecer a relação entre variabilidade intraindividual e interindividual, seguindo a ideia de Heuvel (1996) para a definição de especificidade de falante; (b) através avaliação da variabilidade das medidas de oito sons vocálicos em posição tônica, [i e ε ə a ɔ o u], sons tradicionalmente mensurados para a CL, bem como da variabilidade das medidas de sons consonantais nasais, [m n], a partir da noção de que a cavidade nasal não apresenta a plasticidade do trato vocal e se comporta acusticamente com menor variabilidade, conforme Glenn e Kleiner (1968), Wolf (1972), Höfker (1977) e Figueiredo (1994), os quais apontam o melhor resultado de consoantes nasais, frente inclusive às vogais, quando considerados índices de variabilidade individual; (c) através das medidas vocálicas, que servirão para estabelecer parâmetros geométricos relativos, especificamente à área do espaço vocálico (doravante VSA) e distâncias euclidianas entre pontos vocálicos específicos desse espaço, com o intuito de procurar proporções geométricas que possam ser utilizadas na CL e (d) através da discussão da capacidade de modelamento da própria teoria fonológica baseada no gesto articulatório, a Fonologia Articulatória¹⁵ (Browman e Goldstein, 1986), que traz no seu cerne a definição de unidade de análise que contempla a manifestação física do gesto e sua representação abstrata. Pretende-se, com base na ideia de que o modelo realmente oferece uma perspectiva mais integrada entre os campos da Fonética e da Fonologia, mais especificamente entre produção e percepção, investigar a possibilidade de a Fonologia Articulatória contribuir à aplicação forense.

¹⁵ Originalmente proposta pelos autores sob essa denominação, “Fonologia Articulatória”, mas consagrada no Brasil como Fonologia Gestual, com a intenção de evitar a associação, inadequada, à Fonética Articulatória (SILVA, 2008, p.4; ALBANO, 2012, p. 14).

Nolan (1995), ao sugerir tendências de avanço na área da identificação forense de falante, afirma que é necessário encontrar novas formas de lidar com a variabilidade individual. Para isso, o autor entende que é necessária uma abordagem que discuta como o material fonético é estruturado pela fonologia e refere a Fonologia Articulatória como aparato teórico que, ao definir o modelamento do trato vocal, permite que a síntese de fala ofereça algum esclarecimento sobre a questão.

O trabalho aqui proposto pretende analisar os fundamentos do aparato teórico da Fonologia Articulatória no que se refere à sua capacidade de modelamento do trato vocal, particularmente em relação aos sons vocálicos, e discutir a abordagem desses sons no modelo. Parte-se da noção de que a capacidade de modelamento pode oferecer esclarecimentos sobre a configuração das estruturas componentes do trato relevantes ao exame de CL, não acessíveis através da análise acústica tradicional das evidências linguísticas.

São objetivos específicos deste trabalho:

- 1) Estabelecer valores mínimos de variabilidade, através dos dados de coleta monitorada, tanto para os sons medidos (consoantes nasais e vogais orais) como para os parâmetros considerados (f_0 , os três primeiros formantes¹⁶). Esses valores mínimos servirão de referência para o procedimento de CL.
- 2) Estabelecer uma escala de variância entre os sons vocálicos orais avaliados, com o intuito de determinar quais os mais consistentes para o procedimento de CL.
- 3) Calcular o VSA e distâncias euclidianas com base nas medidas vocálicas e verificar se a variabilidade dessas medidas são adequadas para a CL, ao comparadas com as demais variabilidades.

¹⁶ Conforme será exposto nos resultados, no Capítulo 5, a variável largura de banda, originalmente pensada para compor o estudo, foi abandonada após os primeiros ensaios.

- 4) Estabelecer uma escala de variância entre os sons consonantais nasais avaliados, com o intuito de determinar quais os mais consistentes para o procedimento de CL.
- 5) Comparar a variabilidade entre sons e classes diferentes, através do cotejamento entre os resultados para sons consonantais nasais com aqueles para os sons vocálicos orais.
- 6) Determinar, analogamente ao que se propõe para os sons consonantais e vocálicos, uma escala de consistência (conforme sua variabilidade individual) para os parâmetros acústicos utilizados na CL.

São hipóteses deste trabalho, fundamentadas no que será exposto nos Capítulos 2 e 3:

1. É possível determinar valores referenciais de variabilidade, com base em coleta de fala monitorada, que sirvam de referência para o exame de CL.
2. Há diferença na variabilidade dos parâmetros acústicos de curto termo de diferentes sons do PB, especificamente os vocálicos orais ([i e ε ə a ɔ o u]), que permita ordená-los. Dentre os vocálicos orais, os abertos apresentam menor variabilidade do que os fechados.
3. A variabilidade do VSA e das distâncias euclidianas vocálicas é análoga à dos sons vocálicos e consonantais, o que permite seu uso na prática forense.
4. Há diferença na variabilidade dos parâmetros acústicos de curto termo de diferentes sons do PB, especificamente os consonantais nasais ([m n]), que permite ordená-los.
5. Os sons consonantais nasais apresentam variabilidade menor do que os vocálicos orais, prestando-se com mais efetividade como parâmetros de comparação. Dentre as consoantes nasais, a bilabial é a que apresenta menor variabilidade.
6. Dentre os parâmetros acústicos, f_0 e F_3 são os que apresentam menor variabilidade.

Este trabalho está composto por cinco capítulos após este introdutório. No Capítulo 2, serão apresentadas as bases da Fonética Acústica e Articulatória que servirão de instrumental para a operacionalização da pesquisa, a partir da descrição geral dos campos, com foco nas vogais orais e nas consoantes nasais, e na descrição da área forense, onde será abordado o uso das vogais orais e das nasais na CL, bem como a questão da variabilidade individual. Serão apresentadas, no Capítulo 3, as bases da Fonologia Articulatória, que servirão de aporte teórico para interpretar os resultados obtidos. Além dessa investigação, pretende-se discutir a dissociação Fonética/Fonologia, uma vez que o modelo teórico pleiteia uma relação entre essas áreas diferente do que as teorias fonológicas tradicionais estabelecem. No Capítulo 4 descreve-se o método que conduziu esta pesquisa e, no Capítulo 5, apresentam-se resultados obtidos e discute-se sua aplicabilidade à prática forense, bem como busca-se confrontá-los com as bases implementacionais da Fonologia Articulatória. O Capítulo 6 encerra este relatório com a apresentação de algumas conclusões e sugestões caminhos futuros de pesquisa.

2 FONÉTICA ACÚSTICA E ARTICULATÓRIA

A fim de basear o trabalho em tela na literatura sobre as condições e limites teóricos da Fonética Acústica e Articulatória, os trabalhos de Fant (1970), Kent e Read (1992) e Stevens (1998), dedicados primordialmente à Fonética Acústica, e os de Peter Ladefoged (1975, 1996) e Ladefoged e Maddieson (2008), voltados para a Fonética Articulatória, guiarão a interpretação acústica e articulatória das mensurações de curto termo dos sons avaliados na pesquisa aqui relatada. Diversos outros trabalhos serão referenciados nesta revisão, porém as linhas sobre as quais todos seguem se entremeiam com as noções propostas pelos autores referidos acima.

Inicialmente será apresentada uma visão geral da Fonética, seguindo com fundamentos da Fonética Acústica e Articulatória. Referências a trabalhos sobre aspectos acústicos e articulatórios das vogais orais e depois das consoantes nasais serão trazidas, sendo que, para cada um desses dois grupos, serão analisadas suas produções de forma geral nas línguas do mundo e de forma específica no português brasileiro (doravante PB). Após abordar a Fonética Acústica e Articulatória dos sons vocálicos e consonantais nasais, serão percorridas pesquisas dedicadas ao campo da Fonética Forense, tomando como referência os trabalhos de Nolan (1983, 1997), Hollien (2002) e Rose (2002), cujos conceitos nortearão parte das análises dos resultados e oferecerão base teórica para considerações acerca da utilização de ferramentas estatísticas voltadas para a CL.

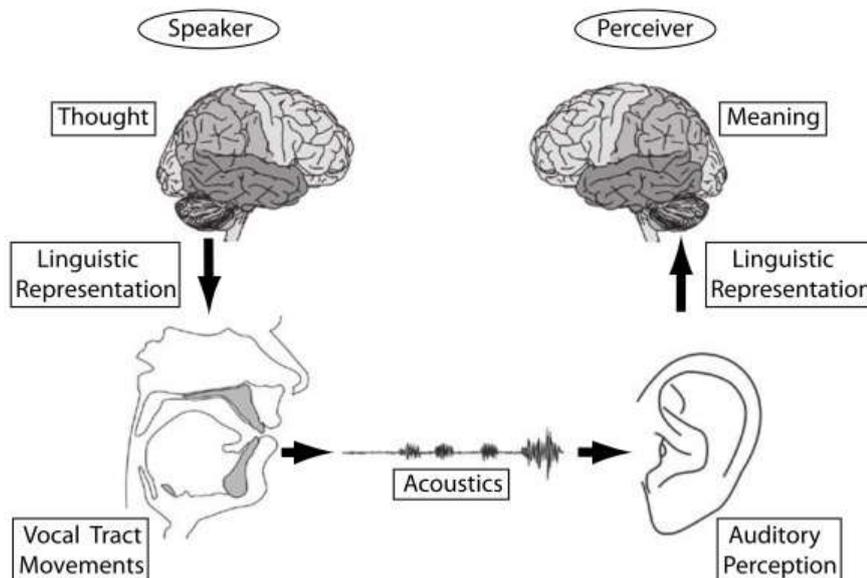
2.1 CONSIDERAÇÕES GERAIS: PRODUÇÃO E PERCEPÇÃO

O senso comum no domínio da Linguística tradicionalmente divide a Fonética em três áreas principais: Fonética Articulatória, Fonética Acústica e Fonética Auditiva (ou Perceptual). De forma geral, os autores concordam na existência desses três campos e alguns sugerem ramos que vão além dessa tríade, como Reetz e Jungman (2009, pg. 1), que referem a Transcrição Fonética como uma área; Cristófaró Silva (2003, pg.23), que considera a Fonética Instrumental como uma das principais áreas de interesse da Fonética e Barbosa e Madureira (2015, pg. 28), que incluem a Fonética Cognitiva como foro de interesse dos estudos fonéticos. A tese aqui proposta toma como ponto de partida a divisão tradicional em três áreas e, considerando a natureza dos questionamentos

levantados e o tratamento dos dados para lidar com tais questões, os achados de estudos relacionados à Fonética Auditiva não farão parte da fundamentação deste trabalho de pesquisa. Assim, a revisão ficará centrada nas áreas de Fonética Articulatória e Fonética Acústica.

No esquema simplificado proposto por Gick et al. (2013, p. 4) na Figura 2.1 a seguir, os autores idealizam o circuito da conversa entre dois indivíduos com o objetivo de ilustrar a ideia de um sistema perspectiva unidirecional de fluxo da informação.

Figura 2.1 – Circuito de conversa entre dois indivíduos numa perspectiva unidirecional.



Fonte: Gick et al. (2013, p. 4).

Nota: *Speaker*: falante; *Thought*: pensamento; *Linguistic Representation*: representação linguística; *Vocal Tract Movements*: movimentos do trato vocal; *Acoustics*: acústica; *Auditory Perception*: percepção auditiva; *Meaning*: significado; *Perceiver*: ouvinte.

Na porção inferior do circuito que aparece na Figura 2.1, veem-se os componentes que se referem às três áreas da Fonética. Assim, o rótulo, à esquerda, referente aos movimentos do trato vocal, indica a competência dos estudos em Fonética Articulatória, o rótulo ao centro, referente ao sinal acústico, indica a competência da Fonética Acústica e o rótulo à direita, referente à percepção auditiva, indica a competência da Fonética Auditiva. A última, por lidar com dados relativos a estudos perceptuais, não servirá como fundamentação para o trabalho a ser aqui proposto, tendo em vista que a CL lida com as informações de registros de áudio e, assim, valoriza os achados articulatórios e acústicos para estabelecer as conclusões acerca da identificação de indivíduos. Os estudos

perceptuais não crescem às análises realizadas na CL, uma vez que não há interesse em determinar como o indivíduo percebe os fenômenos da língua, mas sim quais os achados acústicos e articulatórios que permitem sua particularização.

Ressalta-se que a ideia expressa na Figura 2.1 tem caráter simplificador, apenas para ilustrar a maneira como as três grandes áreas de estudos linguísticos são tradicionalmente delimitadas, sendo que essa divisão pode ser questionada conforme a perspectiva teórica adotada, como será abordado no Capítulo 3, na descrição dos princípios da Fonologia Articulatória.

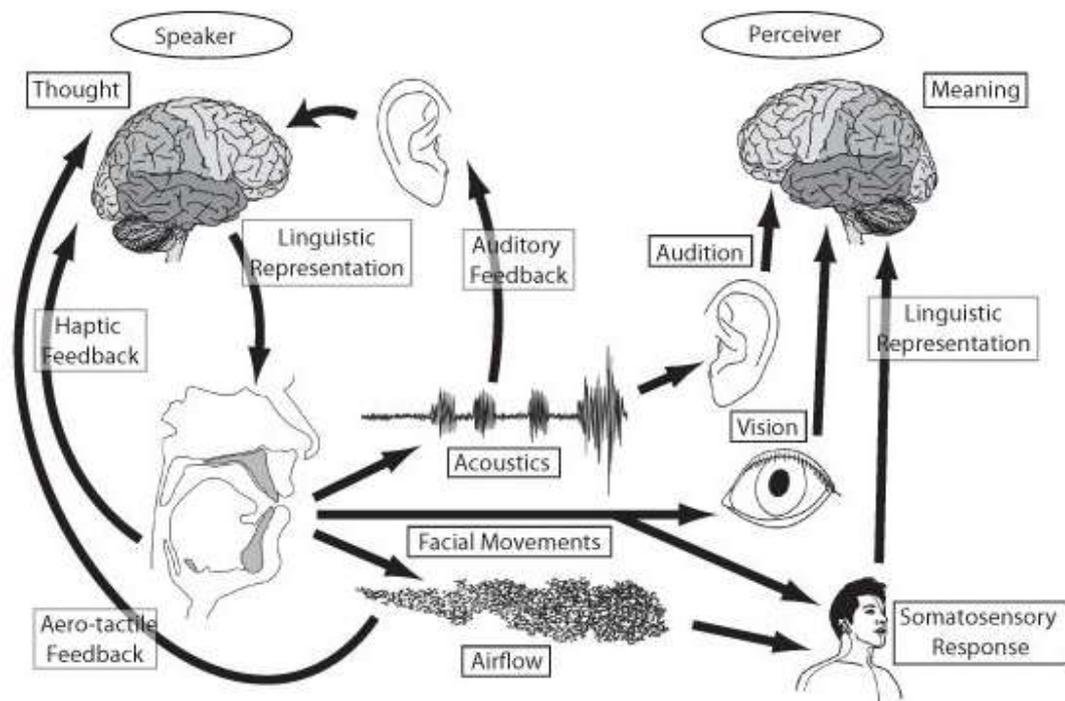
O esquema da Figura 2.1 pressupõe uma pré-alimentação, ou seja, um confronto do sinal a ser emitido com uma imagem acústica resgatada na memória, conforme Albano (2001, p. 100) define ao expressar a tradução para o termo *feed forward*. O ciclo de fala envolve muito mais aspectos do que esses expressos de maneira simplificada anteriormente, como bem referem Gick et al. (2013):

Although the simplicity of a feed-forward model is appealing, we know that producing speech is not strictly linear and unidirectional. Rather, when we speak, we are also constantly monitoring and adjusting what we're doing as we move along the chain. We do this by using our senses to perceive what we are doing. This is called *feedback*. In a feedback system, control is based on observed results, rather than on a predetermined plan. The relationship between feedforward and feedback control in speech is complex. Also, speech perception feedback is *multimodal*. That is, we use not just our sense of hearing when we perceive and produce speech, but all of our sense modalities – even some you may not have heard of before. Thus, while the speech chain as a whole is generally linear, each link in the chain – and each step in the process of speech communication – is a loop [...]. We can think of each link of the chain as a *feedback loop*.¹⁷ (GICK ET AL., 2013, p. 4)

¹⁷ Embora a simplicidade de um modelo pré-alimentado seja tentadora, nós sabemos que a produção da fala não é linear e unidirecional. Ao contrário, quando nós falamos, nós estamos constantemente monitorando e ajustando aquilo que estamos fazendo à medida percorremos a cadeia. Nós fazemos isso utilizando nossos sentidos para perceber o que estamos fazendo. Isso é chamado *retroalimentação*. Em um sistema retroalimentado, o controle é feito com base nos resultados observados ao invés de ser feito com base em plano predeterminado. A relação entre controle pós alimentado e retroalimentado é complexa. Além disso, a retroalimentação na percepção da fala é *multimodal*. Ou seja, nós usamos não apenas nosso sentido da audição quando percebemos e produzimos a fala, mas todas nossas modalidades sensitivas – mesmo alguma de que você não tenha ouvido falar antes. Assim, enquanto a cadeia de fala como um todo

A Figura 2.2 a seguir, que expande a Figura 2.1 a partir do detalhamento dos processos mentais, fisiológicos e físicos envolvidos na comunicação verbal oral situa a *acústica* em relação às demais etapas do circuito da fala sob uma perspectiva multimodal de *feedback*. Para a CL, a *acústica* desempenha papel fundamental pois ocupa a posição, no esquema, em que o registro de que os peritos dispõem é efetuado. Os vestígios examinados na CL sempre são registros de áudio, na maioria dos casos interceptações telefônicas. Outros registros, como palatografia, pressão intraoral, fluxo aéreo, eletromiografia, para citar alguns exemplos, são fundamentais para pesquisas em laboratório, não apenas no campo da Linguística, com o objetivo de analisar outras posições do esquema, que inclui no circuito mecanismos de *feedback* e permite um

Figura 2.2 – Circuito da fala sob uma perspectiva multimodal de *feedback*.



Fonte: Gick et al. (2013, p. 5).

Nota: *Speaker*: falante; *Thought*: pensamento; *Haptic Feedback*: retroalimentação propioceptiva; *Linguistic Representation*: representação linguística; *Aero-tactile Feedback*: retroalimentação aero-tátil; *Auditory Feedback*: retroalimentação auditiva; *Acoustics*: acústica; *Facial Movements*: movimentos faciais; *Airflow*: fluxo aéreo; *Somatosensory Response*: resposta somatossensorial; *Vision*: visão; *Audition*: audição; *Meaning*: significado; *Perceiver*: ouvinte.

é geralmente linear, cada elo na cadeia – e cada passo no processo de comunicação falada – é um laço. Nós podemos pensar em cada elo da cadeia como um *laço retroalimentado*.

vislumbre da multimodalidade presente entre o pensamento do falante e a compreensão do receptor (*perceiver*, no esquema da Figura 2.1). Como será visto adiante, no Capítulo 3, a Fonologia Articulatória é uma área de pesquisa que busca contemplar a multimodalidade ilustrada no esquema ao utilizar-se de diversas formas de obtenção de informação através de dispositivos de monitoramento de outros níveis que não apenas o acústico para fundamentar os seus argumentos teóricos. A partir da interpretação conjugada das diversas dimensões de dados, a Fonologia Articulatória fundamenta seus conceitos.

Tendo em vista que a proposta desta tese lida com a relação entre a realização articulatória e os respectivos achados acústicos, particularmente no que tange às vogais orais e às consoantes nasais, será explorada a literatura relativa à Fonética Articulatória e Acústica dessas classes de produções sonoras, para, na sequência, apresentar como esses conceitos colaboram com a aplicação forense, especificamente com o exame pericial de CL.

2.2 FUNDAMENTOS TEÓRICOS

A comunicação humana através de um sistema linguístico verbal oral, a que comumente chamamos de fala, não surgiu na evolução da espécie de uma hora para outra. Foi gradualmente se especializando da mesma forma que as estruturas orgânicas que estão envolvidas na sua produção. Os processos evolutivos, porém, diferem em sua essência: enquanto as estruturas orgânicas adquiriram sua forma e função por adaptação evolutiva, a fala foi o resultado da “exaptação”¹⁸ evolutiva desses mesmos órgãos, conforme Gould e Vrba (1982). Exaptação, segundo definem, é o processo pelo qual passa uma estrutura ao atingir uma função que não é aquela original para a qual se adaptou (*stricto sensu*), motivado pela seleção natural.

Diante do exposto, não se pode afirmar que exista um órgão específico da fala, pois essa função é o resultado de um processo evolutivo de exaptação em que diversos órgãos, que têm originalmente uma função fisiológica específica e fundamental para o organismo humano, desempenham funções outras. Diafragma e pulmões, responsáveis

¹⁸ Termo proposto a fim de evitar o uso equivocado, inclusive em trabalhos científicos, da noção de adaptação.

pela pressão aérea para a fonação, têm como origem as funções de ventilação e respiração; pregas vocais e demais músculos da laringe, produtores da fonte sonora da fonação, são parte de um delicado sistema esfínteriano de proteção da traqueia contra partículas estranhas à respiração; língua, lábios (bem como outros músculos da face) e estruturas dentárias, que funcionam como articuladores para produção dos diferentes sons das línguas, executam a mastigação e deglutição. Até mesmo o véu palatino, a cavidade nasal e seios da face têm uma participação no momento da fala não prevista por suas funções primárias.

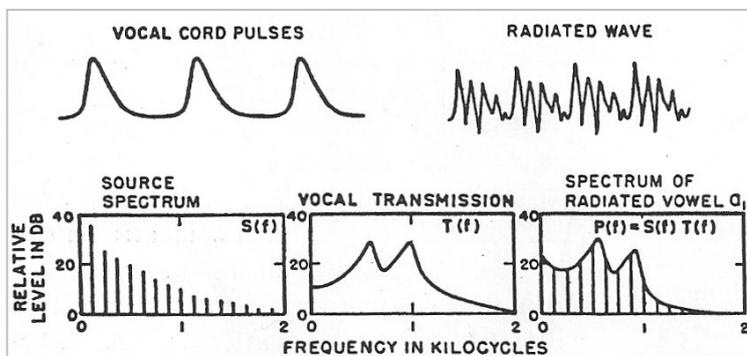
O som resultante do ato da fala será efeito de uma combinação do movimento sincronizado das estruturas exaptadas, que determinarão formatos específicos do trato vocal e combinações de acoplamentos entre as cavidades. Dessa maneira, os seres humanos produzem os diferentes sons utilizados na fala, os quais compõem o material de estudo da Fonética Acústica e, como quaisquer outras ondas sonoras, podem ser estudados através de suas características e propriedades físicas (frequência, período, intensidade, tempo, ressonância). A Fonética Acústica apenas, porém, não é capaz de explicar uma série de fenômenos resultantes de tal combinação de movimentos, uma vez que a sincronia das partes envolvidas na produção sonora não pode ser recuperada apenas com a base no dado acústico. É nesse ponto que a Fonologia Articulatória pode colaborar com a explicação dos fenômenos linguísticos de coarticulação, como será abordado no capítulo seguinte.

Conforme Ladefoged (1996, p. 117), o trato vocal é fechado na extremidade onde se encontram as pregas vocais, que compõem a fonte sonora, e aberto na extremidade dos lábios. Assim, fica estabelecida uma câmara de ressonância de formato complexo e variável, que pode ter a disposição de seus componentes alterada voluntariamente. Simplificadamente, do ponto de vista acústico, a estrutura pode ser considerada como um tubo com extremidade fechada na origem do som e aberta na saída. Os sons produzidos no seu interior entrarão em ressonância de acordo com o formato do tubo, de maneira que qualquer alteração no seu formato produzirá alteração correspondente no resultado.

A relação entre a fonte sonora e as características de ressonância do trato vocal é o princípio da Teoria Fonte-Filtro, proposta por Fant (1970, p. 15), que introduz esse trabalho afirmando: “The speech wave is the response of the vocal tract filter systems to one or more sound sources. This simple rule [...] implies that the speech wave may be uniquely specified in terms of source and filter characteristics¹⁹.”

Como se pode observar no esquema apresentado na Figura 2.3, a seguir, Fant (loc. cit.) considera a onda sonora emitida durante a fala como um produto, $P(f)$, do espectro da fonte, $S(f)$, pela função de transferência (ou característica do filtro), $T(f)$.

Figura 2.3– Ilustração da Teoria Fonte-Filtro



Fonte: Fant (1970, p. 19).

Nota: Simplificação da decomposição do espectro de um som vozeado com dois formantes

O autor expressa sua análise utilizando, além da terminologia acústica, terminologia de Engenharia Elétrica, ao representar as estruturas e características de fonte e filtro através de circuitos elétricos. Realiza cálculos baseados em medidas do trato vocal obtidas em dados de raios-X, utilizando como sujeito de pesquisa um falante do russo. Com isso, obtém modelamentos e traça considerações específicas para 1) vogais orais, 2) vogais nasais e consoantes nasais, 3) líquidas e 4) fricativas, africadas e plosivas. Os estudos expressos em 1) e 2) serão de particular interesse para este trabalho.

¹⁹ “A onda de fala é a resposta dos sistemas de filtro do trato vocal a uma ou mais fontes sonoras. Essa regra simples [...] implica que a onda de fala talvez seja unicamente especificada em termos de características de fonte e filtro”. Traduzido pelo autor.

Para o entendimento dos fenômenos envolvidos na produção da fala, necessário se faz o esclarecimento de como o sinal acústico resultante da fala deve ser analisado para a obtenção das informações a respeito tanto da fonte sonora como do filtro aplicado sobre tal fonte, o que será descrito a seguir.

O som da voz humana é uma onda quasi-periódica complexa e o processo de análise acústica utilizado em Fonética baseia-se na Análise de Fourier (LIEBERMAN E BLUMSTEIN, 1988, p. 24; REETZ E JONGMAN, 2009, p. 143; ELLIS, 2011, p. 764), que determina a representação da onda sonora complexa através da sua decomposição em diversas ondas senoidais (ou “ondas simples”). Para os sons vozeados, a frequência mais baixa nessa decomposição é chamada de frequência fundamental (representada por f_0). No caso dos sons periódicos, os valores múltiplos da frequência fundamental, que compõem a onda complexa original, são chamados de harmônicos.

A implementação matemática da Análise de Fourier dá-se através da transformada de Fourier, fórmula matemática que se presta à aplicação sobre sinais contínuos. Para a aplicação em sinais discretos, como aqueles digitalizados para a análise em computadores, é utilizada a transformada discreta de Fourier (DFT, para *Discrete Fourier Transform*), da qual as FFTs (para *Fast Fourier Transform*) são os algoritmos utilizados em boa parte dos programas computacionais (LIEBERMAN E BLUMSTEIN, 1988, p. 84).

A ferramenta de análise acústica mais difundida é o espectrograma, que oferece grande possibilidade de detalhamento das características temporais, de intensidade sonora e de frequência do sinal sonoro. É um recurso gráfico que apresenta tais grandezas simultaneamente, sendo o tempo e a frequência representados, respectivamente, nos eixos horizontal e vertical e a intensidade representada através de escala de tonalidade de cinza (tomando-se o branco como limite inferior e o preto como limite superior)²⁰.

Com o intuito de exemplificar o recurso e observar as características do gráfico, a Figura 2.4 a seguir apresenta a representação espectrográfica da frase-veículo “Ele dirá

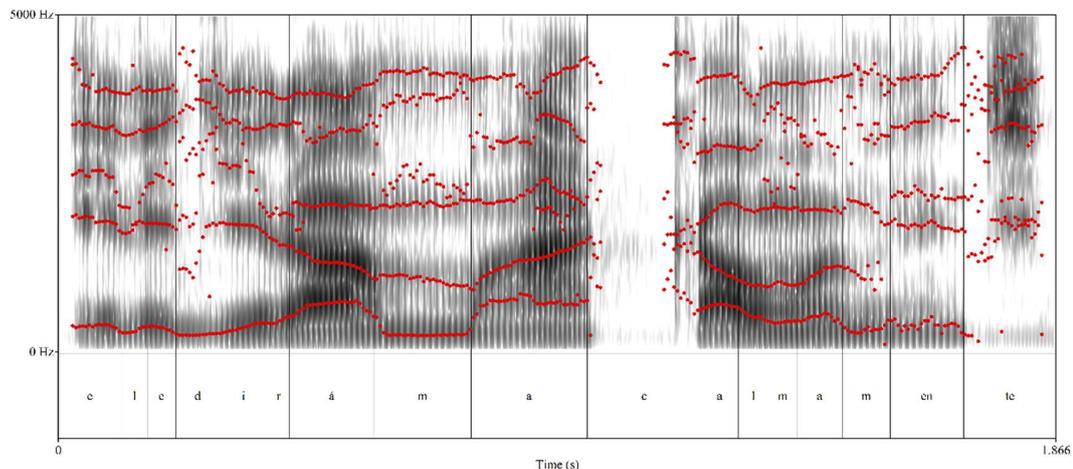
²⁰ Existem programas computacionais (Ocenaudio, Adobe Audition, entre outros) que permitem a representação de intensidade através de paletas de cores, que podem ser configuradas conforme a necessidade do usuário, porém a representação em tons de cinza é a mais difundida.

MA calmamente” (a porção em caixa alta refere-se à sílaba investigada), presente no *corpus* desta tese, cuja composição será abordada no capítulo de metodologia, onde se inclui a descrição do instrumento de coleta utilizado neste trabalho.

No espectrograma, as faixas de frequências em que percebemos a maior concentração de energia, e que ocorrem com maior evidência em sons vocálicos e soantes, são chamadas de formantes, que, na Figura 2.4, são as faixas horizontais em tons de cinza, sobre as quais estão expressos os históricos de formantes, em vermelho, calculados por codificação preditiva linear (explicada a seguir), e estão relacionados com as propriedades de ressonância do trato vocal.

Conforme Stevens e House (1976, p. 70), “the term formant has been used principally to indicate a concentration of spectral energy in a narrow frequency region of a speech signal²¹”.

Figura 2.4 – Espectrograma da frase “Ele dirá MA calmamente”.



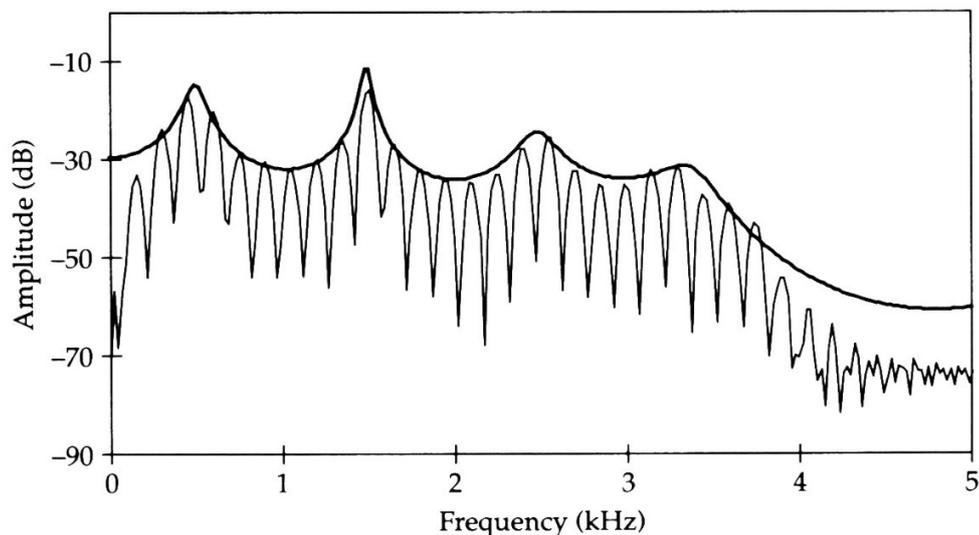
Fonte: O Autor (2017).

Nota: As marcas em vermelho referem-se ao histórico dos formantes, nesse caso configurado para apresentar até F5. A segmentação apresentada é meramente ilustrativa e não houve preocupação com a precisão na separação dos segmentos. A transcrição está expressa em caracteres do alfabeto gráfico e não do alfabeto fonético.

²¹ “O termo formante tem sido usado principalmente para indicar a concentração de energia espectral numa região estreita de frequência de um sinal de fala”. Tradução do autor.

A configuração do trato vocal determinará seu comportamento como filtro de ressonância sobre o espectro, que é o conjunto dos harmônicos que compõem a onda. O espectro está intimamente relacionado à fonte sonora e usualmente é calculado com base na FFT, já referida. Como resultado, o filtro reforçará aqueles harmônicos próximos das frequências dos formantes e atenuará os demais. A Figura 2.5, a seguir, apresenta a ação do filtro (linha superior, mais forte) sobre o espectro calculado por FFT (linha inferior, mais suave), na realização de um [ə].

Figura 2.5 – Relação fonte/filtro para [ə].

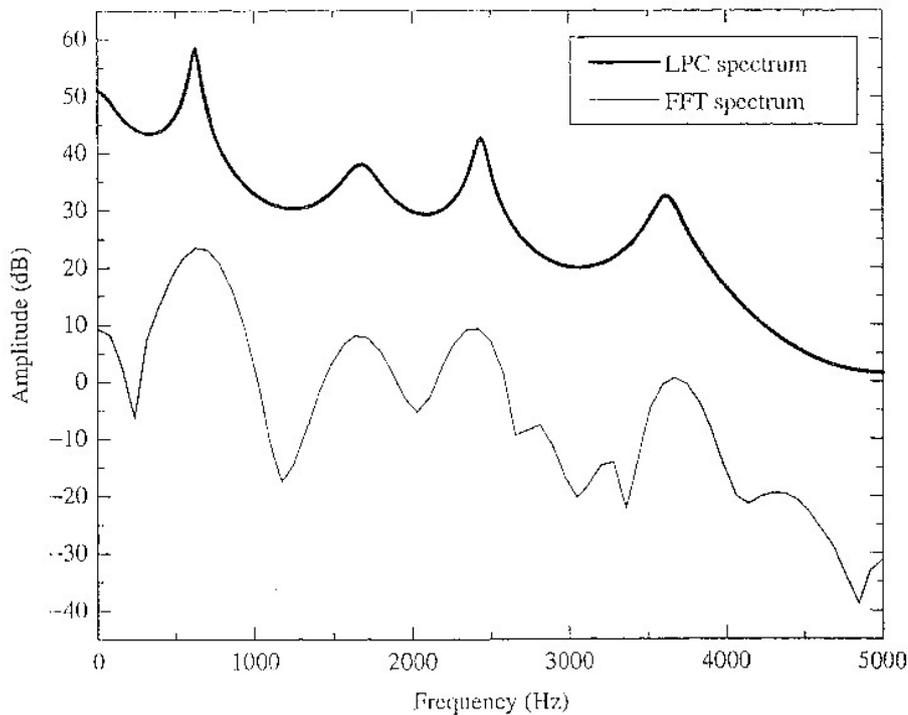


Fonte: Johnson (2003, p. 100).

A codificação preditiva linear (doravante, LPC, acrônimo da expressão, em língua inglesa, *Linear Predictive Coding*) é a ferramenta que realiza uma estimativa da função de filtro do trato vocal (LIEBERMAN e BLUMSTEIN, 1988, p. 86). O resultado do LPC é apresentado em um plano cartesiano onde o eixo vertical indica a intensidade e o eixo horizontal a frequência, de maneira que os picos da curva resultante indicam os formantes, ou seja, os valores em torno dos quais há maior efeito de ressonância. É a ferramenta de análise que será utilizada nesta pesquisa para obter-se os valores de formantes e está intimamente relacionada à conformação do trato vocal. A Figura 2.6 a

seguir ilustra um exemplo da relação entre LPC e FFT para a determinação de formantes. Como se percebe pela ilustração, o LPC tem melhor definição na região de energia formântica e oferece valores mais precisos de formantes, o que se pode constatar pelos picos formados para cada curva. A curva do LPC apresenta picos com delineamento mais agudo, facilitando a determinação do ponto de máxima intensidade.

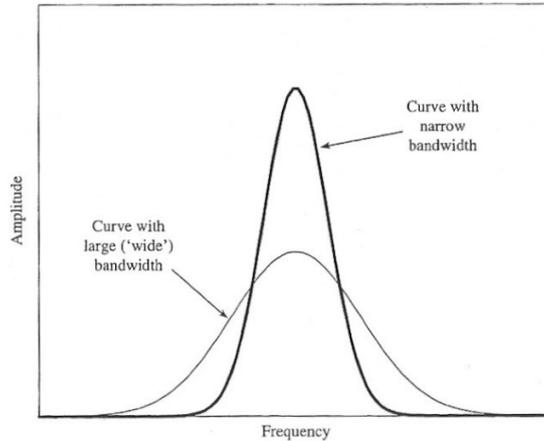
Figura 2.6 – Relação LPC/FFT para determinação de formantes



Fonte: Thomas (2011, p. 42).

Finalmente, para cada valor de formante obtido pelo cálculo do LPC, há um valor de largura de banda correspondente, que é, aproximadamente, a largura do formante 3 dB abaixo de seu pico (THOMAS. 2011, p. 47). A largura de banda indica a faixa, em torno da frequência do respectivo formante, que sofre o efeito de ressonância do filtro, no caso, do trato vocal. Quanto menor o valor da largura de banda, mais definido será o formante, pois uma maior quantidade de energia será concentrada em uma região mais estreita da dimensão da frequência, como se pode observar na Figura 2.7 a seguir, ao comparar-se a curva com a largura de banda mais estreita, aquela com o traço mais espesso, com a curva com a largura de banda mais larga, aquela com o traço mais tênue.

Figura 2.7 – Ilustração do conceito de largura de banda



Fonte: Thomas (2011, p. 47).

As análises acústicas que serão utilizadas neste trabalho se fundamentarão em medidas acústicas como frequência fundamental (f_0), valores de formantes (F_1 , F_2 e F_3) e largura de banda de cada formante (B_1 , B_2 e B_3). As frequências dos formantes são de grande relevância para a caracterização do falante, pois são determinadas pelas duas dimensões do trato vocal e pela maneira como os órgãos vocais se configuram para a produção de cada som (NOLAN, 2002, p. 78; NOLAN e GRIGORAS, 2005)). F_1 e F_2 permanecem mais preservados no sinal acústico e apresentam relações acústico-articulatórias diretas com as dimensões vertical (F_1) e horizontal do trato vocal (F_2), enquanto que f_0 é o mais robusto dos parâmetros (NOLAN, 1993; HOLLIEN, 2002; ROSE, 2002). Com relação à largura de banda dos formantes, Stevens (1998, p. 161) afirma que esse valor sofre influência direta das propriedades físicas do trato vocal e suas paredes e representa a atenuação do sinal resultante da rigidez da superfície, da fricção/viscosidade das paredes do trato e até mesmo do calor do ar por onde se propaga o sinal. Como são influências orgânicas, apresentam grande variabilidade intersujeitos e por isso serão medidas. Esses parâmetros acústicos serão obtidos em porções determinadas do sinal de fala e compõem o que Nolan (1983) refere como sendo parâmetros de curto termo, justamente por serem obtidos em porções curtas, em oposição a parâmetros de longo termo, que se referem a aspectos suprasegmentais e também àquelas características que definem a qualidade vocal.

Os parâmetros de curto termo serão obtidos sobre sons vocálicos orais, pois são as realizações tradicionalmente avaliadas na CL, e sobre sons consonantais nasais, pois são realizações que também apresentam ressonância formântica e cuja variabilidade propõe-se confrontar com a dos sons vocálicos aqui pesquisados.

As seções seguintes abordarão, primeiro, as características acústico-articulatórias particulares às vogais orais e, em seguida, aquelas particulares às consoantes nasais. Por fim, o capítulo encerrar-se-á com uma seção dedicada a descrever o uso dessas produções sonoras na situação forense.

2.3 VOGAIS

A proposta desta seção é caracterizar as vogais do ponto de vista acústico-articulatório e buscar achados de pesquisas que adotaram uma perspectiva acústico-articulatória para esses sons, principalmente no que diz respeito ao PB. Antes, porém, é necessária a compreensão do sistema acústico-articulatório no que se refere à relação íntima que as consoantes e as vogais mantêm entre si.

Em boa parte das línguas do mundo, as vogais podem ser pronunciadas isoladamente em contraposição às consoantes, que apenas podem ser pronunciadas acompanhadas de uma vogal (LADEFOGED, 1996, p. 281). Esse relacionamento entre consoantes e vogais é apresentado pelo Manual da Associação Fonética Internacional (doravante a ser referido apenas como IPA, 2009), que define a fala como uma sequência de estreitamentos, próprios da natureza consonantal, e aberturas, próprias da natureza vocálica, do trato vocal, que se caracteriza por ser a passagem através da qual o ar flui durante a articulação dos sons necessários para a comunicação verbal oral. Uma palavra como *bacana* ([ba'kəne], [bə'kəne] ou [ba'kẽne]) é um exemplo similar ao oferecido pelo IPA (2009, p. 6)²², em que ocorrem três oclusões do trato vocal, uma com os lábios, uma com a porção posterior da língua e uma com a porção anterior da língua, seguidas,

²² O exemplo referido no original é *banana* (com transcrição fonética referente à pronúncia em língua inglesa). Tomei um exemplo que, além de referir-se ao PB, oferece três pontos de articulação (o original oferece dois). Em razão de possível questionamento a respeito das transcrições das variantes da palavra escolhida, a opção de transcrição das vogais centrais no PB será tema de discussão adiante nesta seção.

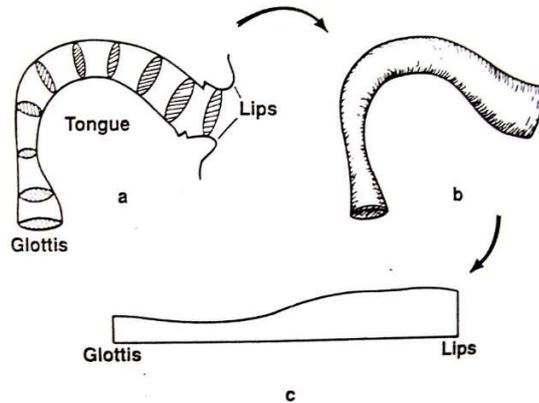
alternadamente, de solturas que permitem que o fluxo do ar percorra o trato sem obstáculos articulatórios.

As vogais são sons que ressoam pela cavidade oral sem a interrupção por processo articulatório, ou seja, em que o estreitamento do trato vocal ocorre em menor grau do que comparativamente às consoantes (IPA, 2009, p. 10). Assim, obedecem a critérios de classificação diversos daqueles que são utilizados para organizar as consoantes, como ponto de articulação, modo de articulação e vozeamento, mais evidentes com relação à participação das estruturas do trato na produção consonantal.

Os critérios de classificação para as vogais orais consideram o espaço estabelecido no trato vocal e o modo como as estruturas acomodam-se para conformar o trato para cada vogal. Dessa maneira, a abertura vertical da cavidade, a posição do corpo da língua em relação ao seu comprimento e o arredondamento dos lábios são os critérios de classificação aplicados aos sons vocálicos orais (IPA, 2009).

Para a análise acústica das vogais, o trato vocal é tomado como um tubo fechado na extremidade onde se encontram as pregas vocais e aberto na outra extremidade, a dos lábios. Dessa maneira, a estrutura é concebida como uma câmara de ressonância de formato complexo que, simplificada, pode ser considerada como tendo o formato de um tubo. É possível imaginar a retificação desse tubo, processo que se vê, na Figura 2.8, a seguir, para representar esquematicamente a relação de seu formato com o som produzido (KENT E READ, 1992). O posicionamento dos articuladores no interior do trato vocal determinará sua conformação e o conseqüente resultado acústico. Cada configuração resultará em ressonância específica e dessa forma uma produção vocálica específica.

Figura 2.8 – Derivação do trato vocal em função de área.

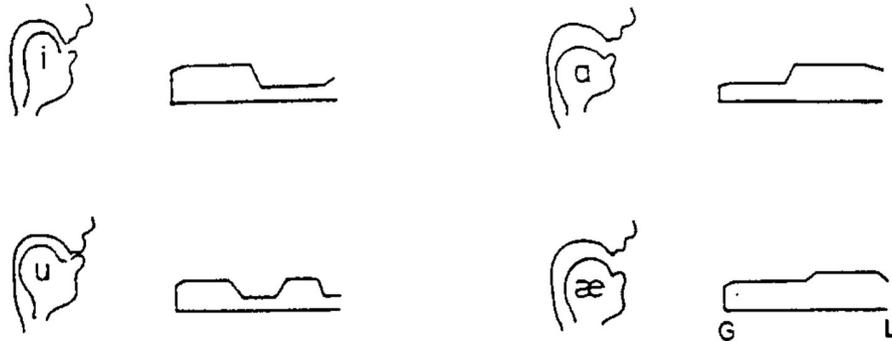


Fonte: Kent e Read (1992, p. 22).

Nota: Os diâmetros das seções são determinados à medida que se avança no trato, partindo da glote em direção aos lábios (a). O tubo curvo resultante da composição das seções (b) é retificado horizontalmente (c).

Os sons produzidos no interior do tubo entrarão em ressonância de acordo com o formato assumido. Assim, qualquer alteração no tubo realizada através dos movimentos articulatorios produzirá alteração no som resultante correspondente. A Figura 2.9, a seguir, ilustra esse princípio para quatro vogais do inglês. No quadrante superior esquerdo da Figura 2.9, observa-se que o posicionamento anteriorizado e alto da língua no trato vocal, à esquerda, resulta em uma configuração do tubo, que esquematicamente está representado à direita, com maior volume na porção próxima da fonte glótica e menor volume na porção mais próxima aos lábios, determinando o que acusticamente o ouvinte percebe como um som vocálico com características de [i]. Equivalentemente, observa-se, nos demais quadrantes da Figura 2.9, as articulações da língua no trato vocal e respectivas configurações de tubo referentes aos sons vocálicos [ɑ], [u] e [æ].

Figura 2.9 – Configurações ilustrativas do trato vocal.

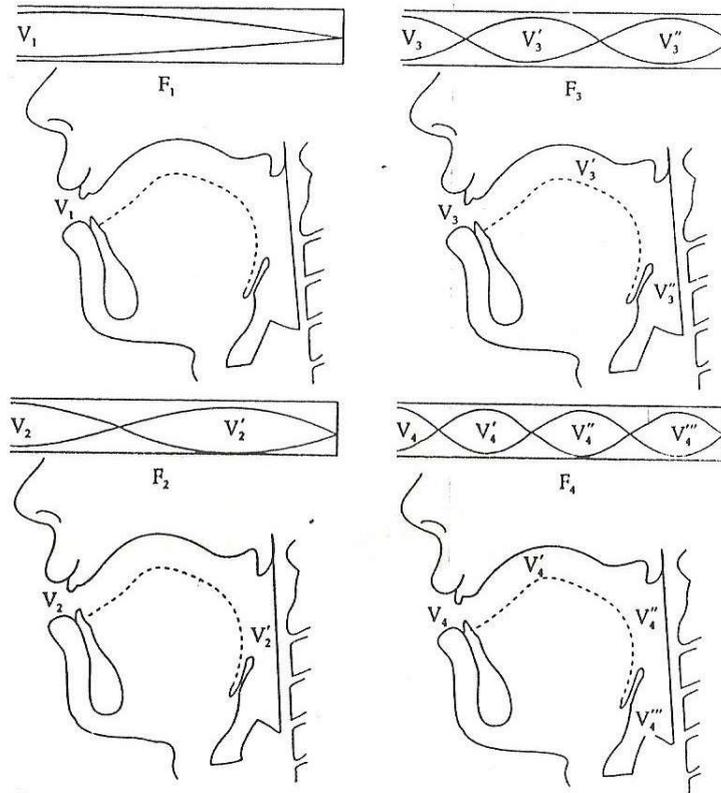


Fonte: Kent e Read (1992, p. 23).

Nota: Configurações articulatorias do trato vocal e respectivas funções de área idealizadas para as vogais do inglês [i] (he), [u] (who), [ɑ] (pa) e [æ] (map). G indica glote e L indica lábios.

As noções abordadas até aqui dizem respeito às características do trato vocal e suas estruturas, ou seja, o espaço físico por onde passa e é alterada a onda sonora originada na fonte glótica. Relacionando à Teoria Fonte-Filtro, a conformação do trato estabelece o tipo de filtro que exercerá a ação sobre a onda sonora, ou fonte. Para explicar como cada alteração, ou constrição, no trato vocal age sobre a onda sonora e as consequências acústicas dessa ação, a Teoria da Perturbação (CHIBA e KAJIYAMA, 1941) descreve a relação entre pressão e velocidade das partículas que transmitem a energia sonora através do trato, conforme ilustra a Figura 2.10 a seguir. Em um tubo uniforme, sem constrições, aberto em uma extremidade e fechado na outra, as ressonâncias naturais apresentam **máxima** pressão na glote e **máxima** velocidade nos lábios e, conseqüentemente, **mínima** velocidade na glote e **mínima** pressão nos lábios. Para cada ordem de ressonância, ou para cada formante, há uma relação estreita entre o comprimento do tubo e o comprimento da onda. Para o primeiro formante, o comprimento do tubo equivale a 1/4 do comprimento da onda. Para o segundo formante a relação é de 3/4 e de 5/4 e 7/4 para o terceiro e quarto formantes, respectivamente, como está expresso nos tubos representados na Figura 2.10.

Figura 2.10 – Acústica do trato vocal, seus nodos e antinodos.



Fonte: Johnson (1997, p. 99).

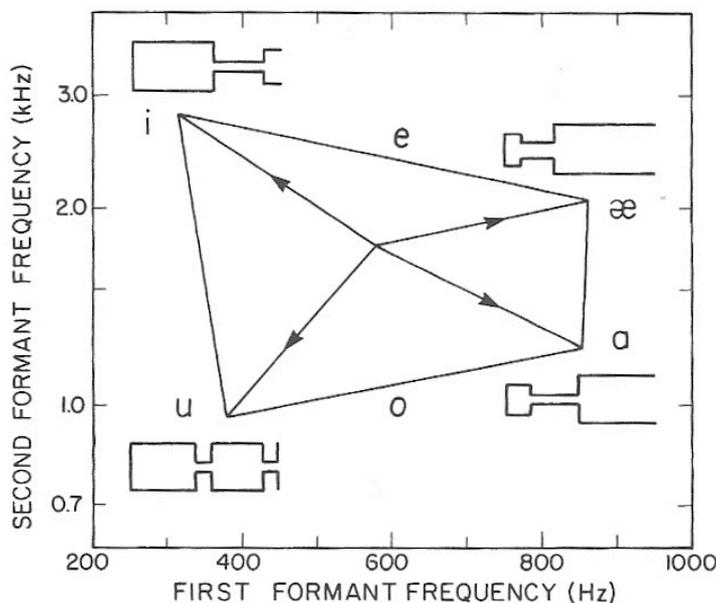
Para cada tubo na Figura 2.10 consta uma representação respectiva do trato vocal e a localização dos nodos, ou pontos de máxima velocidade (indicados pelos “V” na Figura 2.11), existindo um nodo para F₁, dois para F₂ e assim por diante. Apesar de não aparecerem indicados, os pontos de máxima pressão, ou antinodos, localizam-se na intersecção das curvas senoidais e, analogamente aos nodos, são um para F₁, dois para F₂ e assim por diante. Constrições no trato vocal têm efeito direto sobre o valor da frequência do formante, respeitando uma relação simples: constrições próximas aos nodos reduzem o valor da frequência do formante e constrições próximas aos antinodos elevam o valor da frequência do formante.

As vogais são classificadas acusticamente de acordo com a conformação do trato e sua consequência para valores dos formantes e de suas respectivas larguras de banda. Medidas da dimensão vertical do trato, influenciadas pela altura, ou abertura, do trato vocal estarão diretamente vinculadas aos valores de F₁, de forma que quanto mais aberta

estiver a cavidade, maior será o valor de F1 e quanto mais fechada, menor o valor de F1. As medidas da dimensão horizontal, influenciadas principalmente pela posição anteroposterior da língua, determinarão os valores de F2, de maneira que as vogais anteriores apresentarão valores de F2 maiores do que as posteriores.

Com o objetivo de ilustrar a relação da articulação com o dado acústico para diferentes vogais do inglês, apresenta-se o esquema de Stevens (1998) na Figura 2.11 a seguir. Três dimensões de informação são apreendidas desse gráfico. A articulação aparece representada, nas posições mais periféricas do gráfico, pelos esquemas de configurações do tubo, equivalentes aos esquemas apresentados na Figura 2.9, indicando as respectivas conformações do trato vocal. Os parâmetros acústicos de F1 e F2 determinam o posicionamento de cada som no espaço acústico. Por fim, a vogal respectiva está indicada junto a cada esquema. Assim, observa-se, na Figura 2.11, a vogal, sua conformação esquemática e a posição acústica.

Figura 2.11 – Relação de F1xF2 das vogais e os respectivos formatos de tubos ressonadores²³

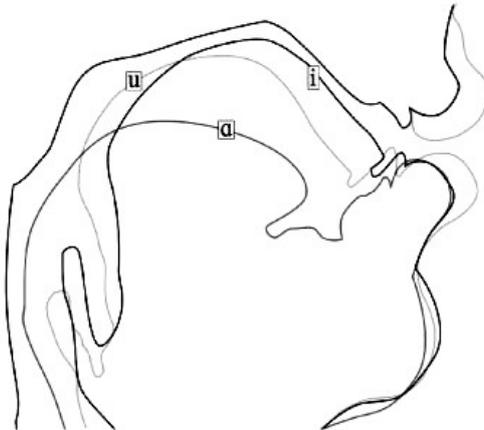


Fonte: Stevens (1998, p.286).

²³ A disposição dos eixos, com F2 na ordenada e F1 na abcissa, e a colocação do ponto de origem no canto inferior esquerdo do gráfico, são elementos comuns na representação cartesiana. Os trabalhos em Linguística, particularmente em Fonética, costumam representar os gráficos de F1xF2 com uma disposição própria, diferente da cartesiana tradicional, com a intenção de posicionar as produções vocálicas de forma análoga às suas conformações de trato vocal. Ao referir as vogais cardeais, a seguir, a questão será retomada com mais detalhamento a respeito de tal representação “invertida” e sua utilidade para os estudos em Fonética.

A transposição da noção articulatória para a respectiva representação fonética fica facilitada em função da evolução na tecnologia de obtenção de imagens em laboratório, o que se observa nas técnicas utilizadas por diversos estudos desde a década de 60, inicialmente por imagens de Raio-X, passando por tomografia computadorizada e chegando a imagens de ressonância magnética (doravante MRI, respeitando acrônimo da língua inglesa) e de ultrassom, conforme relata Stone (2011, pp. 10 et seq.). Particularmente, as imagens de MRI oferecem uma referência mais precisa do posicionamento da língua para a articulação das vogais e permitem uma representação gráfica como a da Figura 2.12, a seguir, montada a partir das respectivas configurações do trato vocal para as vogais [u], [i] e [a], onde as imagens obtidas com base em MRI servem de base para a produção do esquema.

Figura 2.12 – Posição de língua para a realização de três vogais.

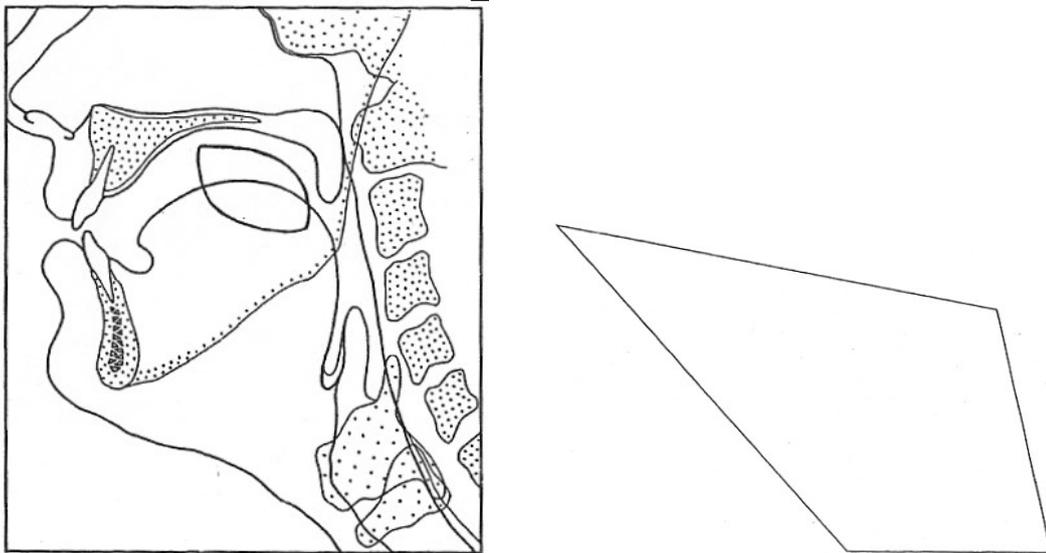


Fonte: Gick et al. (2013, pos.3326)
Nota: Obtida por superposição de MRI.

A relação espacial-articulatória das três vogais resulta em um triângulo vocálico imaginário onde tais vogais ocupam os vértices. Assim, tem-se uma ideia do espaço vocálico ocupado especificamente pelo falante modelo para elaboração da Figura 2.12. No sentido de representar as possibilidades das línguas de forma geral, o espaço vocálico é representado, como explica Laver (1994, p. 272), tendo como referência o ponto mais alto do corpo da língua e seu posicionamento na cavidade oral conforme a vogal produzida. Em uma visão sagital da cavidade, o conjunto das possibilidades de

posicionamento do ponto mais alto do corpo da língua ocupa uma área aproximadamente oval, que tradicionalmente se representa como um quadrilátero irregular, conforme apresentado na Figura 2.13, a seguir.

Figura 2.13 – Espaço vocálico segundo posição do ponto mais alto do corpo da língua.

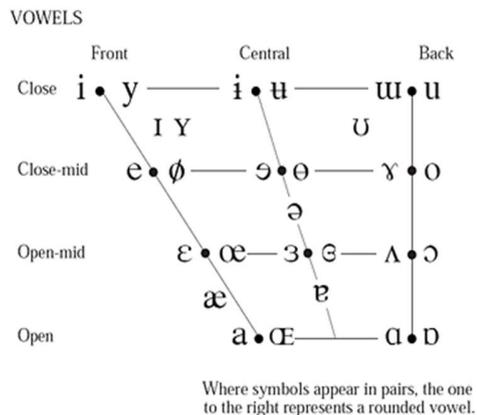


Fonte: Laver (1994, pp. 272-3)

Nota: À esquerda, o espaço vocálico descrito pelo posicionamento do ponto mais alto do corpo da língua; à direita, a representação esquemática desse espaço.

O quadrilátero apresentado na Figura 2.13, à direita, serve como referência para a localização das produções vocálicas existentes nas línguas do mundo, sendo que os vértices, partindo do mais alto e seguindo no sentido anti-horário, correspondem aos sons de [i, a, ɒ, u]. A Figura 2.14, a seguir, mostra a disposição completa das vogais segundo o IPA (2009), em que os elementos aparecem organizados dentro do espaço possível de produções vocálicas, respeitando os limites do quadrilátero exposto na Figura 2.13. Apesar das correções, uma vez que as vogais de altura similares são alinhadas horizontalmente com o objetivo de deixar a disposição simétrica, percebe-se a relação dos limites da área ocupada pelas vogais com o quadrilátero apresentado na Figura 2.13, além dos quais não é possível a realização de vogais.

Figura 2.14 – Trapézio com o alfabeto fonético das vogais.

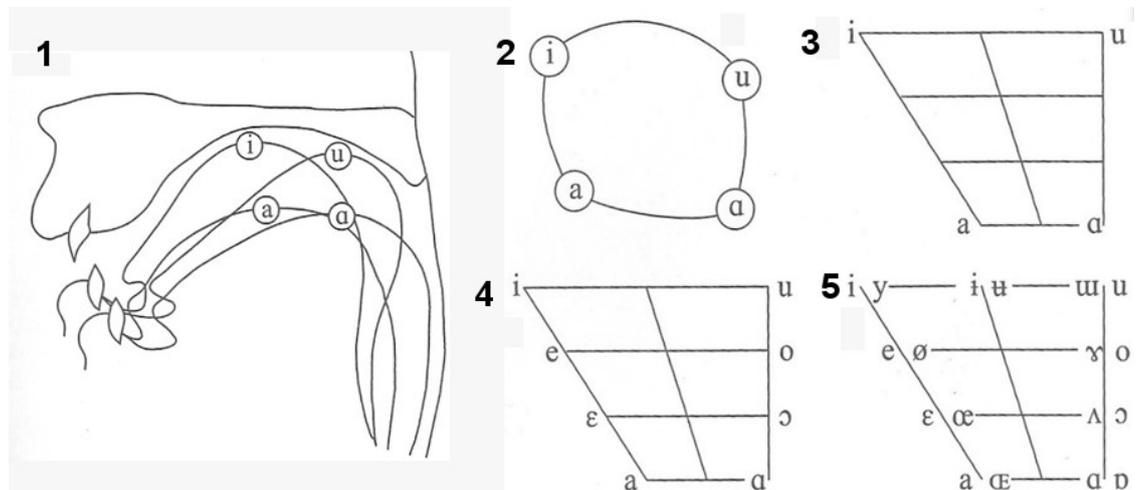


Fonte: IPA (2009, p. ix).

O IPA (2009, pp. 10 e 11) mostra a relação entre o espaço acústico-articulatório possível e o quadro do alfabeto de vogais através de raciocínio análogo, partindo do espaço vocálico estabelecido a partir das quatro extremidades de realização vocálica, de acordo com o posicionamento do ponto mais alto do corpo da língua, e, conseqüentemente, estabelecendo as margens de possibilidade de realização de vogais para os sistemas linguísticos do mundo. A seguir são dispostas as demais possibilidades de realizações vocálicas simetricamente dentro dos limites inicialmente estabelecidos, conforme apresentado na Figura 2.15 a seguir. O resultado dessa operação é o quadro completo das vogais, que oferece diretamente as características de anterioridade e abertura de cada vogal e, não tão diretamente, a característica de arredondamento, já que é por definição que as arredondadas permaneçam à direita do par em cada posição.

O quadro das vogais proposto por IPA (2009) baseia-se no instrumento das Vogais Cardeais (doravante VC), cujo histórico, quase centenário, está descrito por Laver (1994, pp. 272-277), que atribui a Daniel Jones a primeira proposta do sistema, elaborado inicialmente para servir de referência para pronúncia de vogais do inglês no ano de 1917. O sistema serve como referência genérica para sistemas linguísticos do mundo e sua elaboração segue critérios definidos, descritos a seguir.

Figura 2.15 – Representações do espaço acústico.



Fonte: Adaptado do IPA (2009, pp. 10 e 11).

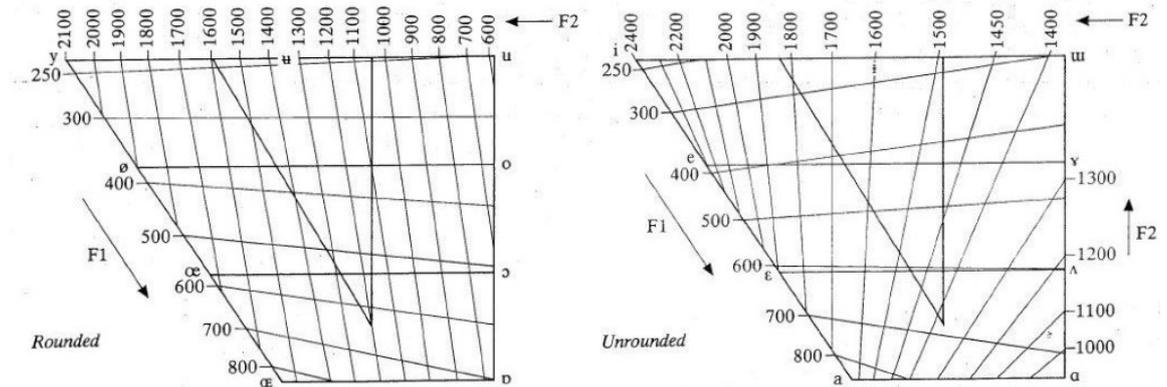
Nota: Seguindo a numeração indicativa, acima e à esquerda de cada representação do espaço vocálico, percebe-se o processo de estabelecimento da posição de referência para cada vogal

Catford (1977) descreve as regras para desenho do quadrilátero, seguindo um princípio de preenchimento que parte das extremidades para o centro do polígono, que tem inclusive suas proporções determinadas, com três unidades de altura por quatro de largura. Preenchendo as oito posições periféricas em primeiro lugar, o autor estabelece que tais vogais sejam as oito “cardeais primárias”, numerando-as de 1 a 8 (de [i] até [u]), partindo da anterior fechada, descendo e subindo novamente até a posterior alta, repetindo o procedimento para aquelas que seriam, então, as “cardeais secundárias”, internas às primeiras e seguindo o mesmo caminho, numeradas de 9 a 16 (de [y] até [ɯ]). As demais vogais são posicionadas conforme as posições das dezesseis primeiras, formando a carta completa das vogais. As VC primárias, conforme o exposto, são pontos de referência definidos através do limite articulatorio (e acústico) que seria possível atingir para o espaço das vogais. Seria, teoricamente, o sistema mais periférico possível (CATFORD, 1977, pp. 173-174).

As representações propostas e revistas até aqui estabelecem a relação entre a realização articulatória e sua representação fonética, sem que se tenha mostrado as referências acústicas (valores de F1 e F2) relativas a cada posição do quadro. A relação

entre a representação dos sons apresentados na Figura 2.15 e o seu esperado comportamento acústico está descrito por Catford (1988, pp. 161-162), uma vez que o autor estabelece os valores específicos de F1 e F2 para as VC primárias e secundárias, conforme apresentado na Figura 2.16 a seguir²⁴.

Figura 2.16 – VC com os respectivos valores para F1 e F2 (valores na tabela abaixo).



[i]	F1 240	F2 2400	[y]	F1 235	F2 2100
[e]	F1 390	F2 2300	[ø]	F1 370	F2 1900
[ɛ]	F1 610	F2 1900	[œ]	F1 585	F2 1710
[a]	F1 850	F2 1610	[æ]	F1 820	F2 1530
[ɑ]	F1 750	F2 940	[ɒ]	F1 700	F2 760
[Λ]	F1 600	F2 1170	[ɔ]	F1 500	F2 700
[ɻ]	F1 460	F2 1310	[o]	F1 360	F2 640
[u]	F1 300	F2 1390	[u]	F1 250	F2 595

Fonte: Catford (1988, pp. 161-162)

Importante ressaltar que a simetria do quadrilátero não respeita uma analogia estrita com a realidade física, uma vez que há menos espaço na porção posterior da cavidade oral do que na porção anterior, o que se percebe ao observar a inclinação das

²⁴ Retomando a nota anterior a respeito da orientação e disposição dos valores de F1 e F2 nos eixos da representação na Figura 3.11, que segue a norma cartesiana usual, e confrontando com a orientação e disposição desses valores nos eixos da Figura 3.16, percebe-se a inversão de eixos e mudança de ponto de origem. Nessa representação, F1 passa para o eixo vertical e F2 para o eixo horizontal e a origem dos eixos está à direita e acima, fazendo com que F1 aumente de valor de cima para baixo e F2 aumente de valor da direita para a esquerda. Essa estratégia de representação também é utilizada para os gráficos de dispersão e de espaços vocálicos, usuais em descrições linguísticas, que têm como objetivo manter o posicionamento dos elementos representados nos gráficos em analogia à sua articulação, tomando como referência um indivíduo em pé e voltado para esquerda.

linhas de F1 em relação às linhas referenciais do quadro, ocasionando maior amplitude de valores para as vogais anteriores do que para as posteriores.

Apesar da origem para a pronúncia específica, o quadro das vogais cardeais estabeleceu-se como um sistema referencial genérico. Assim, essas vogais não têm servido como alvos fonéticos, mas sim como parâmetros descritivos que têm sido confrontados pelos estudos de sons das línguas, como será visto adiante, na seção 2.3.2, em pesquisas da língua portuguesa, principalmente referentes ao PB.

A respeito de sistemas linguísticos vocálicos e sua distribuição de elementos, o modelo proposto por Lindblom (1986) descreve a ocupação do espaço acústico pelas vogais em 209 línguas e propõe uma equação para prever ocupação do espaço de acordo com o número de vogais, na tentativa de estabelecer universais linguísticos para o fenômeno. A respeito da categorização de vogais dos sistemas linguísticos, o autor afirma que se pode considerar uma matriz de possibilidades vocálicas com 42 posições, resultado da combinação de sete graus de abertura, três graus de posicionamento anteroposterior e dois graus de arredondamento ($7 \times 3 \times 2 = 42$). Dessas possibilidades, 37 estão preenchidas pelas vogais das línguas consideradas e 18 são as mais recorrentes nas línguas (LINDBLOM, 1986, p. 14). De maneira análoga, Ladefoged e Maddieson (2008) afirmam que são 60 as possibilidades vocálicas, resultado da combinação de cinco graus de abertura, três de posicionamento anteroposterior e quatro graus de arredondamento, esse com dois parâmetros, compressão e protrusão. Os autores, porém, admitem que é tão improvável de acontecer a ocupação de todas as posições que a hipótese pode ser considerada impossível (LADEFOGED E MADDIESON, 2008, p. 296). Ainda sobre as possibilidades de categorias vocálicas, o IPA apresenta 28 sons vocálicos em seu quadro, o que pode ser constatado através da observação da Figura 2.14.

As possibilidades de posicionamento dos sons vocálicos no espaço acústico são muitas se considerarmos a diversidade de sons das línguas. Os números recém apresentados expressam a dificuldade de se categorizar tantos elementos em um espaço diminuto como é a cavidade oral quando se busca a descrição de aspectos universais das línguas. O argumento de Koenig (2004) de que uma vogal classificada em uma língua

segundo uma representação não necessariamente ocupa a mesma posição no espaço acústico do que a mesma representação vocálica em outra língua, porém, merece consideração. Ou seja, não há que se pensar em um conjunto de vogais universais, ou símbolos vocálicos universais, mas sim em seres humanos que produzem e interpretam tais sons. Com o objetivo de oferecer alguma informação sobre o resultado acústico da produção vocálica e a categorização dos sons vocálicos, serão apresentados resultados e proposta discussão sobre essa questão na seção 5.3, Capítulo 5.

A noção de espaço da cavidade oral oferecida pelas medições acústicas vocálicas tem utilidade evidente como referencial concreto a respeito da articulação dos sons, servindo como ferramenta importante para as descrições das línguas e seu uso pode inclusive extrapolar o campo da linguística, como será comentado na seção seguinte.

2.3.1 Espaço acústico como medida aplicada

A partir de estudos dedicados a explicar fenômenos articulatórios fisiológicos bem como aqueles decorrentes de males neurológicos, pesquisadores inferem medidas espaciais a partir de informações acústicas vocálicas.

A área do quadrilátero formado pelas quatro vogais extremas do espaço vocálico do inglês americano, calculada a partir dos valores de F1 e F2 dessas vogais (aquelas apresentadas na Figura 2.15), serviu de parâmetro para o acompanhamento evolutivo de terapias para o tratamento de males como esclerose lateral amiotrófica e disartrias, doenças neurológicas que acometem a musculatura de forma geral e incluem entre seus sintomas a ininteligibilidade da fala, determinada pela imprecisão articulatória decorrente do comprometimento da musculatura da língua especificamente (KENT E KIM, 2003 p. 436).

Também seguindo a ideia de que medidas representativas da geometria vocálica podem servir de referenciais evolutivos, Sapir et al. (2010) avaliam a eficácia de tratamento sobre a fala de disártricos utilizando-se de quatro parâmetros: a Razão de Centralização Formântica (FCR, do inglês), a área de espaço vocálico, ou VSA, o LnVSA (VSA logarítmico) e a razão $F2i/F2u$ – valor do segundo formante de [i] dividido pelo valor do segundo formante de [u]. Os autores entendem que a fala disártrica é marcada por *undershoot* articulatório, que é a denominação para o movimento em direção a um alvo

articulatório mas que não atinge o alvo (THOMAS, 2011, p. 174), e que esse fenômeno ocasiona centralização na articulação das vogais. Assim, propõem que os parâmetros mencionados, por permitirem interpretação relacionada à centralização, podem indicar a eficácia dos tratamentos que buscam reforçar a precisão articulatória.

Todos os parâmetros recém referidos são operações matemáticas que envolvem mais de uma medida acústica para obter um parâmetro único. A aritmética, seja qual for o parâmetro escolhido, ocasiona naturalmente uma normalização dos dados, minimizando a variabilidade individual. Tal normalização é inclusive desejada dependendo do objetivo do estudo. No caso do estudo de Sapir et al. (2010), os autores estabeleceram uma comparação com o objetivo de determinar o melhor parâmetro, ou seja, aquele que sofre menos interferência dos efeitos de variabilidade interindividual, uma vez que seu estudo busca avaliar um grupo de maneira uniforme antes e após o tratamento para os sintomas de fala ocasionados pela disartria. O FCR obteve o melhor índice, seguido da razão $F2i/F2u$, do LnVSA e, por fim, da VSA.

Fora do campo das doenças com sintomas articulatórios, Pettinato et al. (2016) fazem uso da VSA para avaliar efeitos de idade, sexo e facilidade de comunicação em grupo de adolescentes com idades entre 9 e 14 anos e buscam explicações para VSA maiores do que os encontrados em adultos. Os autores entendem que os valores encontrados não se explicam apenas por diferenças nas dimensões do trato vocal entre crianças e adultos. Em suas considerações referem o *overshoot*, fenômeno oposto ao *undershoot* mencionado anteriormente, ou seja, uma sobrearticulação, característico da fala infantil como uma das possibilidades de explicação para o VSA maior nas crianças, em mais um exemplo da propriedade explicativa dessa medida no que se refere a inferências sobre a geometria do trato oral.

A respeito do prejuízo que a variabilidade individual²⁵ traz para a definição de conclusões com base nas medidas geométricas como indicador de melhora na fala do grupo de disártricos, VSA foi considerado pior em função de sua suscetibilidade à

²⁵ O prejuízo que a variabilidade individual ocasiona sobre o VSA é justamente um benefício para a CL, pois, como será mostrado na seção 2.4.4, Capítulo 2, a variabilidade individual é elemento que determina a qualidade do parâmetro de pesquisa. O mesmo tema será referido com relação às consoantes nasais, na seção 2.4.3, neste capítulo, ao abordar-se a variabilidade individual.

variação individual entre os sujeitos da pesquisa. Esse fator, que é indesejado para os estudos de Sapir et al. (2010), é justamente o que interessa ao estudo proposto por esse relatório. Dessa forma, espera-se que o VSA possa ser um parâmetro que ofereça contribuição para o conhecimento e prática forense, com aplicação específica na CL. O VSA, que nada mais é do que a área, no gráfico cartesiano formado pelo espaço acústico de $F2 \times F1$, do triângulo formado ao considerar-se as vogais [i, a, u] como seus vértices, isso considerando-se o PB, uma vez que para uma língua como o inglês, por exemplo, será formado um quadrilátero por quatro vértices vocálicos. Para o cálculo da área desse triângulo, uma das fórmulas disponíveis necessita das medidas das três arestas que compõem a forma geométrica, que são obtidas pelo cálculo das distâncias euclidianas (doravante DE)²⁶ entre as vogais extremas do espaço vocálico do PB. Uma vez que também são medidas que representam características anatômicas do trato vocal, tais distâncias serão testadas como parâmetros acústicos, de maneira análoga ao que ocorrerá com o VSA²⁷.

2.3.2 Descrições do espaço vocálico do PB

Os estudos a respeito do comportamento acústico das vogais do PB são ainda pouco numerosos na literatura, podendo ser citados aqueles particularmente dedicados à descrição de vogais orais (MIRANDA, 2012; ESCUDERO ET AL., 2009; RAUBER, 2006; PEREIRA, 2001; MORAES, CALLOU E LEITE, 1996; LIMA, 1991). Os trabalhos de Santos (2013), Sousa (1994), Seara (2000), Jesus (1999) e Cagliari (1977) estudam a nasalidade e incluem coleta de dados e descrição de vogais orais em suas pesquisas, com o objetivo de traçar paralelos entre os sistemas oral e nasal, porém apenas os dois primeiros coletam dados referentes às sete vogais tônicas orais.

Seara (2000) coleta dados das cinco vogais orais com correspondência às vogais nasais e Jesus (1999) e Cagliari (1977) coletam dados das três vogais orais das extremidades, ou seja, as duas altas e a baixa. Em abordagem clínica voltada para a voz,

²⁶ $\sqrt{(F2v1 - F2v2)^2 + (F1v1 - F1v2)^2}$ é a equação para DE, no espaço vocálico formado pelo plano cartesiano de $F2 \times F1$ e sendo $v1, v2$ pontos, ou vogais, nesse espaço.

²⁷ $\sqrt{p(p-a)(p-b)(p-c)}$ é a equação para o VSA, sendo a, b, c as arestas do triângulo, ou as três DE entre os três vértices do triângulo vocálico [a], [i] e [u], e p o semiperímetro, ou a soma das arestas dividido por 2.

porém oferecendo medidas formânticas vocálicas, Gonçalves et al. (2009) também apresentam descrição acústica do sistema vocálico tônico. Finalmente e mais recentemente, Barbosa e Madureira (2015), em um trabalho de cunho didático-experimental mais abrangente dos sons consonantais e vocálicos do português, incluem uma descrição das sete vogais orais. De forma geral, porém, os estudos referidos apresentam achados circunscritos a cada pesquisa. A diversidade de delineamentos de amostras e estratégias adotadas torna a equiparação dos valores acústicos médios encontrados nas pesquisas uma tarefa merecedora de ressalvas, que devem ser feitas de maneira a observar-se as particularidades metodológicas de cada estudo. Dessa maneira, não há garantia de que a igualdade/diversidade entre os estudos não seja fruto do perfil da amostra e/ou da metodologia adotados. Serão apresentados sumariamente os estudos, com exceção daqueles em que a apresentação de medidas não contemple o sistema vocálico completo do PB.

A fim de caracterizar acusticamente os sistemas tônico, pretônico e postônico do PB falado em cinco capitais brasileiras (Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife), Moraes, Callou e Leite (1996) extraíram amostras de vogais de um *corpus* de fala espontânea, composto por entrevistas informais com quinze locutores masculinos de formação universitária (três de cada área urbana), estratificados por três faixas etárias (25-35 anos, 36-56 anos e 56 anos em diante). Foram medidas, para cada falante, quinze ocorrências de cada vogal em cada posição tônica, totalizando 1.575 vogais tônicas, 1.395 vogais pretônicas e 675 vogais postônicas²⁸.

Os autores analisaram acusticamente os sistemas tônicos das cinco capitais e compararam seus resultados aos de Delgado Martins (1973) para as vogais tônicas do português europeu (doravante PE) e às VC, conforme valores descritos por Catford (1977). A comparação teve que comportar uma distinção relacionada à simetria de cada sistema, uma vez que o sistema de referência é composto por oito vogais e os sistemas de língua portuguesa são compostos por sete. Como se observa no Quadro 2.1 a seguir, há três linhas para os valores das vogais abertas, uma vez que as cardiais primárias

²⁸ Não há especificação na metodologia se são postônicas finais ou não-finais, porém no decorrer da apresentação dos resultados há uma menção a vogais finais, não ficando claro se tratava-se de uma referência a toda a amostra de 675 ocorrências.

possuem duas abertas e incluem uma mais anterior, representada por [a], e outra mais posterior, representada por [ɑ], enquanto que os sistemas da língua portuguesa apresentam apenas uma, representada por [e], existindo correspondência para comparação, no caso das abertas, apenas entre PE e PB. No caso das demais vogais, a correspondência se verifica entre os três sistemas comparados.

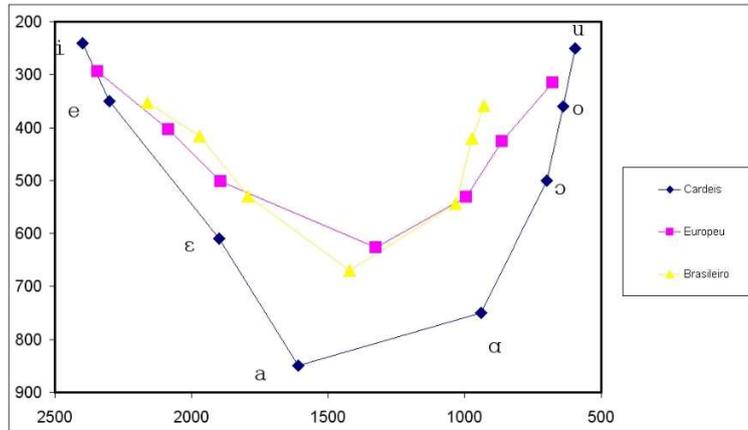
Quadro 2.1 – F1 e F2 para VC, PE e PB.

	VC		PE		PB	
	F1	F2	F1	F2	F1	F2
i	240	2400	294	2343	352	2162
e	390	2300	403	2084	416	1971
ɛ	610	1900	501	1893	530	1793
a	850	1610				
ɐ			626	1326	670	1420
ɑ	750	940				
ɔ	500	700	531	994	543	1033
o	360	640	426	864	419	973
u	250	595	315	678	358	928

Fonte: Moraes, Callou e Leite (1996).

A comparação dos três sistemas vocálicos fica mais clara quando se observa os gráficos referentes aos espaços vocálicos para cada sistema, conforme apresentados na Figura 2.17 a seguir. Além da evidente diferença, na porção central, entre o sistema de oito e os de sete elementos, onde as centrais do PB e PE são próximas entre si e são mais fechadas e em posição intermediária em relação ao F2 das duas abertas das VC, o PB tem suas vogais fechadas em posição mais centralizada do que as do PE. Esse último achado, porém, não encontra respaldo no trabalho de Escudero et al. (2009), que será abordado a seguir.

Figura 2.17 - Comparação dos sistemas de vogais (cardeais, PE e PB).

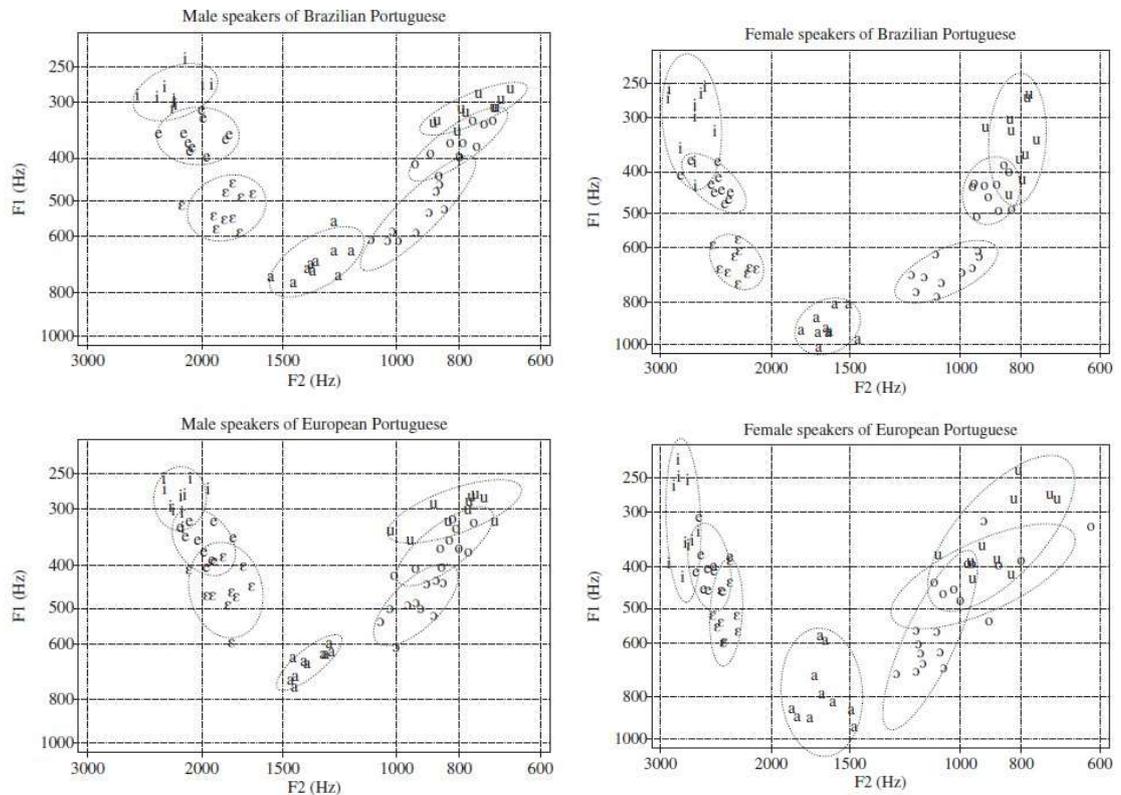


Fonte: Moraes, Callou e Leite (1996).

Escudero et al. (2009) realizaram estudo semelhante, no sentido da descrição do sistema vocálico, ao de Moraes, Callou e Leite (1996), porém com vogais tônicas medidas em ocorrências de fala produzida através de frases-veículo. O estudo compara as sete vogais periféricas do PE e do PB, investiga as diferenças entre sexos e constata propriedades intrínsecas das vogais. Os autores coletaram vinte elocuições para cada vogal de dez homens e dez mulheres para cada idioma do português, com falantes das cidades de São Paulo e Lisboa. Com relação às propriedades intrínsecas das vogais, os autores concluíram que as vogais altas são mais curtas e apresentam f_0 mais alto do que as vogais baixas. Os resultados que interessam à tese aqui proposta, porém, são aqueles que ficam evidentes nas Figuras 2.17 e 2.18 a seguir, e dizem respeito à comparação entre o sistema vocálico do PB e o do PE, bem como à comparação entre os sexos.

A Figura 2.18 a seguir apresenta quatro gráficos de dispersão, os dois à esquerda para dados dos homens, os dois à direita para os dados das mulheres, os dois de cima para dados do PB e os dois de baixo para os dados do PE, sendo que cada ocorrência se refere à média de F1 e F2 de cada indivíduo participante do estudo. Essa disposição dos gráficos permite a comparação entre os gêneros, no sentido horizontal, e entre os sistemas vocálicos das duas línguas, no sentido vertical.

Figura 2.18 - Comparação das distribuições de vogais (entre os dois gêneros e entre o PB e PE).

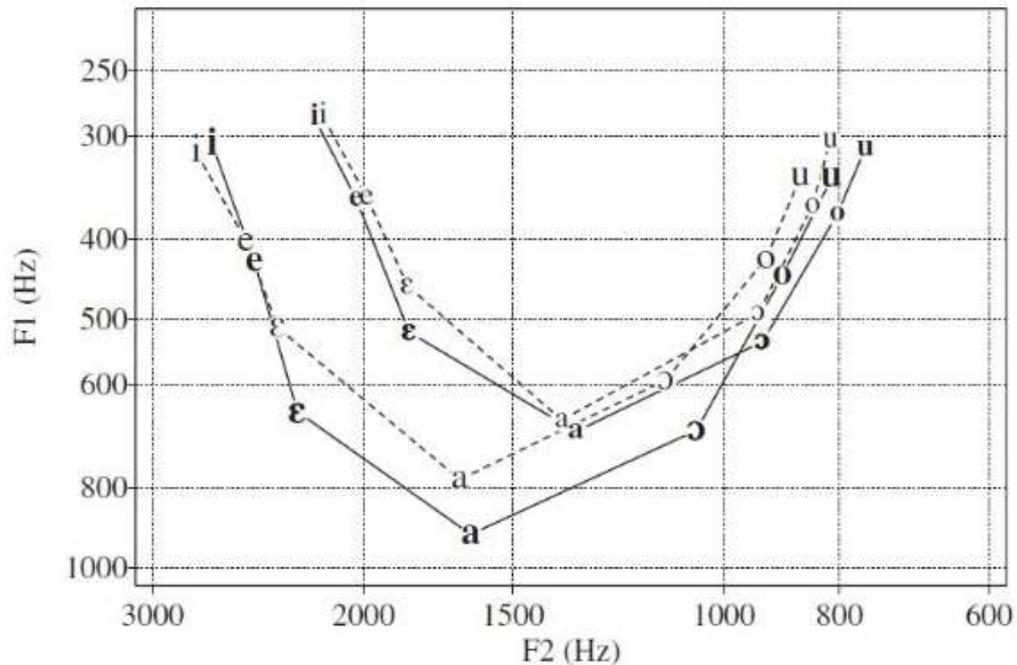


Fonte: Adaptado de Escudero et al. (2009, pp. 1384-5).

Com relação à variável gênero, as duas comparações, dentro das línguas, indicam o maior espaço acústico ocupado pelas mulheres do que pelos homens, o que se explica pela dimensão de trato menor, em média, na mulher do que no homem.

Como o objetivo do trabalho de Escudero et al. (2009) é a descrição e comparação de propriedades dos sistemas vocálicos de ambos dialetos, a Figura 2.19 a seguir é aquela que resume seus achados e foi obtida através do cálculo das médias de F1 e F2 para cada sexo e para cada dialeto, de forma que as quatro linhas desse gráfico se referem aos quatro gráficos apresentados na Figura 2.18, sendo que as linhas pontilhadas referem-se ao PE, as linhas cheias, ao PB, as fontes maiores, às mulheres e as fontes menores, aos homens. Como resultado em relação à comparação dos sistemas, os autores concluem que há indicativo de diferença para as vogais posteriores e que as anteriores são semelhantes para os dois dialetos.

Figura 2.19 - Comparação dos sistemas de vogais.



Fonte: Escudero et al. (2009, p. 1387).

Como se percebe nos trabalhos referidos, de Moraes, Callou e Leite (1996) e de Escudero (2009), aparentemente não há problemas em se traçar analogias entre as pesquisas. Uma atenção mais detida, porém, revela que o lapso temporal, a diversidade regional dos participantes e a estratégia de coleta, uma vez que se trata de fala espontânea no caso do primeiro e frase veículo no caso do segundo, prejudica o cotejo a ponto de permanecer a dúvida sobre a possibilidade do pareamento dos resultados.

Dentre os trabalhos referidos, são passíveis de comparação o de Pereira (2001) e Lima (1991). Isso se deve ao fato de que Pereira (2001), ao coletar dados de fala espontânea e de fala monitorada de cinco indivíduos de Florianópolis, intenta reproduzir o delineamento de Lima (1991). A diferença temporal é levada em consideração, uma vez que houve a intenção de verificar a possibilidade uma situação de mudança linguística em progresso. O fato de o estudo mais recente incluir a fala espontânea, o que no estudo de Lima (1991) não ocorreu, pode ser resolvido uma vez que há registro em Pereira (2001) de fala monitorada em separado, apesar de ser importante ressaltar que os instrumentos de coleta não foram iguais, ocasionando contextos circundantes às

vogais analisadas que não eram páreos. Pereira (2001) concluiu favoravelmente à sua hipótese de que a mudança em progresso se devia à influência migratória de habitantes porto-alegrenses para Florianópolis, com [i], [a] e [u] mais anteriores, [e] e [o] mais posteriores e [ɛ] e [ɔ] mais altas do que as vogais de Lima (1991).

Os estudos mais recentes mostram que a evolução tecnológica tem um efeito indireto. O acesso a equipamentos mais eficazes influenciou a capacitação do recurso humano, que desenvolveu estratégias de automação de procedimentos de análise, como fica evidente nos trabalhos de Rauber (2006) e Escudero et al. (2009), em que os autores descrevem e apresentam rotina para estabelecimento de parâmetros de configuração do LPC, que torna o processo de medição de formantes mais confiável. A comparação entre esses dois trabalhos também é plausível, uma vez que a metodologia é muito semelhante, com diferença nas regiões de origem/residência dos participantes, uma vez que Rauber (2006) coleta dados de seis participantes masculinos e seis femininos, divididos igualmente entre residentes no Rio Grande do Sul, Santa Catarina e Paraná, enquanto que Escudero et al. (2009) coletam dados de 40 participantes, 20 falantes do PB e 20 do PE, com dez homens e dez mulheres em cada dialeto. Os falantes do PB são todos de São Paulo e os do PE são de Lisboa.

Miranda (2012) coleta dados de oito participantes do PB e oito do inglês americano, divididos igualmente em homens e mulheres, para estudar a pronúncia de vogais do PB e do inglês americano por falantes nativos e de segunda língua, em relação às duas línguas, de forma cruzada. Para permitir a equivalência de seus resultados para comparações, a autora adota estratégia de normalização de Lobanov²⁹, a qual permite minimizar os efeitos individuais do trato vocal. O método de normalização calcula, para cada indivíduo e para cada vogal, os valores dos formantes separadamente e é considerado intrínseco ao falante e ao formante, uma vez que a equação considera valores do próprio falante e do próprio formante apenas, e extrínseco à vogal, pois a

²⁹ A fórmula da normalização é $F_n[V]N = ([V] - média)/DP$, onde $F_n[V]N$ é o valor normalizado para $F_n[V]$ (frequência formântica da vogal V ; *média* é o valor médio para o formante n do falante que produziu a ocorrência e DP é o desvio padrão do formante n . Há páginas na *Internet* que calculam a normalização se forem oferecidos os parâmetros.

fórmula considera a média e o desvio padrão de todas as vogais no cálculo para cada vogal.

A esse respeito, cabe ressaltar que os resultados sobre os quais foram aplicados processos de normalização não são interessantes para a aplicação forense, uma vez que a normalização justamente mascara, ou minimiza, as características individuais. Porém, trabalhos que apresentem gráficos de dispersão antes e após a aplicação da normalização poderiam servir como critério de avaliação das vogais para a CL. O critério é a dispersão das vogais normalizadas, de maneira que aquelas que apresentem maior dispersão, mesmo após a aplicação da normalização, podem ser consideradas mais apropriadas à CL, pois seriam mais resistentes à normalização. Essa dispersão, que eliminaria a influência individual, seria característica da língua e favoreceria, a princípio, a relação de variabilidade inter e intra indivíduo. O trabalho de Miranda (2012), porém, não apresenta os dados anteriores e posteriores à normalização de forma que permita essa consideração.

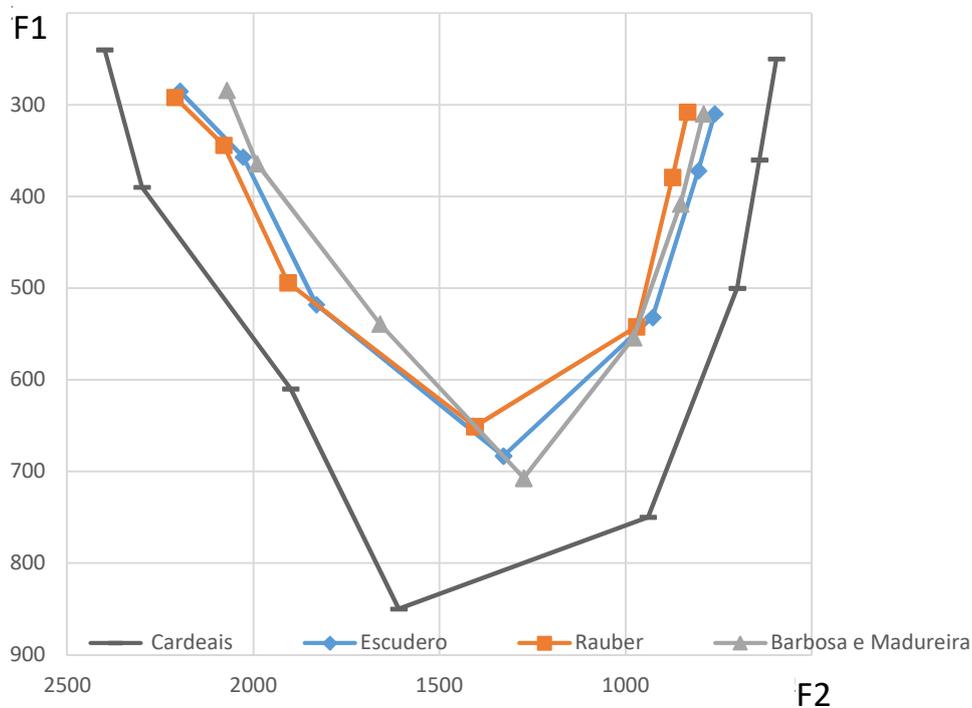
A respeito dos trabalhos de Santos (2013), que estuda a nasalidade e oralidade vocálica, e Sousa (1994), que foca na nasalidade vocálica, mas coleta dados de vogais orais, aponta-se o método de coleta como um importante entrave para comparação dos resultados, uma vez que o mais atual realizou entrevista de experiência pessoal e o mais antigo gravou a realização de logatomas isolados formados pela combinação da plosiva bilabial desvozeada e as sete vogais orais do PB. Ainda é relevante nesses estudos o lapso temporal de dezenove anos, além do o fator regional, uma vez que Santos (2013) coleta dados do PB de indivíduos trabalhadores rurais da região de Fortaleza das Nogueiras, Maranhão, com idade acima de 50 anos, e Sousa (1994) realiza quatro experimentos com nove indivíduos diferentes, sendo que o quarto experimento, em que coleta os dados vocálicos, é realizado com quatro indivíduos, um nascido em Recife, que se mudou para o estado de São Paulo; um Goiano que morou em Brasília e, à época residia em Campinas; um paulista e um paranaense.

Indicativo da diferença de recursos tecnológicos à época é o relato de Sousa (1994) a respeito das medidas formânticas. A autora explica que mediu os formantes, com um espectrógrafo digital *Kay Elemetrics Sona-Graph*, em duas faixas de frequência,

uma de zero a 8KHz e outra de zero a 4KHz, uma vez que na primeira, mais ampla, os saltos do cursor eram de 40Hz, enquanto que na segunda, mais estreita, os saltos eram de 20Hz, minimizando os erros de medição. As faixas de frequência não são um fator que foram considerados como elemento comparativo nesse levantamento bibliográfico, uma vez que, para os trabalhos posteriores ao de Sousa (1994) a tecnologia dirimiu os problemas de precisão das medidas.

Barbosa e Madureira (2015) comparam os resultados encontrados através do estudo de quatro indivíduos de diferentes estados com os resultados de Escudero et al. (2009). Assim, a comparação com os dados de Rauber (2006) mostra uma discreta similaridade entre os espaços vocálicos resultantes dos três trabalhos no que diz respeito aos sons posteriores, [ɔ], [o] e [u], e aos anteriores altos, [i] e [e], e discrepâncias relativamente à baixa, [a], e a médias baixa anterior [ɛ], como sugere a apresentação dos espaços vocálicos na Figura 2.20 a seguir.

Figura 2.20 - Comparação dos sistemas de vogais.



Fonte: O Autor (2017).

Por fim, apresentando um estudo que não foi realizado com a perspectiva linguística, mas sim voltado para a clínica dos distúrbios da voz, Gonçalves et al. (2009) mediram as vogais [i], [e], [ɛ], [a], [ɔ], [o], [u] sustentadas por alguns segundos, a partir de trechos de fala de cinco segundos, emitidos por 23 homens e 23 mulheres paulistanos. O objetivo dos autores era avaliar a função de transferência do trato vocal em relação à fonte glótica. Os autores concluíram que o [ɛ] é a vogal em que a conformação do trato tem menor influência de ressonância sobre o som produzido pelas pregas vocais. Tal achado pode não parecer relevante do ponto de vista linguístico descritivo, porém suscita curiosidade para quem se dedica à área forense, uma vez que pode sugerir uma aplicação para as medidas de fonte sonora à CL, especificamente no momento da coleta de material padrão, que não prevê gravação de vogal sustentada (a respeito do material padrão, ver Capítulo 2, seção 2.5.1), ou, mais relevante para esta tese, pode sugerir uma conformação vocálica em que a fonologia da língua exerça menor influência sobre as características individuais, favorecendo a manifestação de uma característica própria de cada indivíduo.

Para uma visão mais geral sobre os trabalhos levantados, apresenta-se um sumário dos delineamentos no Quadro 2.2, a seguir, de forma a favorecer o entendimento dessa diversidade. Assim, apresenta-se, nas colunas, da esquerda para a direita: **Autor**; **Ano**; **Norm**, a informação a respeito de os dados terem sofrido ou não algum processo de normalização; **N por sexo**, número de participantes por sexo, M para masculino e F para Feminino, que, somados, indicam o total daquele estudo; **N por região**, critério geográfico, com siglas para referência aos estados e a cidade aparecendo por extenso, com número de participantes à esquerda nos casos em que não fica óbvia a composição; **N PB masc**, número de participantes masculinos falantes do PB; modo de coleta da vogal; **Medição**, estratégia de medição da vogal em relação ao instante.

Quadro 2.2 – Visão geral dos trabalhos em acústica de vogais.

Autor	Ano	Norm	N por sexo	N por região	N PB masc	Modo de coleta da vogal	Medição
Escudero	2009	não	20 M 20 F	20 PE 20 PB	10	palavra isolada + frase veículo - logatomas	40% central da vogal
Rauber	2006	não	6M 6F	4 RS 4 SC 4 PR	6	palavra isolada + frase veículo - logatomas	40% central da vogal
Barbosa e Madureira	2015	não	4M	MG,SP,PI, CE	4	frase veículo - logatoma	ponto médio
Lima	1991	não	5M	Florianópolis	5	frases preparadas	ponto médio
Pereira	2001	não ³⁰	5M	Florianópolis	5	frases preparadas + vernaculo	ponto médio
Moraes, Callou e Leite	1996	não	15M	Recife, Rio, Sampa, Salvador, Poa	15	fala espontânea - NURC	ponto médio
Gonçalves et al.	1995	não	23M 23F	São Paulo	23	vogal sustentada	trecho
Miranda	2012	Lobanov	4M 4F	Vitória	4	frase veículo	trecho médio
Santos	2013	não	10 M 10 F	10 PE 10 PB (MA)	5	fala espontânea (entrevista)	ponto médio
Sousa	1994	não	9M	diferentes regiões	4	logatoma silábico (pV#) em frase veículo	ponto médio

Fonte: O Autor (2017).

Com a mesma intenção do quadro anterior, a saber, a de meramente oferecer um resumo dos resultados dos trabalhos apresentados e também a de permitir a conferência das comparações entre aqueles textos que se sugeriu a possibilidade, apresenta-se os valores médios de F1 e F2 no Quadro 2.3 a seguir. Juntamente com esses resultados, apresentam-se os valores referenciais das vogais cardeais, na primeira linha, para permitir alguma forma de comparação com os trabalhos que foram considerados sem condições para tal, devido aos seus delineamentos, de maneira que, no mínimo, seja possível confrontar cada resultado com os limites vocálicos que as vogais cardeais representam.

³⁰ Pereira refere o uso da escala MEL para a apresentação de seus dados em gráfico. Os dados, porém quando da apresentação dos valores médios tabulados, foram apresentados em Hz, antes da apresentação do gráfico e sem indicação de aplicação de qualquer normalização.

Quadro 2.3 – Vogais cardeais e resumo dos valores médios de vogais dos trabalhos referidos.

Autor	Fmt	[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
Cardeais	F1	240	390	610	850 750	500	360	250
	F2	2400	2300	1900	1610 940	700	640	595
Escudero 2009	F1	285	357	518	683	532	372	310
	F2	2198	2028	1831	1329	927	804	761
Rauber 2006	F1	292	344	494	651	542	379	308
	F2	2212	2080	1908	1405	971	874	834
Barbosa e Madureira 2015	F1	284	364	539	707	554	408	310
	F2	2072	1991	1660	1274	979	852	791
Lima 1991	F1	332	424	549	620	569	437	328
	F2	2070	1860	1567	1287	1009	915	788
Pereira 2001	F1	324	418	448	650	414	421	357
	F2	2252	1744	1617	1440	879	889	919
Moraes, Callou e Leite 1996	F1	353	417	530	670	544	420	359
	F2	2162	1971	1793	1420	1033	973	929
Recife	F1	397	449	561	730	568	454	403
	F2	2235	2004	1850	1460	1110	1031	939
Salvador	F1	320	390	480	643	503	400	346
	F2	2110	1883	1716	1303	986	953	930
Rio de Janeiro	F1	336	400	533	653	546	410	350
	F2	2196	2016	1833	1426	1020	976	943
São Paulo	F1	336	403	550	706	570	410	336
	F2	2053	1953	1750	1396	990	913	933
Porto Alegre	F1	373	440	526	616	530	423	356
	F2	2213	1996	1816	1513	1056	990	896
Gonçalves et al 1995	F1	298	407	588	754	580	412	345
	F2	2151	1956	1745	1279	947	833	800
Miranda 2012	F1	343	439	590	764	594	434	355
	F2	2113	2142	1782	1396	1075	956	984
Santos 2013	F1	322	486	614	726	581	440	384
	F2	2159	2008	1591	1369	998	832	865
Sousa 1994	F1	291	351	550	738	576	410	299
	F2	2111	2023	1826	1271	870	745	640

Fonte: Adaptado pelo Autor (2017).

Os valores apresentados no Quadro 2.3 referem-se, para cada estudo, ao grupo de participantes masculinos falantes do PB, a despeito de que alguns estudos tenham dedicado a tomar medidas de participantes femininas e do PE. As quantidades de participantes masculinos falantes do PB aparecem expressas na coluna “PB masc” do Quadro 2.2, anteriormente apresentado. A disposição dos dados no Quadro 2.3 está feita no sentido de permitir, pelo menos, olhar os valores de mesma natureza acústica, sem que se tire qualquer conclusão categórica a seu respeito, uma vez que cada um dos estudos apresentados refere, de uma maneira ou de outra, ao realizarem seus levantamentos bibliográficos, a importância em se analisar os dados com cuidado quando confrontados com outros trabalhos. No que diz respeito à CL e à prática forense, um

panorama como esse pode pelo menos servir de indicador para o delineamento de pesquisas a serem produzidas no sentido de oferecer material descritivo do PB como apoio para a formulação de perfis linguísticos.

Foram apresentados os princípios acústicos e articulatórios para a produção das vogais e como as medidas acústicas desses sons podem oferecer informações sobre a geometria do espaço acústico vocálico, com particular interesse no VSA e distâncias euclidianas. Em seguida foi ilustrado seu comportamento vocálico nos sistemas linguísticos do PB. Antes de abordar a aplicação forense, serão apresentadas as características referentes às consoantes nasais que podem colaborar com a CL.

2.4 CONSOANTES NASAIS

Do ponto de vista articulatório, os sons consonânticos nasais no PB são três: [m] [n] e [ɲ]. Quando considerada a posição inicial da sílaba, a gramática da língua aceita tanto no início como no meio de palavra a nasal bilabial (como em “mala” e “amor”) e a nasal alveolar (como em “nada” e “caneta”), enquanto que o palatal ocorre quase que exclusivamente no meio da palavra (como em “manhã”).

O comportamento acústico-articulatório das vogais foi explanado na seção anterior, que iniciou com a apresentação da relação consoante/vogal no fluxo da fala como um processo de interrupção seguido de abertura do trato vocal (IPA, 2009). Nessa perspectiva, a produção consonantal adquire uma característica de “ruído” que interrompe o fluxo vocálico. Se as vogais têm como sua característica de produção o trato vocal aberto e sem obstáculos à passagem do ar e ressonância, as consoantes se caracterizam justamente pela constrição do trato mediante diferentes estratégias, ou modos de articulação, e/ou locais, ou pontos de articulação (LADEFOGED, 1996). Algumas consoantes apresentam propriedades acústicas que as caracterizam semelhantemente ao simples ruído, como ocorre com as fricativas desvozeadas, enquanto outras apresentam características sonoras que permitem a extração de medidas acústicas de forma similar às vogais, como é o caso das nasais, fato que aponta para a possibilidade de incluir o estudo desses sons no presente trabalho.

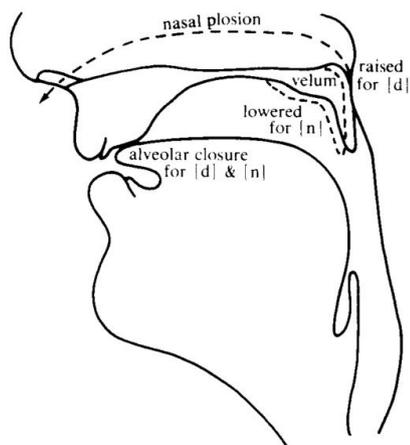
Diferentemente das demais consoantes, em que a articulação se dá através de um obstáculo ao fluxo aéreo que não altera o lugar de saída da ressonância, a cavidade oral,

as nasais se caracterizam pelo acoplamento da cavidade nasal à oral, desviando o caminho percorrido pela corrente de ar e a saída da ressonância. O acoplamento da cavidade nasal é comandado pelo véu palatino, composto por um conjunto de músculos cuja ação pode resultar na permissão, no caso dos sons nasais ou nasalizados, ou obstrução, no caso dos sons orais, à passagem do ar pela cavidade nasal. No caso das consoantes nasais, há um simultâneo impedimento do fluxo pela cavidade oral, sendo que o ponto de articulação, com as possibilidades do bilabial ao velar, determinará o formato da cavidade oral no acoplamento à nasal. Caso a cavidade oral não apresente obstáculo simultaneamente à abertura velofaríngea, o resultado será uma vogal nasal ou nasalizada. Como exemplo, a Figura 2.21, a seguir, ilustra a opção de permitir ou não a passagem pela cavidade nasal, para dois sons oclusivos com mesmo ponto de articulação: a plosiva alveolar vozeada e a nasal alveolar.

Em uma análise inicial, pode-se imaginar um sistema relativamente simples, em que uma válvula é aberta ou fechada para a passagem ou impedimento do fluxo aéreo. A respeito dessa suposta simplicidade do processo de produção da nasalidade, seja consoante ou vocálica, Ohala e Ohala (1993) afirmam

physiologically a nasal speech sound is quite simple: it just involves lowering of the soft palate to a degree sufficient to couple the oral and nasal cavities acoustically. With a concomitant oral closure, a nasal consonant is produced; without it, a nasal vowel³¹.

Figura 2.21 - Comparação de oclusivas nasal e oral com mesmo ponto de articulação.

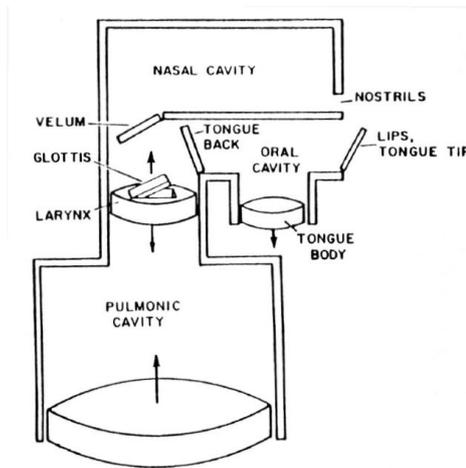


Fonte: Ladefoged (1996, p. 49).

³¹ Fisiologicamente, um som nasal de fala é bastante simples: apenas envolve o abaixamento do véu palatino até um grau suficiente para acoplar acusticamente as cavidades oral e nasal. Com um fechamento bucal concomitante, é produzida uma consoante nasal; sem ele, uma vogal nasal.

A representação esquemática do trato vocal, conforme apresentada na Figura 2.22 a seguir, evidencia um princípio que me levou à formulação de uma das hipóteses do presente trabalho: a característica estática da cavidade nasal e sua consequente invariabilidade, ou variabilidade mínima. O mecanismo apresentado com elementos simplificados de articulação, além de útil para o entendimento preliminar do processo de nasalização, serve para tentativas de modelamento do sistema (OHALA e OHALA, 1993).

Figura 2.22 – Mecanismo aerodinâmico do trato vocal.



Fonte: Ohala e Ohala (1993, p. 227).

Percebe-se, através da Figura 2.22, que o mecanismo de produção dos sons da fala compreende uma série de câmaras capazes de terem seu volume e, conseqüentemente, sua pressão de ar alterados de acordo com a disposição de estruturas, que funcionam segundo os princípios de pistões. Tais câmaras são interconectadas através de válvulas que regulam o fluxo de ar injetivo ou ejetivo. As paredes componentes da cavidade nasal são fixas e a participação da cavidade na produção dos sons da fala é controlada pela válvula vélica, permanecendo o volume da cavidade, assim, inalterado. A única variabilidade intraindividual provinda da articulação que se pode esperar em relação à cavidade nasal é aquela resultante da ação conjunta dos músculos do véu palatino e que determinará o grau de abertura dessa válvula do sistema. A esse respeito, os trabalhos que seguem tecem considerações que podem colaborar para o entendimento de seu funcionamento (KRAKOW E HUFFMAN, 1993;

GICK ET AL., 2013; BELL-BERTI, 1993; LIEBERMAN E BLUMSTEIN, 1988; CAGLIARI, 1977).

No que diz respeito ao movimento de abertura do véu palatino, o processo não é uma simples abertura ou fechamento da passagem, mas pode ser efetuado através de uma diversidade de movimentos. Krakow e Huffman (1993, p 9) afirmam que a contração do elevador do palato é principalmente observada na elevação e o relaxamento é principalmente observado no abaixamento do véu palatino, quando examinadas cadeias de sons orais e nasais. Seus exames da função velofaríngea através de eletromiografia ainda incluíram os músculos palatoglosso, palatofaríngeo, constritor faríngeo superior e constritor faríngeo médio.

Além de referirem uma série de músculos envolvidos na função velofaríngea, Gick et al. (2013) apontam o elevador do palato como o principal músculo para a realização do vedamento velofaríngeo e afirmam que as estratégias de fechamento são diversas, havendo a participação de um ou mais músculos na atividade motora, que podem realizar um movimento articulatoriamente mais simples, como um “alçapão”, ou mais complexo, como um esfíncter, existindo estratégias intermediárias entre essas duas. Bell-Berti (1993, p. 67) ainda refere a participação do músculo palatoglosso em uma provável função de abaixamento do véu palatino, contrapondo as hipóteses de que o abaixamento se daria apenas por relaxamento da musculatura.

Estudos mostram que a função velofaríngea será subestimada se apenas for considerada como uma mera porta que permite ou impede a passagem do ar pela cavidade nasal. Lieberman e Blumstein (1988) indicam o elevador do palato como o músculo responsável pela nasalidade, que tensiona quando da produção oral e relaxa quando da produção nasal. O seu movimento, porém, não é simplesmente binário e sim gradual. Os autores referem que as vogais [i] e [u], por seus traços de altura, exigem um maior tensionamento do músculo do que a vogal [a]. Ou seja, o vedamento nasal é mais exigido quando da produção das vogais altas.

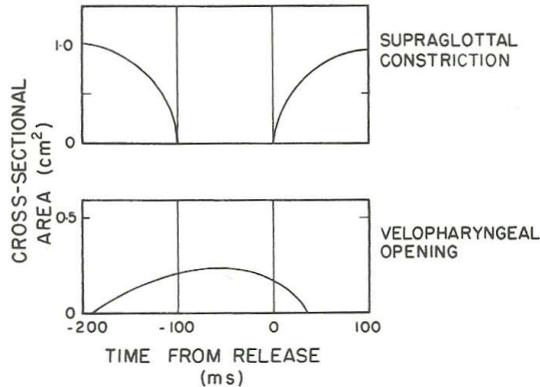
Cagliari (1977, p. 159) também menciona a existência de diferentes graus de abertura e propõe uma escala que varia da posição mais alta do véu palatino até sua posição mais baixa, totalizando onze graus de abertura, conforme critérios de pressão aérea: 1 – sopro, referente à posição mais alta; 2 – oclusivas desvozeadas; 3 – oclusivas

vozeadas; 4 – fricativas desvozeadas; 5 – fricativas vozeadas; 6 – vogais orais fechadas; 7 – vogais orais abertas; 8 – vogais nasalizadas fechadas; 9 – vogais nasalizadas abertas; 10 – nasais; 11 – respiração, referente à posição mais baixa. Bell-Berti (1993, p. 69) ordena os sons conforme a posição vélica – iniciando na posição mais baixa do véu e terminando na mais alta – sendo a mais baixa para as consoantes nasais, subindo gradualmente para as vogais baixas, vogais altas (as nasais antes das orais) e chegando na posição mais alta para as obstruintes.

Depois de terem sido abordadas as estratégias articulatórias, apresentar-se-ão as respectivas manifestações acústicas das consoantes nasais. Por serem sons de características soantes, em que há vibração das pregas vocais ao mesmo tempo em que há uma conformação do trato vocal que permite a ocorrência da ressonância, as consoantes nasais apresentam formantes com energia que, apesar de ser menos intensa que a das vogais, permite sua análise espectrográfica clara. Os formantes relativos à cavidade nasal aparecem em regiões características, de forma geral, em torno dos 250Hz e outra em torno dos 1000Hz, segundo Stevens (1998). Além disso, seu espectro mostrará a presença de bandas de frequência onde haverá a atenuação da energia resultante na cavidade oral: os antiformantes. Tal atenuação se deve à participação da cavidade nasal no sistema de ressonância, onde as antirressonâncias provenientes da região nasal atuam como atenuadoras das ressonâncias orais (FANT, 1960; KUROWSKI e BLUMSTEIN, 1993; OHALA e OHALA, 1993; STEVENS, 1998).

Stevens (1998, p. 488 et seq.) descreve o acoplamento das cavidades e sua ação sobre as ressonâncias resultantes da emissão das consoantes nasais bilabial, alveolar e velar, sempre em contexto intervocálico. Em primeiro lugar, o autor relaciona a ação da constricção supraglotal e da abertura velofaríngea, através das medidas de área das seções transversais respectivas, em um intervalo de 300ms em que o ponto de origem é a soltura (tradução minha para “release”), conforme o esquema apresentado na Figura 2.23, a seguir.

Figura 2.23 – Relação da constrição supraglotal e abertura vélica.



Fonte: Stevens (1998, p. 488).

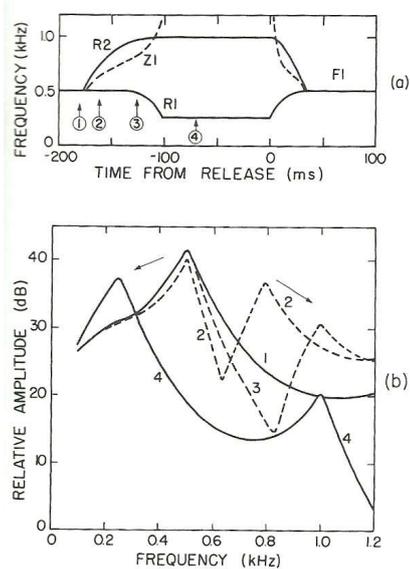
Nota: *Cross-sectional área*: área da seção transversal; *Supraglottal constriction*: constrição supraglotal; *Velopharyngeal opening*: abertura velofaríngea; *Time from release*: tempo desde a soltura.

Na Figura 2.23 fica clara a relação inversa de fechamento oral/abertura vélica, sendo que os terços do gráfico se referem à sequência vogal-nasal-vogal. A fase de abertura vélica, ou abaixamento, é mais lenta do que a fase de fechamento, fato constatado pela assimetria da curva.

Stevens (1998), tomando como referência temporal os instantes apresentados na Figura 2.23, explica o comportamento das ressonâncias através da análise da sua evolução na faixa de tempo estabelecida (300 ms). Esse mesmo princípio é seguido nas três figuras que seguem, de modo que, nos gráficos que apresentam, no quadro superior de cada figura, a faixa de frequência em torno do F1 da vogal e os instantes marcados por números referem-se às curvas numeradas no gráfico dos espectros apresentado no quadro inferior da respectiva figura.

A Figura 2.24, a seguir, refere-se aos instantes de passagem da vogal para a consoante nasal, neste caso, sem discriminação entre o ponto de articulação. Percebe-se como a ressonância relativa a F1 da vogal gradualmente bifurca-se e permanece com duas regiões de ressonância determinadas durante a consoante nasal.

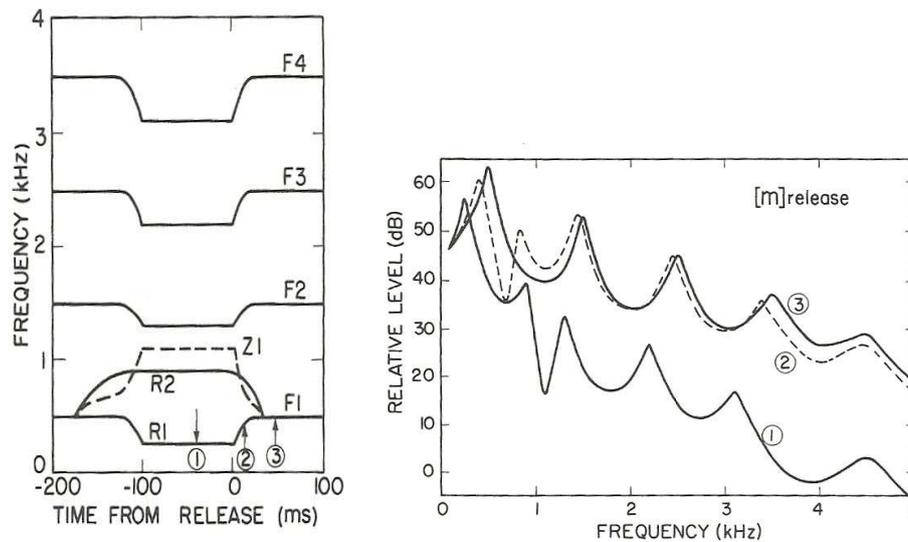
Figura 2.24 – Comportamento das ressonâncias no início da nasal (“implosion”, no inglês).



Fonte: Stevens (1998, p. 491).

Movimento inverso ao observado na Figura 2.24 é o apresentado nas Figuras 2.25 e 2.26 a seguir, com as regiões de ressonância da nasal voltando a se encontrar na faixa referente ao F1 da vogal. A Figura 2.25 refere-se à consoante nasal bilabial e a Figura 2.26, à alveolar. Stevens (1998) refere a discreta diferença percebida entre os gráficos dos espectros da bilabial e da alveolar nos instantes de número 1.

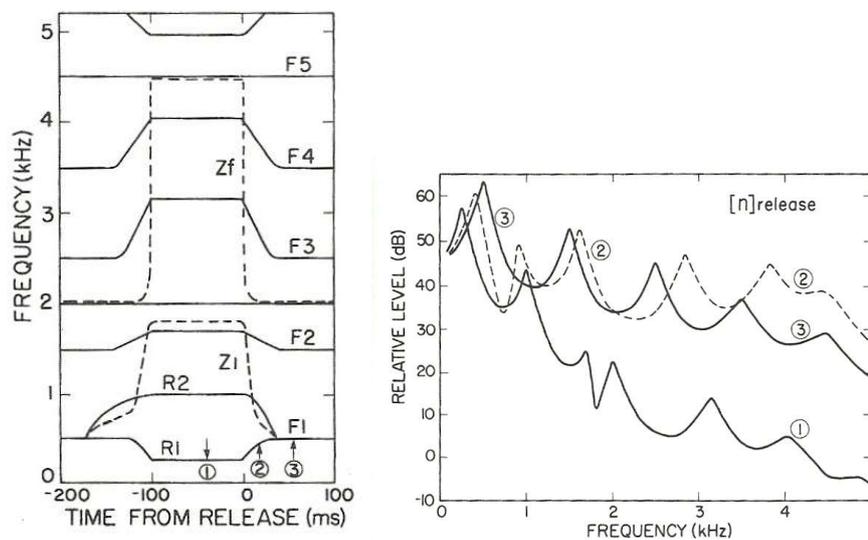
Figura 2.25 – Comportamento das ressonâncias na soltura de [m].



Fonte: Adaptado de Stevens (1998, p. 495 e 497).

A faixa de ressonância indicada por Z1 (linha pontilhada da imagem à esquerda) nas duas figuras refere-se ao zero da função de transferência da cavidade nasal, o antiformante, visível nos espectros (quadro à direita em cada figura) como um vale mais proeminente que os demais, em torno dos 1KHz para a bilabial e em torno dos 2KHz para a alveolar.

Figura 2.26 – Comportamento das ressonâncias na soltura de [n].



Fonte: Adaptado de Stevens (1998, p. 500 e 501).

Os estudos a respeito da nasalidade consonantal buscam estabelecer os invariantes linguísticos que caracterizariam as consoantes nasais como classe sonora e também aquelas características acústicas particulares a cada som, [m], [n] e [ŋ], principalmente os dois primeiros (FUJIMURA, 1962; KUROWSKI e BLUMSTEIN, 1993).

A caracterização individual dos sons nasais pode se dar pela localização dos antiformantes, mas não se deve excluir a importância da transição formântica para as vogais seguintes como estratégia para essa individualização. O murmúrio nasal é uma porção do contínuo acústico que é influenciada pela consoante nasal, sendo tarefa complexa seu estudo, o que exige cuidado devido ao efeito da variabilidade contextual e individual (FUJIMURA, 1962, p. 1875). Kurowski e Blumstein (1993) estudam os invariantes linguísticos relacionados ao ponto de articulação das nasais, partindo de analogia com achados relativos aos pontos de articulação das plosivas e indicam a

dificuldade em determinar invariantes relacionados ao modo de articulação. As autoras ainda referem limitações nos achados relativos à variabilidade dos contextos vocálicos ao apontar a maior variabilidade dos parâmetros avaliados nos contextos envolvendo as nasais seguidas de [i] do que quando seguidas de [a] e [u]. Também referem as diferenças nos tamanhos dos aparatos vocálicos individuais como complicadores para os estudos.

Para o PB, Sousa (1994) e Seara (2000) testam o comportamento acústico das consoantes nasais com o intuito de entender a nasalidade de forma geral e, assim, dedicam-se ao seu foco de pesquisa, a nasalidade vocálica.

Sousa (1994) realiza quatro experimentos para estudar a nasalidade no PB. No terceiro experimento, dedicado exclusivamente à medição dos valores formânticos das consoantes nasais, a autora mede os formantes nasais para os três sons, [m] e [n] em início de palavra, e [ŋ] intervocálico, precedendo as sete vogais tônicas, em palavras eliciadas através de frase-veículo. Seus resultados são consistentes com os estudos para as consoantes nasais, que apresentam FN1³² abaixo dos 300Hz como o formante mais proeminente.

Seara (2000) realiza uma revisão do comportamento acústico das nasais, medindo inclusive outros parâmetros além dos valores dos formantes e largura de banda, como intensidade e duração. Apesar de concentrar seu foco principalmente sobre o estudo da nasalidade das vogais, a autora dedica um capítulo à análise das consoantes nasais palatal, bilabial e alveolar na busca por parâmetros invariantes acústicos que permitissem caracterizar cada um dos sons diferentemente, estabelecendo a distinção entre as três consoantes. Após descartar a análise da nasal palatal, que apresenta a intervenção de uma vogal [ĩ] muito nasalizada e que dificultava a extração de medidas, avalia elocuições em contextos tônico e em contexto átono, eliciados através de frases-veículo, de cinco falantes masculinos florianopolitanos. Ressalva-se que a autora utilizou uma palavra para cada contexto, “mata, mito, muda, nata, Anita, numa” para o contexto tônico e “tâmara, química, túmulo, cânula” para o contexto pretônico, deixando inexplicada a decisão de não ter coletado contextos pretônicos para [na] e [ni]. Seus resultados para as consoantes

³² Primeiro formante da consoante nasal. Ver nota 33.

nasais aparecem resumidos nos quadros apresentados a seguir. A autora analisa para as consoantes nasais, além de amplitude e duração, f_0 , formantes nasais e largura de banda dos formantes, separadamente para [m] e [n] antes de [a, i, u].

Para a obtenção dos formantes das nasais³³, Seara (2000) utilizou-se da estratégia, também apontada por Thomas (2011, p. 43), de estimar número adicional de coeficientes do LPC para a análise dos formantes³⁴. Para os formantes das vogais orais e transição de formantes consonantais são necessários de dez a doze coeficientes para o LPC, quando se estima cinco formantes em uma faixa de análise de 5000Hz para um adulto masculino típico. Quando a intenção for analisar consoantes nasais, deve-se estimar número adicional de coeficientes para que o LPC possa oferecer um acoplamento adequado em relação ao espectro do FFT.

O Quadro 2.4 a seguir apresenta os resultados para [m] diante de [a] e, dentre os quadros posteriores a esse, apresentam-se os resultados conforme segue: o Quadro 2.5 apresenta os resultados para [m] diante de [i]; o Quadro 2.6 apresenta os resultados para [m] diante de [u]; o Quadro 2.7 apresenta os resultados para [n] diante de [a]; o Quadro 2.8 apresenta os resultados para [n] diante de [i]; e o Quadro 2.9 apresenta os resultados para [n] diante de [u].

³³ Os formantes nasais, no caso das consoantes, são todos os formantes obtidos pelo LPC, conforme aparecem nos quadros de 3.4 a 3.9, diferentemente das vogais nasais, em que há a ocorrência dos formantes vocálicos simultaneamente aos formantes nasais, o que exige a comparação com o correlato vocálico oral para que, através da comparação dos dois resultados, infira-se o formante nasal.

³⁴ A estratégia utilizada por Seara (2000), além de estimar número adicional de coeficientes para o LPC, sobrepunha essa curva ao FFT, para, além de determinar os formantes nasais, determinar as regiões dos antiformantes nasais. Tal abordagem não foi adotada no presente estudo em função da quantidade de dados.

Quadro 2.4 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [a] com suas respectivas larguras de banda. [Hz]

Contextos tônico e átono											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	126,97	261,06	44	840,09	119	1789,86	277	2463,20	303	3526,69	252
	135,94	243,97	81	600,13	99	1091,88	162	2252,00	231	3311,43	185
Desvio-Padrão	15,67	24,49	10,33	81,05	45,29	329,85	39,74	232,56	45,63	244,84	79,42
	22,79	35,43	56,28	109,68	55,54	105,06	72,78	199,25	83,26	128,17	-
Variância	245,60	599,90	106,8	6568,7	2051	108800	1579	54086,2	2082	59948	6307
	519,35	1255,38	3168	12031	3085	11037,6	5297	39701,5	6932	16427,4	-
Coeficiente Variação	12,3%	9,38%	23,5%	9,65%	35,5%	18,4%	14,3%	9,44%	15%	6,94%	31,5%
	16,76%	14,52%	69,5%	18,28%	56,1%	9,62%	44,9%	8,85%	36%	3,87%	-

Fonte: Seara (2000, p. 170).

Quadro 2.5 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [i] com suas respectivas larguras de banda. [Hz]

Contextos tônico e átono											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	123,20	257,97	33	918,75	318	1661,41	162	2245,77	388	3578,82	313
	144,00	252,14	54	821,57	66	1585,14	329	2274,77	189	3630,50	375
Desvio-Padrão	10,44	22,53	8,33	160,85	52,29	273,88	70,63	121,62	47,82	273,66	118,53
	24,26	33,85	36,45	117,55	7,15	169,92	40,07	179,52	53,19	241,08	83,13
Variância	109,05	507,50	69,39	25873,8	2733,80	75010	71,17	14791,8	2287,14	74892,5	14050,7
	588,65	1145,89	1528,78	13818,4	51,17	28872,9	1605,48	32228,5	3993,53	58121,4	6910,87
Coef. Variação	8,47%	8,73%	27,77%	17,51%	16,44%	16,48%	43,60%	5,41%	12,32%	7,65%	37,87%
	17,33%	13,42%	67,50%	14,31%	10,83%	10,72%	12,18%	7,89%	28,14%	6,64%	22,17%

Fonte: Seara (2000, p. 171)

Quadro 2.6 – Frequência fundamental e frequência dos formantes nasais de [m] diante de [u] com suas respectivas larguras de banda. [Hz]

Contextos tônico e átono											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	130,43	261,89	60	677,81	302	1746,89	274	2477,33	309	3461,28	303
	144,71	253,48	52	675,69	324	1723,21	279	2646,23	130	3504,48	-
Desvio-Padrão	12,65	30,87	20,43	96,63	112,66	402,90	115,47	146,50	91,09	307,80	90,06
	33,39	29,78	17,04	119,61	53,19	412,24	38,93	140,08	63,95	243,23	-
Variância	160,10	953,20	417,46	9337,60	12692,4	162332	13333,1	21462,1	8296,85	94742,3	8111,61
	1114,80	887,10	290,24	14307,8	2829,45	169942	1515,45	19622,5	4089,23	59158,8	-
Coef. Variação	9,70%	11,79%	34,05%	14,26%	37,31%	23,06%	42,14%	5,91%	29,48%	8,89%	29,72%
	23,07%	11,75%	32,77%	17,70%	16,37%	23,92%	13,95%	5,29%	49,19%	6,94%	-

Fonte: Seara (2000, p. 171).

Quadro 2.7 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [a] com suas respectivas larguras de banda³⁵. [Hz]

Contexto tônico											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	123,83	259,51	58	1265,03	295	2226,94	362	2711,70	349	3552,74	310
DP	14,92	23,97	30,86	133,02	111,01	144,49	100,2	132,79	73,92	130,98	92,60
Variância	222,66	574,79	952,1	17693,4	12324,2	20878,0	10048,1	17632,5	5464,16	17157,0	8574,60
CV	11,47%	9,24%	53,21%	10,51%	37,63%	6,49%	27,69%	4,90%	21,18%	3,69%	29,87%

Fonte: Seara (2000, p. 172).

Quadro 2.8 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [i] com suas respectivas larguras de banda. [Hz]

Contexto tônico											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	133,86	247,26	67	1094,03	281	2003,71	277	2782,44	241	3759,42	390
DP	25,66	36,94	22,38	252,69	79,23	219,90	55,78	155,32	77,35	139,29	48,85
Variância	658,42	1364,73	500,86	63854,1	6277,65	48355,8	3111,65	24125,3	5983,86	19402,3	2386,33
CV	19,17%	14,94%	33,40	23,10%	28,20%	10,97%	20,14%	5,58%	32,09%	3,70%	12,53%

Fonte: Seara (2000, p. 172)

Quadro 2.9 – Frequência fundamental e frequência dos formantes nasais de [n] diante de [u] com suas respectivas larguras de banda. [Hz]

Contextos tônico e átono											
Estadísticas	F0	FN1	B1	FN2	B2	FN3	B3	FN4	B4	FN5	B5
Média	131,80	273,37	57	980,82	81	1912,70	223	2745,15	144	3529,50	-
	149,23	258,71	28	1047,17	218	2204,72	261	2819,65	57	3517,72	-
Desvio-Padrão	14,22	21,99	12,50	141,09	25,87	301,40	21,42	335,62	46,63	277,46	-
	40,60	32,21	4,44	93,15	54,58	217,55	42,07	247,15	20,50	268,99	-
Variância	202,20	483,70	156,36	19906,3	669,46	90841,2	458,69	112643	2174,14	76984,6	-
	1648,71	1037,45	19,68	8676,56	2979,07	47326,6	1769,88	61084,2	420,44	72355,1	-
Coeficiente Variação	10,79%	8,04%	21,93	14,38%	31,94%	15,76%	9,60%	12,23%	32,38%	7,86%	-
	27,21%	12,45%	15,86%	8,89%	25,04%	9,87%	16,12%	8,76%	35,96%	7,65%	-

Fonte: Seara (2000, p. 173)

³⁵ Os Quadros 3.8 e 3.9 apresentam dados somente referentes ao contexto tônico das sílabas [na] e [ni], uma vez que as palavras componentes do instrumento de coleta não apresentavam opções que incluíssem as sílabas em contexto pós-tônico.

Os Quadros de 2.4 a 2.9 apresentam os valores para f_0 , formantes e larguras de banda, da primeira à quinta faixa de frequência, dispostos, em cada quadro, de cima para baixo, com a média em primeiro, seguido do desvio padrão e da variância³⁶, com essas três medidas expressas em Hz, terminando com o coeficiente de variação. Para aqueles quadros em que constam os contextos tônico e átono, o primeiro fica acima do segundo na linha correspondente à medida e, para os quadros referentes a [na] e [ni], apenas os resultados para contexto tônico foram obtidos. Esse fator não apresenta relevância para a pesquisa em tela, pois apenas os valores de f_0 , dos três primeiros formantes e suas respectivas larguras de banda, medidos neste trabalho, serão comparados com os resultados obtidos por Seara (2000).

A intenção de Seara (2000) com o experimento era que o levantamento efetuado indicasse parâmetros acústicos mais consistentes, o que foi avaliado segundo o coeficiente de variação, para cada uma das duas consoantes nasais analisadas com o objetivo de caracterizá-las acusticamente e de permitir seu reconhecimento através desses parâmetros. A autora descreve os valores encontrados em cada caso, porém os achados que são relevantes para a pesquisa em tela dizem respeito ao coeficiente de variação que, para os melhores casos em Seara (2000), ficou em torno dos 10% e, de forma geral, ficou na faixa dos 10 a 20%, valores que serão tomados como referência no Capítulo 5. Finalizando sua análise das consoantes nasais, a autora conclui, para as consoantes nasais, que a variabilidade interindividual resultante das diferenças das conformações de cavidades nasais dos indivíduos prejudica o estabelecimento de invariantes acústicas que determinem as produções sonoras. Conforme será abordado na seção seguinte, esse achado é importante para o raciocínio que embasa a CL.

Esta seção apresentou os elementos acústicos e articulatórios das vogais orais e das consoantes nasais, a partir de estudos para o PB. Os conceitos e representações abordados são fundamentais para o conjunto de procedimentos que compõe o exame de CL. A seção seguinte tem por objetivo esclarecer a maneira como esses conceitos e representações podem servir aos peritos que se dedicam à identificação através da comunicação verbal oral.

³⁶ Desvio padrão e variância são medidas redundantes, uma vez que essa é o quadrado daquele. Uma vez que os quadros foram retirados integralmente de Seara (2000), decidiu-se por manter a apresentação original.

2.5 APLICAÇÃO FORENSE

2.5.1 Introdução

As análises focadas nos fenômenos linguísticos perceptíveis ao exame realizado através da oitiva, ou seja, aquelas que não necessitam obrigatoriamente do nível de detalhamento de parâmetros acústicos como os aqui levantados, têm sido cada vez mais decisivas no que diz respeito à consistência das conclusões alcançadas nos casos forenses de CL. De qualquer maneira, a análise de parâmetros acústicos segmentais ainda é fundamental para a CL, seja como acessória à análise dos fenômenos linguísticos possíveis de serem levantados através da oitiva direta, seja como ferramenta principal de embasamento das conclusões determinadas nos laudos periciais de CL (GOLD E FRENCH, 2011; MORRISON, 2014).

Vários são os parâmetros possíveis de serem utilizados para compor a CL, que serão avaliados conforme a quantidade e qualidade do material acústico disponível para o exame. Para o PB, Figueiredo (1994) realizou diversos ensaios para avaliar o desempenho dos parâmetros formantes vocálicos, frequência fundamental, espectro de longo termo, consoantes nasais, tempo de vozeamento e aspectos rítmicos-temporais, dentre eles velocidade de fala. French et al. (2010, pp. 146-147) lista os diferentes características da fala que podem compor a CL, referindo a qualidade vocal, a entonação, o *pitch* – obtido através da média e da variação da f_0 , a taxa de articulação³⁷, as características rítmicas, processos de assimilação e elisão relacionados ao encadeamento de fala, as características consonantais (p. ex. plosões, durações de nasais, líquidas e fricativas em contextos fonológicos específicos, tempo de vozeamento), as características vocálicas, as informações linguísticas de níveis que extrapolem o nível fonético-fonológico (p. ex. marcadores discursivos, escolhas lexicais, variantes morfossintáticas, entre outras), as patologias e características não linguísticas manifestadas no áudio como elementos a serem considerados para o exame. Com relação às características das vogais, o autor menciona a configuração dos formantes e larguras de bandas. As análises das vogais podem ir além da avaliação de suas

³⁷ A respeito das diferenças entre as definições de taxa de articulação e taxa de elocução, recomendo o quadro explicativo de Gonçalves (2013, p.61).

configurações de formantes (FRENCH ET AL., 2010), apesar dessa análise ser a que tradicionalmente se utiliza uma vez que apresenta um desempenho consistente (FIGUEIREDO, 1994). Quanto às consoantes nasais, French et al. (2010) refere-as apenas quanto sua duração e Figueiredo (1994) conclui que seu desempenho discriminante é inferior aos demais testados. Nesta seção, após uma breve introdução sobre a CL, apresentaremos considerações sobre a utilização das vogais e das consoantes nasais com objetivo forense.

A respeito do procedimento envolvido para a consecução do exame de CL, diversos aspectos devem ser observados em cada etapa. Recebido o registro de áudio em que consta a fala que está sob questionamento, que na prática forense recebe a denominação de **material questionado**, e realizada cópia digital para preservar o original, procede-se à triagem do conteúdo para retirar todos os elementos que prejudiquem a qualidade acústica do objeto de análise. Ruídos ambientais, falas sobrepostas, falhas de captura, sejam do sistema de captura ou do posicionamento em relação ao dispositivo de captura, são removidos, mantendo-se apenas o que se denomina “fala líquida”, que é a porção de fala constante no registro original que apresentava qualidade acústica para a CL. Esse material servirá de referência para a etapa seguinte em que o perito procederá à gravação do denominado **material padrão**³⁸. A coleta do material padrão é feita em duas situações, uma coleta de fala repetida, em que o indivíduo repetirá trechos conforme o transcrito da fala líquida, com modelo oferecido oralmente e não por escrito, que é utilizado para comparações de trechos foneticamente análogos, e outra em que se procede à coleta de fala espontânea, em que é gravada uma situação de entrevista em que se pretende obter a fala mais espontânea possível, considerando, por um lado, as orientações para a entrevista sociolinguística, em que o vernáculo é mais proximamente obtido quanto menor atenção for dispendida

³⁸ Os grifos ressaltando a nomenclatura “questionado” e “padrão” foram aplicados, pois essa relação, questionado/padrão, é uma das bases da criminalística, no que diz respeito a todos os exames comparativos, desde a mais corriqueira marca em local, a impressão digital, até os exames que requerem equipamentos sofisticados, como o DNA, ou conhecimento multidisciplinar profundo, como é o caso da CL. Os termos são muito comuns no âmbito forense e serão recorrentes neste relatório, nesta seção e na discussão, principalmente. Como a solicitação do exame pericial usualmente vem na forma de questão, ou quesito, consagrou-se o termo “questionado”. O segundo termo, “padrão”, tem o sentido de referência, pois é aquele material que o perito irá coletar controladamente para garantir sua aplicabilidade ou, eventualmente, quando não for o caso da coleta de padrão, sobre o qual tem certeza de origem.

no monitoramento da própria fala (LABOV, 1972, p. 208), e, por outro lado, levando em conta o contexto criminal, onde a situação de coleta é naturalmente constrangedora para o sujeito da pesquisa, seja vítima ou suspeito. A coleta de fala espontânea busca uma aproximação do participante ao seu vernáculo através de estratégias de envolvimento do indivíduo em sua narrativa e em técnicas de entrevista. Acima de tudo passa pela percepção do entrevistador em conduzir um assunto em direções que façam o participante retirar sua atenção e controle sobre sua fala. A entrevista como um evento de fala tem um caráter especial, de forma que a naturalidade e a informalidade devem estar em questão (FEAGIN, 2004).

A respeito dos estilos de fala com os quais a CL lida, adota-se a ideia de Warner (2012) e Schilling-Estes (2008) sobre o contínuo de tipos de fala, em que se pode imaginar, numa extremidade do contínuo, a leitura de monossílabos sem sentido e, na outra extremidade, uma conversa informal (WARNER, 2012, p. 622). Assim, é possível referenciar, em relação a tal contínuo, além das três situações da CL – frases repetidas, fala espontânea e gravação desavisada – a coleta realizada pela pesquisa em tela, com monossílabos lidos em frase-veículo. Dessa maneira, numa extremidade do referido contínuo fica a coleta de fala como a que será realizada pela pesquisa em tela, e na outra, o material questionado, já que é uma gravação sem a ciência por parte de quem está sendo gravado, e, entre as duas extremidades, as estratégias utilizadas para a coleta de material padrão de fala, com a coleta de fala repetida mais próxima da primeira extremidade do contínuo, e a coleta de fala espontânea, mais próxima da segunda extremidade. Essa posição dos diferentes materiais de fala utilizados na CL deve ser considerada durante o exame. Com relação à leitura de monossílabos, estratégia utilizada no presente estudo, é a abordagem escolhida uma vez que se deseja obter um resultado de variabilidade mínima para os parâmetros acústicos.

2.5.2 As vogais como parâmetros de identificação

As vogais não carregam tanto conteúdo informativo da fala quanto as consoantes e, dessa forma, a variação vocálica seria mais tolerada pelos falantes. Adicionalmente, as vogais apresentam mais características espectrais discriminativas, o que as tornaria mais atrativas do ponto de vista da identificação de falante (BEIGI, 2011, p. 128). Dessa

maneira, os sons vocálicos e seus parâmetros acústico-articulatórios carregariam aspectos individuais valiosos para a atividade forense.

Os parâmetros normalmente medidos são os formantes e suas larguras de banda (FRENCH ET AL., 2010, p. 147). Relacionados diretamente com o formato do trato vocal, os dois primeiros formantes das vogais orais são parâmetros consagrados na CL uma vez que são representativos das dimensões vertical (primeiro formante) e horizontal (segundo formante) da cavidade oral, dimensões que variam principalmente como consequência da posição da língua nessa cavidade. O posicionamento da língua ocasionará constrictões próximas aos nodos ou antinodos, conforme apresentado na seção 2.3, Capítulo 2, determinando uma conformação do trato vocal que servirá como elemento de ressonância, atenuando ou reforçando frequências formânticas específicas. Para cada vogal haverá um formato específico do trato vocal que pode ser inferido pela conformação dos formantes no espectrograma.

Estudos dedicados a avaliar a importância dos formantes vocálicos para a CL não são profusos. Rose (2007) estuda o poder discriminante das vogais do inglês australiano em um experimento de larga escala, que envolve 171 falantes masculinos que pronunciaram duas ocorrências, uma em isolamento e outra em final de frase com acento nessa posição, de palavras com forma “hVd”, com vogal na posição de V. Foram pesquisados os onze monotongos vogais da língua, cinco tensos (/i: ʉ: ə: a: o:/) e seis frouxos (/ɪ e æ a o ʊ/), bem como os seis ditongos (/aɪ ɛɪ ɔɪ ɐɪ ɪə ɛə/). Foram considerados para o estudo os três primeiros formantes, porém nem todas as vogais tiveram os três formantes medidos, como o caso de /ʊ/, que não apresentava o F3 com condições de mensuração. Como resultado principal do estudo, a vogal /i:/ apresentou o melhor poder discriminante em termos de taxa de erro igual³⁹.

Com o objetivo de efetuar a CL em um caso real, Nolan e Grigoras (2005) descrevem as análises realizadas em quatro ligações telefônicas atribuídas a um suspeito falante de inglês australiano. Além de procederem a uma descrição das características linguísticas dialetais das duas amostras, os autores realizaram a análise acústica baseados em gráficos de dispersão de F1x F2 para a vogal /ɪ/ e gráficos de

³⁹ Medida que confronta o número de falsos positivos e falsos negativos de um teste, onde a equivalência desses erros é o referencial.

dispersão para ditongos, em que os eixos são formados pelas medidas de F2 no início do ditongo para o eixo horizontal e de F2 no fim do ditongo para o eixo vertical. As áreas de ocupação bem determinadas e separadas em todos os gráficos de dispersão permitiram que os autores concluíssem pela não semelhança das amostras de fala, isentando o suspeito da acusação.

O trabalho de Bosch (2003) é aquele que investiga, no Catalão, os mesmos parâmetros acústicos a que se propõe a pesquisa em tela, frequência fundamental e os três primeiros formantes vocálicos. As vogais analisadas são as tônicas [i, e, ε, a, o, u] e as átonas [ɪ, ʊ, ə] e foram obtidas a partir de leitura realizada por cinco indivíduos masculinos e cinco femininos. As variáveis falante, lugar de articulação e modo de articulação foram analisadas através de análise de variância multifatorial (MANOVA) em sua interação com as variáveis dependentes f0, F1, F2 e F3.

O autor conclui que f0 e F3 apresentam as interações mais significativas com a variável falante, que indica esses parâmetros acústicos como os mais relevantes para a CL. A respeito da metodologia adotada pelo autor para a análise acústica, em que foram utilizados os recursos de Texgrid e de *script* do programa computacional que serviu à pesquisa⁴⁰, é importante ressaltar os parâmetros definidos para o cálculo do LPC. Foram definidos de forma fixa os parâmetros “número de formantes” em cinco e “valor máximo de formante” em 5000Hz para os falantes masculinos, o que permite que o processo de leitura dos valores formânticos seja realizado automaticamente, diferentemente do procedimento realizado na pesquisa em tela, como será descrito no Capítulo 4, em que esses parâmetros foram configurados individualmente para cada som medido, permitindo o melhor acoplamento do LPC ao espectrograma, conforme indicado por Thomas (2011, p. 41).

Os valores dos formantes são, entre vários dos citados anteriormente, elementos de cotejo utilizados profusamente na CL. Para que o cotejo seja representativo, os elementos que inserem variação nesses valores devem ser controlados (NOLAN 1983, 1997; HOLLIEN, 2002; ROSE, 2002). Essa noção baseia a coleta de padrão de fala

⁴⁰ A respeito do programa de análise acústica e de suas ferramentas de TextGrid e *scripting* são apresentadas as explicações no Capítulo 4.

repetida para a realização de comparações acústicas, bem como fundamenta a configuração voltada para a variabilidade individual, que é o tema da pesquisa em tela.

A quantificação dos formantes é realizada de forma isolada para cada som, sem que haja uma estratégia associativa de seus valores, além da representação no gráfico de F1x F2, na tentativa de se obter a representação de características particularizantes. A Razão de Centralização Formântica (FCR, do inglês) e a área de espaço vocálico (VSA, do inglês) oferecem elementos que permitem interpretar os valores formânticos integradamente em um valor. Sabe-se que o FCR sofre menor influência da variabilidade individual do que o VSA, conforme apresentado na seção 2.3.1, neste capítulo (SAPIR ET AL., 2010; PETTINATO ET AL., 2016), o que torna a última medida, bem como seus componentes, as distâncias euclidianas, teoricamente candidatas a parâmetro de análise com chances de colaborar com a perícia de CL.

Com relação ao ponto de medição na vogal, normalmente adota-se a estratégia proposta por Labov et al. (1972) e consagrada nas pesquisas linguísticas de obter-se o valor do formante no ponto em que o F1 da vogal atinge seu máximo, uma vez que seria aí que a vogal sofreria a menor influência coarticulatória dos contextos circundantes (THOMAS, 2011, p. 150). A estratégia de medição na zona de estabilidade da vogal pode ser questionada, uma vez que a intenção da pesquisa linguística é obter um valor representativo da língua, enquanto que, para o pesquisador forense, o objetivo não é encontrar um valor comum e sim um valor representativo do indivíduo, com suas idiossincrasias. Para a tarefa de identificação, uma estratégia como tomar os ditongos proferidos pelo suspeito e pelo autor, e compará-los mediante gráficos de dispersão de F2 no início do ditongo por F2 no fim do ditongo (NOLAN e GRIGORAS, 2005) permite a percepção da transição do formante para cada indivíduo do cotejo, oferecendo critério consistente para a determinação da semelhança das falas. Os aspectos dinâmicos da fala, principalmente as transições formânticas, são elementos que naturalmente carregam as particularidades dos falantes (NOLAN ET AL., 2006; McDOUGALL, 2005).

Para o presente estudo adotar-se-á a estratégia de trabalhar com três pontos de medida em torno do ponto médio da vogal, com o objetivo de tomar valores representativos do trecho, conforme propõe Thomas (2011), e não apenas um instante, o que será descrito na seção 4.3, Capítulo 4.

2.5.3 As nasais como parâmetros de identificação

A relativa rigidez da cavidade nasal assegura uma baixa variabilidade intraindividual em relação às características acústicas de ressonância nasal ao mesmo tempo em que sua estrutura sinuosa e dimensão interna apresentam uma complexidade suficiente para conferir alta variabilidade interindividual (ROSE, 2002, p. 135). Seara (2000) afirma que a variabilidade interindividual resultante das diferenças das conformações de cavidades nasais dos indivíduos prejudica o estabelecimento de invariantes acústicas que determinem as produções sonoras, o que leva a pensar que essas diferenças possam ser individualizantes. Apesar desse aparente contexto favorável, a prática forense, pelo menos no que diz respeito à realidade no Brasil, não apresenta indicativos de utilização dos formantes das consoantes nasais como elementos de comparação.

Os trabalhos encontrados a respeito da avaliação da nasalidade consonantal como parâmetro para o procedimento de reconhecimento aplicado ao campo forense, são voltados para a verificação automática de falantes, que não é o foco do trabalho aqui proposto, porém podem oferecer alguma luz sobre a questão. Glenn e Kleiner (1968), Wolf (1972), Höfker (1977) e Su et al. (1975) pesquisaram, entre outros parâmetros, espectros contínuos de frequências das consoantes nasais, com limites inferiores e superiores diferentes em cada um desses estudos, mas em todos o espectro avaliado permaneceu entre 150 e 7000Hz. Esses estudos também têm em comum a estratégia de representação do espectro contínuo, através da sua estratificação em faixas de frequências menores, também, conforme cada estudo, que são representados por vetores matemáticos que contêm, para cada estrato do contínuo da frequência, um valor de intensidade. Assim, não são os formantes nasais especificamente que são avaliados como parâmetros identificadores, mas os vetores definidos quando comparados com bases de dados pré-estabelecidas. Essa é uma estratégia que melhor se presta a um sistema com paradigma de reconhecimento automático, que tem a capacidade de analisar um volume de dados muito grande quando comparado ao procedimento manual, ou “artesanal”, de CL. O que diferencia cada um dos trabalhos a seguir descritos são,

fundamentalmente, as características de composição dos vetores, bem como os delineamentos relativos às coletas e metodologias de análise.

Glenn e Kleiner (1968) utilizam amostras da consoante [n] obtidas em ambientes fonéticos diferentes. Seu estudo inclui 30 falantes, 20 homens e 10 mulheres, do inglês americano, que leram duas listas de dez palavras com a consoante nasal em contextos fonéticos e posições diferentes em relação à sílaba (p. ex. “mountain”, “nominal” e “moment”, para a lista de referência e “nine”, “mention” e “munition” para a lista teste). De cada ocorrência foi obtido um espectro estratificado em 25 faixas de 100Hz, com 64 níveis de amplitude para cada faixa, contidas na faixa de 1000-3500Hz. Essa estratificação era representada por um vetor de dimensões 25 (faixas) x 64 (níveis de amplitude). Vetores de referência foram estabelecidos a partir da leitura da primeira lista de palavras e vetores de teste a partir da leitura da segunda lista, com 300 ocorrências em cada conjunto. Para o conjunto de referência, foi estabelecida a média das dez ocorrências de cada falante. Após, compararam-se as 300 ocorrências do conjunto de vetores de teste com o conjunto de referência, obtendo-se um índice de acerto de 43%. Em seguida, compararam-se com o conjunto de referência 150 valores, obtidos a partir da combinação de cinco médias de duas ocorrências dos 10 valores de cada falante, com o índice de acerto de 62%. O mesmo foi feito com a média de 5 valores, resultando em 60 itens comparados com o conjunto de referência, resultando em 82% de acerto e, finalmente, a média dos 10 valores de cada falante do conjunto de teste foram comparadas com o conjunto de referência, resultando no índice de acerto de 93%.

Em estudo que avaliou os espectros vocálicos, os espectros consonantais nasais e a frequência fundamental em dados extraídos de sentenças curtas lidas por 21 indivíduos masculinos falantes de inglês americano, Wolf (1972) refere o melhor desempenho dos parâmetros nasais em relação aos demais, sendo superados apenas pela frequência fundamental. Para a testagem das nasais, o autor utilizou-se de estratégia baseada no estudo de Glenn e Kleiner (1968), porém estratificou os espectros de [m], entre 150 e 4900 Hz, em 30 faixas, e de [n], entre 150 e 3400 Hz, em 25 faixas.

O comportamento das consoantes nasais deve ser considerado em relação à vogal que o segue, já que, entre [m] e [n], o som bilabial apresenta maior tendência à coarticulação, uma vez que a língua permanece livre durante a articulação dessa nasal,

ao contrário da alveolar, em que a língua está envolvida na articulação desse som. A dinâmica coarticulatória pode servir melhor para a identificação do que a pesquisa da nasal isolada (SU et al., 1974, p. 1881).

Höfker (1977), em estudo com 12 falantes, registrou 24 sons do alemão pronunciados de forma isolada, com dez repetições, em uma câmara anecoica e obteve, como resultado do ordenamento desses 24 sons, as três nasais ([m], [n] e [ŋ]) nas primeiras posições, mostrando bom desempenho desses sons, que superaram inclusive as vogais.

Os trabalhos apresentados referem-se a delineamentos de pesquisa voltados para o reconhecimento automático de falantes, procedimento que difere da atividade normalmente utilizada na CL, que é manual, ou melhor, artesanal, no sentido de que os cotejos são efetuados de maneira minuciosa e sem auxílio de ferramentas automáticas. Porém, os achados apresentados apontam para a importância das consoantes nasais para o processo de reconhecimento de falante, seja por verificação ou por identificação.

Quando se considera o reconhecimento realizado através da percepção, mais uma vez percebe-se que as consoantes nasais são parâmetros consistentes. Amino et al. (2012) propuseram pesquisa, para o japonês, em que dez indivíduos produziram cinco repetições de logatomas formados por sequência /VCVCVCV/ inseridas em frase-veículo, em que V eram posições ocupadas pela vogal [a] e C eram ocupadas por uma das oito consoantes coronais ou a nasal bilabial. O ponto coronal foi escolhido por ser o que mais apresenta itens no inventário do japonês. Foi recortada apenas uma sílaba da sequência para cada uma das cinco repetições e apresentada isoladamente para os sujeitos que deviam reconhecer os locutores após uma fase de treinamento, através da oitiva de frases-modelo que serviam como referência para a comparação. Foram realizados dois experimentos com esse material, um onde o ouvinte conhecia o falante e outro, com grupo menor de ouvintes, em que não havia familiaridade com os falantes. Conforme o esperado, o índice de acertos, ou de reconhecimentos positivos, foi maior no primeiro caso, de indivíduos conhecidos, do que no segundo, de indivíduos desconhecidos para o avaliador. Nos dois experimentos, a nasal foi o elemento consonantal, entre as oito coronais, que apresentou mais destaque (AMINO et al., 2012).

Figueiredo (1994) refere o possível valor de se utilizar os sons consonantais nasais como recurso de identificação ao citar Nolan (1983), que resumiu os resultados dos estudos de Glenn e Kleiner (1968), Wolf (1972) e Höfker (1977), e afirma que: "These results indicate that quite accurate speaker identification can be achieved on the basis of spectral information taken from individual segments of an utterance, in this case nasals"⁴¹ (NOLAN, 1983, p. 76). Figueiredo (1994) testou diversos parâmetros para a identificação de falantes e descreveu, entre vários experimentos, um realizado com nove indivíduos que leram uma lista de quinze palavras, que não tiveram controle de contexto vocálico no entorno da ocorrência da nasal [n], única testada. Figueiredo (1994) obteve desempenho da nasal como parâmetro para CL abaixo dos demais parâmetros e conclui que o *corpus* limitado para esse teste, tanto no que se refere ao número de indivíduos quanto ao número de ocorrências produzidas, indica um campo fértil para futuros estudos.

Como se vê, a rigidez da cavidade nasal não é garantia de invariabilidade acústica. Apesar do acoplamento da cavidade nasal à cavidade oral resultar em um comportamento formântico menos variável intraindividualmente, oferece um resultado acústico com menor energia. O único fator que pode fazer variar a energia nasal, é a tensão e conseqüente grau de abertura do véu palatino (LIBERMAN e BLUMSTEIN, 1988). Estratégias articulatórias diferentes são utilizadas por indivíduos diferentes para alcançar os mesmos objetivos acústicos. É possível que um indivíduo com cavidade nasal menor realize movimento velofaríngeo mais amplo e compensação do volume oral para obter a mesma nasalidade que outro indivíduo com cavidade nasal maior (ENGWALL 2003, p. 9).

Tendo em vista a proposta do estudo em tela de avaliar o comportamento dos formantes vocálicos, que será feito pelo cálculo do LPC, e considerando a estratégia proposta por Seara (2000) de medição dos formantes consonantais nasais através da extrapolação de coeficientes do LPC, ou seja, utilizando do mesmo recurso de medição dos formantes vocálicos, a mesma ferramenta de medição será utilizada, na pesquisa em tela, para as duas situações, o que facilita a coleta de uma quantidade de dados como a

⁴¹ Esses resultados indicam que a identificação de falante acurada pode ser alcançada com base na informação espectral obtida a partir de sons individuais de uma produção, nesse caso as nasais.

que se propõe o trabalho, conforme será descrito no Capítulo 4. A quantidade de dados a serem analisados está intimamente relacionada à capacidade de inferência sobre a variabilidade individual, e essas noções, quantidade de dados e variabilidade individual, são fundamentais à CL, como será exposto na seção que segue.

2.5.4 Relação entre a variabilidade inter e intraindividual

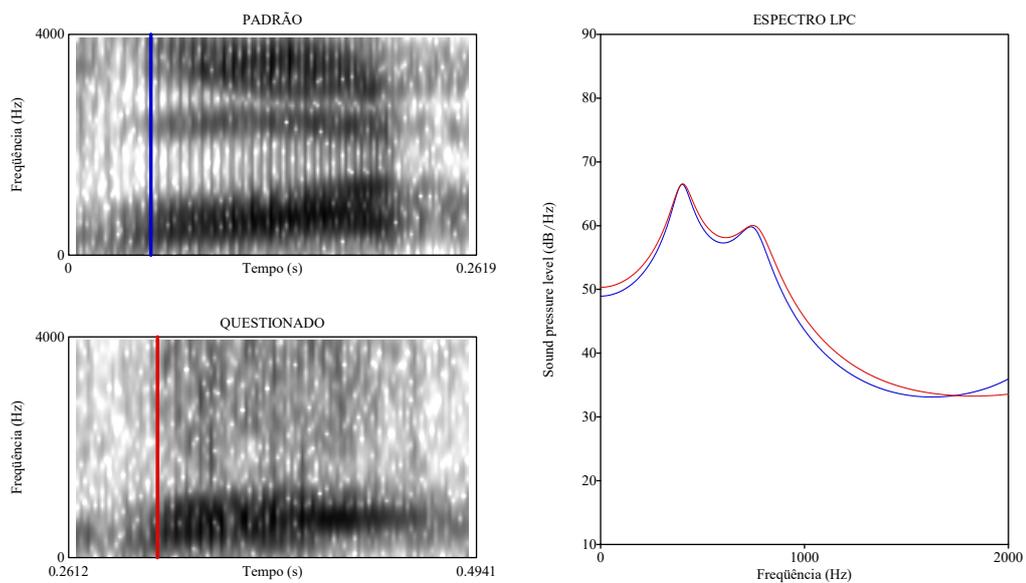
Devido à capacidade de realizarmos, corriqueiramente e sem auxílio de recursos tecnológicos, o reconhecimento leigo de locutores, a intuição pode levar-nos a imaginar que um indivíduo possui uma voz única, como se fosse um elemento próprio característico, de maneira análoga à impressão digital. Essa é uma simplificação, primeiramente porque a voz de uma pessoa não é constante como o é a impressão digital e, além disso, porque não se sabe ao certo em que medida a variação na fala de um indivíduo causa uma sobreposição de aspectos de sua voz em relação aos outros indivíduos de sua comunidade de fala (NOLAN, 2002, p. 12).

A analogia com a impressão digital foi inaugurada por Kersta (1962) e se perpetua em algumas comunidades em função da situação em que surgiu na época. Com o avanço da tecnologia após investimento com objetivos militares, pois se pretendia reconhecer as vozes nas mensagens interceptadas por rádio, em meados da década de 40 surgiu o sonógrafo, aparelho que permitia “enxergar” os sons (Hollien, 2002). Kersta (1962) foi quem pleiteou a força do espectrograma como um exame páreo ao da impressão digital, batizando o exame com nome alusivo ao exame da impressão digital (*voiceprint* em analogia a *fingerprint*). A questão que se coloca é justamente entender o motivo pelo qual a impressão digital tem maior consistência do que a amostra de fala e voz. A resposta para tal questão fundamenta-se na relação entre variabilidade intraindividual e variabilidade interindividual. No caso da digital, a primeira variabilidade é muito pequena, enquanto a variabilidade interindividual é muito grande. No caso da fala e da voz, a variabilidade intraindividual não é pequena, em função da plasticidade que é característica dos órgãos fonatórios (NOLAN, 2002). Associada a isso, a variabilidade interindividual, apesar de existir em função das diferentes constituições orgânicas e das diferentes experiências linguísticas e cognitivas, é minimizada por mecanismos intuitivos de normalização sempre que escutamos a fala, uma vez que o objetivo da comunicação

é normalizar a produção para permitir e facilitar a percepção, ou seja, a comunicação torna comum o sinal de fala.

A língua é o órgão que confere boa parte da variabilidade à fala, pois apresenta grande mobilidade, característica que exige uma série de cuidados durante o processo de identificação, como, por exemplo, o controle dos contextos linguísticos em que ocorre o som analisado, como já foi referido na introdução da seção 2.5, Capítulo 2. A exemplo desse controle de contexto, a Figura 2.27 a seguir apresenta um elemento de confronto de vestígios acústicos utilizado em um caso real de CL (Laudo Pericial 1590/2013, expedido pelo Departamento de Criminalística do Rio Grande do Sul). São apresentados dois espectrogramas, ambos referentes à primeira sílaba da palavra “voto”, proferida no material padrão e no material questionado. À esquerda da imagem, o espectrograma superior refere-se ao material padrão e o inferior, ao material questionado. À direita, apresenta-se o gráfico LPC para as duas amostras, referente aos pontos indicados pelos cursores coloridos nos espectrogramas, o azul para o material padrão e o vermelho para o questionado. Tendo em vista o controle dos contextos conforme recomenda Nolan (1983), os pontos dos espectrogramas a partir dos quais se calcula o LPC devem ser, como de fato são, na Figura 2.26, correspondentes.

Figura 2.27 – Exemplo de cotejo utilizado em laudo de CL.



Fonte: O Autor (2013), em Laudo Pericial de número 1590/2013, expedido pelo Departamento de Criminalística do Rio Grande do Sul.

Esse confronto das falas é, porém, quantitativamente tênue, apesar de ser um indício de convergência entre as amostras. Apenas com um conjunto grande de vestígios é que se podem fazer afirmações a respeito da semelhança das amostras.

Conforme Nolan (1983, p. 26), Rose (2002, p. 11) e Hollien (2002, p. 7), o princípio fundamental para a particularização de um indivíduo é que a variação intraindividual do parâmetro analisado seja menor do que a variação interindividual, conforme foi exemplificado através da analogia com impressão digital, em que a variação intraindividual é pequena, já que o mesmo dedo produzirá impressões muito semelhantes, e a variação interindividual é grande, uma vez que o dedo equivalente de pessoas diferentes sempre produzirá impressões diferentes.

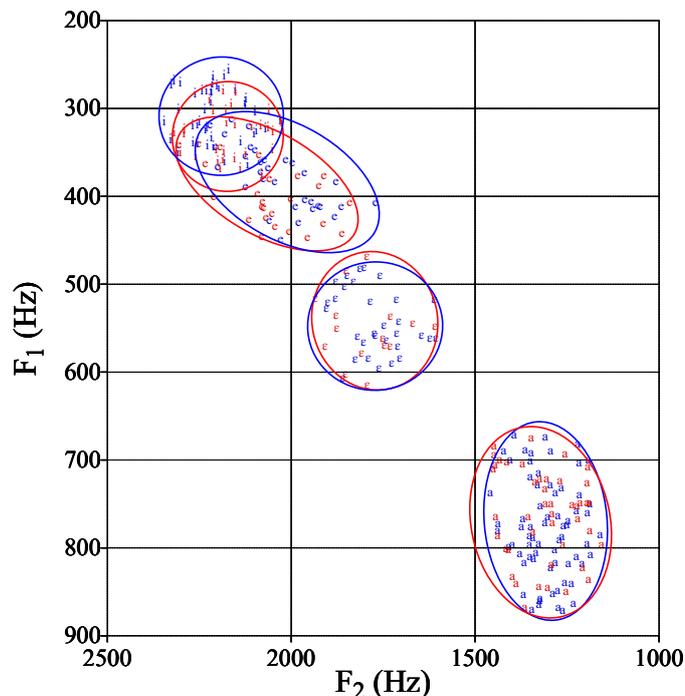
A definição do conceito de especificidade de falante como sendo a razão entre a variabilidade interindividual e a variabilidade intraindividual, que pode ser utilizada tanto para o processo automático quanto para o processo manual é fundamental para o exame de CL, como mostrou Heuvel (1996). O autor fez um estudo de diferentes parâmetros de vogais e consoantes da língua holandesa, realizando experimentos que avaliaram o espectro e a duração desses sons. Sua pesquisa, porém, dedicou-se a achados destinados ao reconhecimento automático de locutor. Os achados acústicos espectrais serviram para que o autor estabelecesse um *ranking* do conjunto conforme o desempenho para a identificação de falantes, determinando quais produções sonoras são mais falante-específicas (em inglês, *speaker-specific*). O resultado colocou três vogais, “a”, “i” e “u”, e três consoantes, “n”, “m” e “s”, em ordem, de maneira que, partindo do mais falante-específico e decrescendo nesse valor até o menos falante-específico, formou-se a sequência “animus”. Já os achados de duração serviram para a discussão do fenômeno de coarticulação. O autor conclui que os sons mais recorrentes na língua têm a tendência a ser mais falante-específicos, o que ele explica por um possível menor monitoramento sobre as produções mais frequentemente realizadas.

O processo de identificação de falantes pode ser entendido se imaginarmos um espaço de fala multidimensional (Nolan 1991), onde as dimensões são estabelecidas por parâmetros fonéticos variáveis. Como exemplo, tomemos o gráfico de dispersão de vogais, onde o primeiro e o segundo formantes formam os eixos ordenados, como um exemplo de um espaço bidimensional da fala tradicionalmente utilizado nos estudos

linguísticos. Conforme McDougall (2005), os falantes ocupam regiões desse espaço que por vezes se sobrepõem e, se está correta a hipótese de que cada indivíduo possui características ímpares, pode-se estabelecer uma região única para cada indivíduo definindo-se um número suficiente de parâmetros fonéticos variáveis.

Utilizando-se do mesmo caso real de CL recém referido (Laudo Pericial 1590/2013, expedido pelo Departamento de Criminalística do Rio Grande do Sul), apresenta-se, na Figura 2.28 a seguir, um gráfico de dispersão de vogais anteriores mais a baixa⁴² do falante cuja fala foi investigada. Em azul as ocorrências de vogal do material padrão e em vermelho as ocorrências das vogais do material questionado.

Figura 2.28 – Exemplo de cotejo utilizado em laudo de CL.



	[i]			[e]			[ɛ]			[a]		
	n	F1	F2									
Padrão	42	326	2146	32	393	1995	31	542	1770	68	757	1365
Questionado	29	348	2108	25	413	2014	15	549	1789	44	742	1373
Diferença (%)		6,75	1,77		5,09	0,95		1,29	1,07		1,98	0,59

Fonte: O Autor (2013), em Laudo Pericial de número 1590/2013, expedido pelo Departamento de Criminalística do Rio Grande do Sul.

⁴² As vogais posteriores foram excluídas da análise pois a qualidade acústica do registro dificultou extrações precisas dessa série.

A proximidade das áreas para as quatro vogais é considerada também como vestígio e convergência das duas amostras, assim como o exemplo do cotejo anterior, apresentado na Figura 2.27. Este, porém, é considerado um vestígio com maior peso do que aquele, em função de carregar em si a variabilidade inerente à fala e, assim, é mais representativo do indivíduo. Com o intuito de dimensionar a representatividade do gráfico, apresenta-se no quadro, na porção inferior da Figura 2.28, os valores médios de F1 e de F2 para cada vogal, as quantidades de sons vocálicos que compõem o gráfico e a diferença percentual⁴³ entre o valor dos formantes do material padrão e do material questionado.

Rose (2002, p. 12) afirma que, teoricamente, é possível realizar o processo de reconhecimento adotando-se quantas dimensões se entender necessárias, o que o autor chama de comparação multidimensional, tomando por base as noções euclidianas de espaço. Dessa forma, seria possível traçar gráficos de dispersão em que os eixos indiquem medidas que não necessariamente F1 e F2, mas quaisquer parâmetros acústicos de natureza contínua. As áreas ocupadas pelos elementos medidos podem ser comparadas e, assim, definidas as diferenças ou semelhanças entre dois conjuntos. Caso as áreas sejam resumidas em pontos no plano cartesiano, pelo estabelecimento de valores médios das ocorrências, torna-se possível a comparação entre tais pontos através da distância euclidiana, por exemplo.

Champod e Evett (2000) sustentam a ideia de que, no âmbito forense, não basta uma hipótese analisada ser determinada como possível, mas além de possível deve ser confrontada com, no mínimo, outra hipótese concorrente. No caso da CL, são exemplos as hipóteses de que as amostras de fala têm mesma origem e de que não têm mesma origem. Nesse confronto de hipóteses reside o ponto que os autores consideram como fundamental para a atuação do perito, uma vez que deve ser capaz de estimar o peso, a chance, o valor de uma hipótese em relação a sua alternativa. A representação formal dessa relação de hipóteses é contemplada pela razão de verossimilhança, base para a inferência bayesiana, que permite calcular a probabilidade *a posteriori* dadas as hipóteses e a evidência baseada em dados numéricos. Nesse sentido, Rose (2002) e

⁴³ Diferença percentual calculada pela fórmula $|((P-Q)/P)*100|$, em que Q se refere ao parâmetro do material questionado e P ao parâmetro do material padrão.

Aitken e Gold (2013) sugerem a abordagem bayesiana como a mais adequada para a avaliação de evidências no campo da Fonética Forense, em que as noções de similaridade, que se refere ao quanto são semelhantes os elementos comparados, e tipicidade, que diz respeito ao peso que possui a evidência comparada, entram em questão. A similaridade diz respeito ao quão semelhante uma evidência é da outra. Porém essa informação não é útil sozinha. Deve-se ter uma noção também da tipicidade da evidência, do peso que essa evidência assume para a comprovação da hipótese de que o padrão e o questionado são de mesma origem. No âmbito dos fenômenos linguísticos, por exemplo, caso duas amostras apresentem apagamento da fricativa alveolar final indicadora de plural, apesar de haver similaridade entre as amostras, a tipicidade é fraca, uma vez que a incidência do fenômeno é grande na língua. Por outro lado, se as amostras apresentam uma realização da líquida lateral em final de sílaba como velarizada para uma comunidade de fala em que não é esperada, a tipicidade já pode ser considerada como mais forte para essa evidência. Similaridade e tipicidade são conceitos que, segundo Eriksson (2012), podem ser entendidos como paralelos a consistência e distintividade. Consistência é a decisão a respeito do quanto o material padrão e o questionado são compatíveis, se foram produzidos pelo mesmo indivíduo. Distintividade é uma avaliação da medida em que as mesmas características são compartilhadas por outros membros da comunidade (ERIKSSON, 2012, p. 62). Esses conceitos serão representados no estudo aqui proposto através das noções de variabilidade intraindividual, relacionada à similaridade, e variabilidade interindividual, relacionada à tipicidade, que terão cálculos definidos apropriadamente na seção 4.5, Capítulo 4 e serão discutidos na seção 5.3.1, Capítulo 5, com base nos resultados da presente pesquisa, na seção 5.1, Capítulo 5.

As bases fonéticas, articulatórias e acústicas, foram apresentadas para entender as medidas que serão realizadas, tanto no que diz respeito aos sons vocálicos, seja com base em suas medidas formânticas diretas para as oito possibilidades em posição tônica encontradas no PB, [i e ε ə a ɔ o u], ou com base nas medidas indiretas obtidas através desses sons vocálicos – o VSA e as DE, quanto no que diz respeito às consoantes nasais. Em seguida, foram apresentadas as aplicações forenses, principalmente dos formantes vocálicos, em que se mostra que a aplicação das medidas de VSA, das DE e dos

formantes das consoantes nasais não têm aplicação corriqueira na CL. Por fim, tecemos considerações sobre a variabilidade intra e interindividual, tema que servirá de linha para determinar a propriedade de cada parâmetro para o uso forense através da relação entre essas variabilidades. No capítulo que segue, serão apresentados os conceitos fundamentais da Fonologia Articulatória bem como alguns aspectos que relacionam essa teoria aos princípios da Fonética Acústica e Articulatória recém apresentados.

3 FONOLOGIA ARTICULATÓRIA

Tradicionalmente, os campos da Fonética e da Fonologia apresentam uma relação de afastamento que beira a impossibilidade de intersecção, resultado da delimitação estabelecida pelas teorias linguísticas sobre seus objetos de análise, sejam fonemas ou traços, e do conseqüente desenvolvimento de raciocínios explicativos sobre tais objetos e seus fenômenos, principalmente no que se refere às teorias de origem estruturalista⁴⁴, que encaram a Fonética como o estudo da manifestação acústica ou da produção física do som, e a Fonologia como o estudo da representação abstrata do sistema sonoro. Frente a esse contexto dicotômico, a Fonologia Articulatória busca estabelecer a quebra de paradigmas teóricos.

No sentido de rever a relação estanque entre Fonética e Fonologia, a Fonologia Articulatória (BROWMAN E GOLDSTEIN, originalmente na obra de 1986 e estendida pelas obras de 1989a, 1990a,b,c, 1992, 1995)⁴⁵, propõe uma revisão fundamental de conceitos, ao determinar uma unidade de análise diferente das tradicionais noções de fonema e traço distintivo. A teoria parte do pressuposto de que a unidade mínima do conhecimento fonológico não é estática e tampouco discreta, mas sim dinâmica e que pode se sobrepor aos seus pares. Assim, os autores definem o *gesto articulatório* como unidade de trabalho, noção que restringe o significado laico do termo, a saber, movimento

⁴⁴ Incorre-se aqui em uma simplificação, uma vez que não é o foco do trabalho uma retomada histórica das teorias. A referência pretende incluir principalmente os expoentes de teorias inegavelmente expressivas da Linguística, encontrados nas obras de Saussure (1995), Trubetzkoy (1964), Jakobson e Halle (1956), Chomsky e Halle (1968) e Clements e Hume (1995). Embora se compreenda que a simplificação possa sofrer críticas, é inquestionável que essas perspectivas têm em comum, e aqui reside o objetivo da simplificação, a separação clara entre fonologia e fonética e tomam como objetos unidades discretas e estáticas, sejam fonemas ou traços. A respeito dessa evolução teórica, Silva (2002) realiza uma retomada histórica digna de leitura, por sua clareza e abrangência.

⁴⁵ Em Browman e Goldstein (2000), os próprios autores referem as obras de 1986, 1989a, 1990c e 1992 como as responsáveis por dar conta da formulação teórica, mas na obra de 1995, eles referem as obras de 1986, 1989a, 1990a,b e 1992. Goldstein e Fowler (2003) referem as obras de 1992 e 1995 como base para a formulação teórica. Cabe ressaltar que, devido à profícua publicação promovida pelo Laboratório Haskins em seu próprio *Status Report (SR)* e ao lapso temporal entre a publicação pelo Laboratório e por veículos externos, algumas incongruências e referências equivocadas foram encontradas no levantamento feito por ocasião deste estudo. O caso mais interessante e que se considera digno de nota ocorre quando Browman e Goldstein (1990c, p. 352) referiram que não haviam, ainda, investigado encontros consonantais em final de sílaba, quando Browman e Goldstein (1988) abordaram, pelo menos, a sobreposição dos gestos na sequência vogal-consoante, o que poderia, pelo menos, oferecer alguma consideração sobre a posição final de sílaba. Ocorre que, ao consultar as versões originais dos trabalhos no *site* do Laboratório, o trabalho publicado em *Papers in Laboratory Phonology I: between the grammar and the physics of speech*, em 1990c, foi publicado originalmente no SR-92, de dezembro de 1987, enquanto Browman e Goldstein (1988) aparece no SR-93/94, de janeiro de 1988.

muscular articulatório, ao estabelecer que esse primitivo linguístico não é a simples execução de movimentos resultantes de uma representação mental, mas que é composto por um conjunto de descrições objetivas, que colocam simultaneamente o gesto articulatório à serviço da abstração e da ação de movimentos articulatórios necessários para se atingir um objetivo articulatório. A Fonologia Articulatória, ao delimitar um objeto que dá conta das dimensões da representação e da produção unificadamente, assume um referencial, distinto em relação às demais teorias, que coloca numa perspectiva completamente diferente o relacionamento dos níveis fonético e fonológico, a ponto de considerar não mais dois, mas apenas um nível. Essa questão será recorrente neste capítulo e discussão breve sobre o tema será proposta no Capítulo 5.

Apesar das referências fundamentais ao aporte teórico da Fonologia Articulatória recaírem obrigatoriamente sobre as dezenas de trabalhos de Browman e Goldstein, cabe destacar que a produção dos estudos sobre os temas componentes da teoria se deu através de uma ação colaborativa promovida pelo Laboratório Haskins⁴⁶, onde pesquisadores conduzem estudos a respeito das bases biológicas da fala, linguagem e leitura. As seções seguintes oferecerão uma ideia da integração de estudos distintos e da maneira como se formaram as bases para a proposição da Fonologia Articulatória. Primeiramente serão oferecidos os argumentos de Fowler (1980) a favor de uma teoria que envolva a noção temporal em sua unidade de análise. Em seguida será apresentada a definição da unidade de análise proposta por Browman e Goldstein (1986), que toma por base as ideias de Fowler (1980) associadas a proposta de um modelo computacional que permite implementar as variáveis articulatórias de maneira simular a fala. Por fim serão apresentadas as vogais e o papel que cumprem dentro da proposta da Fonologia Articulatória.

3.1 A CONTESTAÇÃO DA UNIDADE DE ANÁLISE TRADICIONAL

Os primeiros conceitos que posteriormente serviriam à delimitação da Fonologia Articulatória foram estabelecidos por Fowler (1980) através da proposta de que as teorias

⁴⁶ O site <http://www.haskins.yale.edu/> descreve a atuação instituição.

tradicionais⁴⁷ falham em suas explanações a respeito de padrões coarticulatórios ou de outras manifestações que envolvam o controle temporal. Além disso, a autora descreve um esquema de princípios que devem reger qualquer teoria que se proponha consistente o suficiente para dar conta da coarticulação. As teorias fonológicas tradicionais apresentam um ponto em comum na relação de suas unidades de análise com a dimensão temporal, sejam essas unidades traços, fonemas ou segmentos. As análises fonológicas clássicas invariavelmente omitem a consideração do tempo na representação de seus primitivos, de forma que a temporalidade⁴⁸ se estabelece de maneira extrínseca aos fenômenos, ou seja, as unidades discretas são ordenadas sobre o eixo do tempo (FOWLER, 1980, p. 113).

Os modelos fonológicos tradicionais concebem o segmento como um elemento estático e discreto, o que, de acordo com Fowler (1980), não pode ser comprovado por manifestações acústicas ou articulatórias, uma vez que a representação discreta e a manifestação concreta não são páreas nem exatamente correspondentes. Ao contrário, os dados acústicos e articulatórios são elementos contínuos sem fronteira obrigatória no decorrer do tempo. Alguns de seus aspectos são claramente observados, dependendo da técnica aplicada na sua observação, como o modo de articulação, o ponto de articulação e o vozeamento. A temporalidade, porém, não é discreta ou evidente à análise do dado acústico. Não há bordas perpendiculares ao eixo do tempo demarcando início e fim de segmentos e, assim, a informação temporal não pode ser observada claramente. É a mente que possui o conceito de segmentos fonológicos que é imposto ao sinal acústico enquanto é percebido (FOWLER, 1980, p. 114). Os modelos fonológicos tradicionais não estabelecem a dimensão do tempo como um elemento previsto em suas propostas para a observação da coarticulação e, por esse motivo, Fowler (1980) classifica esses modelos teóricos como “teorias de tempo extrínseco”. Uma teoria cujo primitivo assuma as características dinâmica e contínua – em oposição às concepções de primitivo estático e discreto – deve considerar a temporalidade como inerente ao planejamento da produção articulatória.

⁴⁷ Ver nota 44, pois “tradicionais” tem a intenção de fazer uma generalização análoga em relação às teorias mencionadas naquele apontamento.

⁴⁸ Tradução do autor para o termo em inglês *timing*, com a intenção de **não** considerar *timing* como tradução de tempo.

Fowler (1980), ao questionar os conceitos básicos nas teorias de tempo extrínseco, considera que, para essas teorias, as unidades fonológicas têm três propriedades que não são verificadas em seus correlatos fonéticos, a saber, tais unidades são discretas, estáticas e livre de contexto. A proposição de uma teoria de tempo intrínseco deve contrapor essas propriedades e definir unidades que sejam contínuas, qualitativamente separadas e serialmente ordenadas, características pleiteadas para contrapor a característica *discreta* da unidade das teorias de tempo extrínseco. A unidade linguística, na teoria de tempo intrínseco, também deve ser dinâmica, opondo a característica *estática*, própria das teorias de tempo extrínseco. Em uma teoria de tempo intrínseco, o primitivo de análise deve ser sensível ao contexto, propriedade que contempla a variabilidade, em oposição à propriedade *livre de contexto* para as teorias de tempo extrínseco. Por fim, uma teoria de tempo intrínseco, apesar de não considerar o segmento como um primitivo, acomoda a noção de segmento ao tratá-lo como elemento qualitativamente separado e ordenado serialmente (FOWLER, 1980, p. 114; FOWLER ET AL., 1980, p. 374).

Além dos princípios e propostas recém expostos, que serviram como delineadores para a denominada “Teoria Geral da Ação”, Fowler et al. (1980) questionam o fato de, para as teorias de tempo extrínseco, o planejamento de uma sucessão de segmentos não corresponder à execução desses segmentos. Essa não correspondência entre plano e execução resulta na obrigatoriedade de um processo de transformação do segmento como unidade abstrata em unidade física, bem como o contrário, de manifestação física em representação abstrata. Tal processo, chamado de tradução, além de prever uma construção/desconstrução que deterioraria as informações acústicas/articulatórias, pois sua quantidade e grau de detalhamento exigiria reduções para viabilizar o processamento dessas informações, seria um procedimento custoso devido à relação entre a complexidade de informação e a velocidade em que se dá o processamento da compreensão e expressão através da comunicação verbal oral. As teorias que comportam algum tipo de tradução, em um ou mais níveis, são denominadas “Teorias de tradução”⁴⁹ (FOWLER ET AL., 1980, p. 376). Os argumentos apresentados em questionamento às abordagens de tempo extrínseco e a favor de um arcabouço teórico

⁴⁹ Tradução do autor para “Translation Theories”.

inovador servem para estabelecer as bases sobre as quais uma teoria de tempo intrínseco deve se configurar e, como será abordado a seguir, a Fonologia Articulatória respeita essas bases para fundamentar-se.

3.2 O GESTO ARTICULATÓRIO E SUA IMPLEMENTAÇÃO

A Teoria Geral da Ação, proposta por Fowler et al. (1980), percebe a dissociação existente entre as definições de unidades linguísticas das teorias de produção, de articulação e de percepção de fala, considera a importância dos estudos de coordenação de movimentos para as teorias linguísticas e propõe uma unidade linguística que baseie o estudo articulatório. Assim, ofereceram os princípios para a definição de uma unidade que preservasse a relação entre abstração e ação. Em continuidade às ideias de Fowler (1980) e Fowler et al. (1980), Browman e Goldstein (1986) alegam que as teorias existentes, apesar de se dedicarem a formalizar com detalhes uma diversidade de explicações para fatos e fenômenos específicos, não se prestaram a resolver o afastamento entre as áreas da Fonética e da Fonologia. Para minimizar ou eliminar esse afastamento, os autores oferecem uma representação fonológica em uma descrição explícita e direta do movimento articulatório nas dimensões do espaço e do tempo, o que simplifica, ou até torna desnecessária, a tradução complexa requerida em uma abordagem como a de regras de implementação, respeitando a ideia de Fowler et al. (1980).

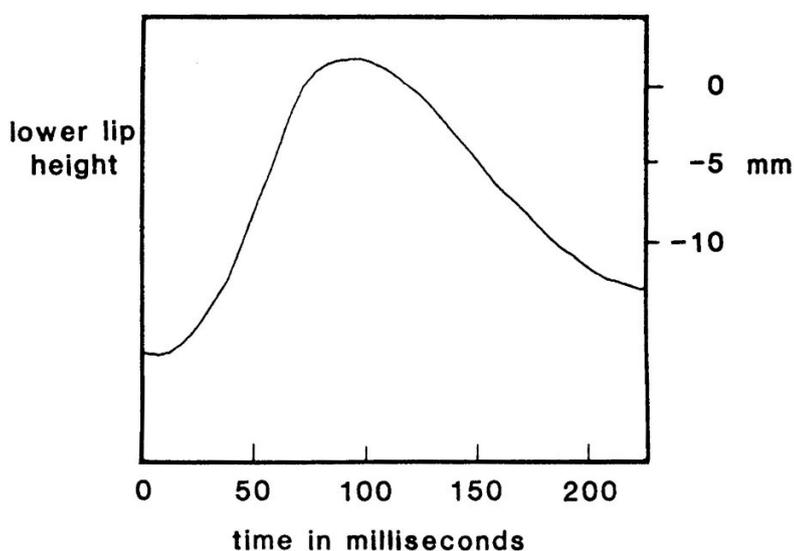
Browman e Goldstein (1986) sugerem que descrever a fala em termos de sobreposição de unidades articulatórias com tempo inerente pode dar conta da variação acústica, missão facilitada pela evolução dos recursos tecnológicos na área de rastreamento de movimentos articulatórios e pela produção de conhecimento para a descrição analítica e matemática de movimentos coordenados.

A formalização matemática do gesto parte do conceito de estruturas coordenativas de Turvey (1977), em estudo que buscou a construção de ações mediadas pela visão através do ajustamento de unidades relativamente autônomas. Essa formalização foi instanciada no modelo dinâmico de tarefa de Saltzman e Kelso (1987), que trata os gestos como estruturas autônomas que geram trajetórias no espaço e tempo sem a

necessidade de tradução entre planejamento e execução (BROWMAN E GOLDSTEIN, 1986, p. 222).

A definição inicial do gesto pressupõe a ausência de um estado de estabilidade no movimento, característica presente em todas as observações da fala, sejam articulatórias ou acústicas, fenômeno ilustrado pela Figura 3.1 a seguir, que apresenta a trajetória do lábio inferior na sequência [abə]. Cada um dos articuladores do trato vocal respeita essa característica de ausência de estado de estabilidade e esses articuladores podem ser medidos e representados analogamente ao que foi feito com o lábio inferior. Se cada um dos segmentos tradicionais pudesse ser representado por um e único gesto, como na relação do segmento bilabial e seu respectivo gesto de fechamento, então a Fonologia Articulatória não ofereceria vantagens representacionais ou implementacionais para o campo da fonologia. A relação entre o gesto e o segmento não é biunívoca. Assim o gesto será a instanciação de uma família de trajetórias relacionadas através de um modelo dinâmico – Dinâmica de Tarefa – descrito a seguir (BROWMAN E GOLDSTEIN, 1986, p. 223 e 224).

Figura 3.1 – Trajetória do lábio inferior para a sequência [abə].



Fonte: Browman e Goldstein (1986, p. 222).

A Dinâmica de Tarefa (SALTZMAN 1986; SALTZMAN E KELSO, 1987) define o movimento em termos da tarefa a ser cumprida, não sendo necessário que haja a definição específica da função para cada estrutura anatômica que compõe o movimento necessário à tarefa em questão. A ação envolvida na fala não precisa dar conta de cada um dos articuladores envolvidos na manifestação sonora desejada. Como a tarefa não visa às partes, nesse caso aos articuladores, mas sim ao conjunto formado por estes e ao resultado de sua ação conjunta, a ação do conjunto orientada para um objetivo nunca é desempenhada exatamente da mesma maneira. A dinâmica que envolve a ação do gesto articulatório pode ser resolvida por um sistema massa-mola proveniente da Física, mais precisamente da dinâmica oscilatória.⁵⁰

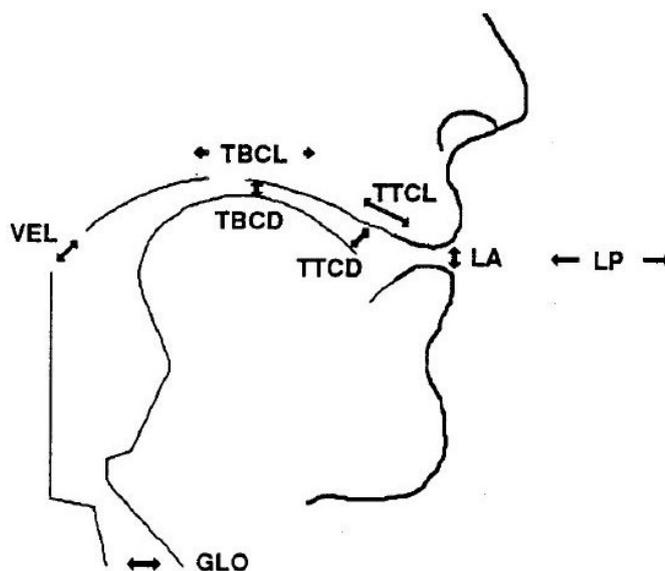
Com o propósito de apresentarem um modelo dinâmico computacional, Browman e Goldstein (1989a, p. 72; 1990c, p. 344; 1992, p. 24) definem o gesto como uma caracterização abstrata de movimentos coordenados de articuladores no trato vocal orientados ao cumprimento de tarefa. Para implementar o modelo computacional, o objetivo de cada gesto é especificado em termos de dimensões independentes da tarefa, denominadas “variáveis do trato vocal”. Cada variável é associada com a configuração de um conjunto articulador cujos movimentos estão vinculados ao valor da respectiva variável. Assim, a partir do objetivo da tarefa do gesto, as variáveis relativas a esse gesto têm seus parâmetros determinados e, com base nesses parâmetros, os articuladores buscam o cumprimento da tarefa. Não há o controle independente de cada articulador, o que seria um processamento muito mais complexo do que esse, orientado ao cumprimento da tarefa. A Figura 3.2 a seguir apresenta, no quadro que aparece na porção superior, na primeira coluna, o conjunto articulador que é representado pelo gesto, denominado conforme o efeitor final do respectivo conjunto, na segunda coluna; as

⁵⁰ A equação diferencial de segunda ordem para o sistema massa-mola com amortecimento, que resulta em uma senóide representativa do movimento oscilatório do gesto, é $m\ddot{x} + b\dot{x} + k(x-x_0) = 0$ onde,
 m = massa do objeto;
 b = amortecimento do sistema;
 k = rigidez da mola;
 x₀ = comprimento da mola na posição de repouso;
 x = deslocamento instantâneo do objeto;
 x' = velocidade instantânea do objeto;
 x'' = aceleração instantânea do objeto.

variáveis do trato vocal que especificam o gesto, antecedidas pelas respectivas siglas, a partir do inglês, e, na terceira coluna, os articuladores envolvidos em cada gesto. Na porção inferior da Figura 3.2, as siglas situam as respectivas variáveis anatomicamente.

Figura 3.2 – Efetores finais, variáveis do trato e articuladores envolvidos.

Efator final	Variáveis do trato	Articuladores envolvidos
Lábios	LP protrusão labial	Lábio inferior, superior e mandíbula
	LA abertura labial	
Ponta da língua	TTCL local de constrição da ponta da língua	Ponta e corpo da língua, mandíbula
	TTCD grau de constrição da ponta da língua	
Corpo da língua	TBCL local de constrição do corpo da língua	Corpo da língua e mandíbula
	TBCD grau de constrição do corpo da língua	
Véu palatino	VEL abertura vélica	Véu palatino
Glote	GLO abertura glotal	Glote



Fonte: Browman e Goldstein (1989a, p. 73; 1990c, p. 344; 1992, p. 24). Adaptação do Autor.

Observando-se o quadro expresso no alto da Figura 3.2, percebe-se que cada gesto pode ser especificado em termos de uma ou duas variáveis do trato. Os gestos orais envolvem pares de variáveis do trato, um membro do par determinando o local da constrição (*constriction location*) e o outro membro do par, o grau da constrição (*constriction degree*). Assim, tem-se: para o gesto de lábios, a protrusão labial (LP) e a abertura labial (LA); para o gesto de ponta da língua, o local de constrição da ponta da língua (TTCL) e o grau de constrição da ponta da língua (TTCD); e, para o gesto de corpo da língua, o local de constrição do corpo da língua (TBCL) e o grau de constrição do

corpo da língua (TBCD). Além disso, esses gestos podem ser representados apenas pela indicação de seu articulador efetor, lábios (LIPS), ponta da língua (TT) e corpo da língua (TB). O gesto vélico (VEL) envolve uma variável, a abertura vélica (VEL) e o gesto glótico (GLO) também envolve uma variável, abertura glotal (GLO), que indicam o tamanho da abertura ou grau de constrição do respectivo gesto.

Como se percebe na Figura 3.3 a seguir, Browman e Goldstein (1989a, p. 74) ainda previram outros recursos para o modelo, no intuito de que pudesse dar conta de maneira mais genérica de fenômenos linguísticos, como o gesto raiz da língua (TR) e suas respectivas dimensões de local e grau de constrição, a dimensão local da constrição para o gesto glotal, além da dimensão formato da constrição (*constriction shape*) para os gestos de ponta da língua e de corpo da língua. Todas essas dimensões e variável não estão implementadas no modelo e são aquelas indicadas por asteriscos na Figura 3.3.

Figura 3.3 – Inventário dos conjuntos articuladores e seus parâmetros.

Gestos:		
Conjunto articulador	Dimensões	
LIPS	(grau constr, lugar constr, LA LP)	rigidez
TT	(grau constr, lugar constr, forma constr,* TTCD TTCL)	rigidez
TB	(grau constr, lugar constr, forma constr,* TBCD TBCL)	rigidez
TR*	(grau constr,* lugar constr,*)	rigidez*
VEL	(grau constr, VEL)	rigidez
GLO	(grau constr, lugar constr,* GLO)	rigidez

Fonte: Browman e Goldstein (1989a, p. 74).

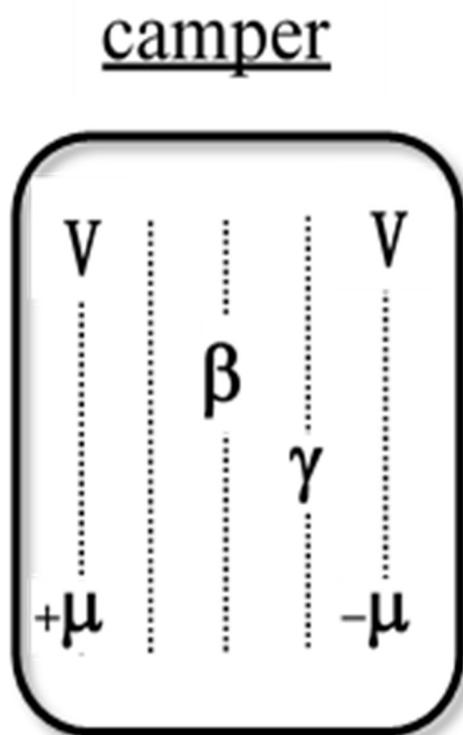
O gesto refere-se a um conjunto articulador que, assim, é composto por: grau de constrição, local de constrição, forma de constrição (não implementada no modelo, como já dito) e rigidez. Todas as variáveis apresentam, pelo menos, o parâmetro grau de constrição. O local de constrição é relevante apenas para os gestos orais e as respectivas variáveis (LP, TTCL e TBCL). Os parâmetros grau e local de constrição determinam o valor x_0 na equação dinâmica e o parâmetro rigidez, o valor k . Para que o modelo dinâmico de tarefa possa resolver a equação diferencial de cada gesto, os parâmetros que formam a equação devem ser oferecidos a esse modelo. Os valores numéricos dos parâmetros são contínuos, em milímetros para as variáveis de grau de constrição, em graus para as variáveis de local de constrição e Hz para rigidez, e por isso não estabelecem classes ou categorias no sistema. Para representar faixas contrastivas de

parâmetros gestuais de maneira discreta são utilizados os descritores gestuais “fechado”, “crítico”, “estreito”, “médio” e “largo” para grau de constrição e “protruso”, “labial”, “dental”, “alveolar”, “pós-alveolar”, “palatal”, “velar”, “uvular” e “faríngeo” para local de constrição (BROWMAN E GOLDSTEIN, 1989a, p. 75).

Quando um regime de controle para um gesto específico é ativado, a equação regula a coordenação dos articuladores para que cumpram a tarefa, ou seja, a formação e soltura da constrição, de acordo com o alvo, definido pelo grau e local de constrição e pela rigidez. Isso ocorre simultaneamente para cada gesto envolvido e os gestos de uma ocorrência de fala são organizados em uma estrutura coordenada superior, ou constelação, representada por uma pauta gestual, que será descrita a seguir.

Várias representações de pautas gestuais são propostas, de acordo com o fenômeno estudado, uma vez que as pautas comportam elementos de outras representações teóricas, e de acordo com a evolução do modelo (BROWMAN E GOLDSTEIN, 1986, 1989a,b, 1990a, 1990b, 1992). Com o objetivo de dar conta das representações de sobreposição dos gestos articulatorios, a pauta gestual recebe sobre si a disposição da constelação de gestos, onde as variáveis do trato podem ser alinhadas em camadas, permitindo a interpretação dos fenômenos articulatorios em que a temporalização de cada gesto está prevista na sua dimensão horizontal. Como os gestos aparecem dispostos em camadas, a relação de sincronização entre os gestos se torna evidente. A pauta não tem a pretensão de ser uma representação profunda de um nível não aparente do gesto, mas de servir como um indicador direto do gesto realizado e ser um elemento mais conveniente a sua descrição. Além disso, a pauta pode apresentar configurações diversas, de acordo com o fenômeno descrito (BROWMAN E GOLDSTEIN, 1990c). As Figuras 3.4 e 3.5 a seguir exemplificam apresentações de pautas diversas e ilustram também a evolução do modelo. A Figura 3.4 apresenta a pauta gestual da palavra “camper”, do inglês, onde percebe-se uma primeira proposta de representação dos gestos (BROWMAN E GOLDSTEIN, 1986). As camadas são ocupadas pelos gestos, em que “V” serve para vogal genérica, “β” para abertura e fechamento de lábios, “γ” para abertura e fechamento de glote e “μ” para o gesto vélico, onde “+” indica a abertura e “-” indica o fechamento.

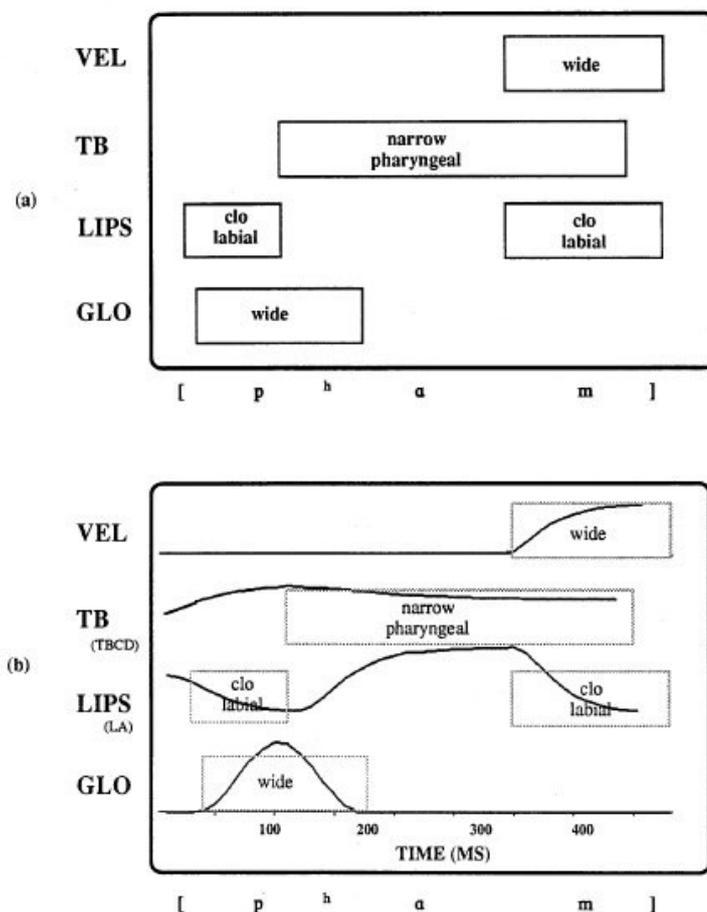
Figura 3.4 – Pauta gestual da palavra “camper”



Fonte: Browman e Goldstein (1986, p. 242).

Na Figura 3.5 a seguir, registra-se a indicação dos efetores fora da pauta propriamente dita. Os gestos são representados em caixas, ilustrando uma proposta de pauta mais evoluída e elaborada, no que diz respeito às bases teóricas do modelo, em relação à Figura 3.4. Nessa representação, onde se verifica a pauta gestual da palavra *palm*, a altura da caixa (sua dimensão vertical), tanto em (a) quanto em (b), indica o alvo (*target*) do gesto; cada caixa aparece confrontada com as trajetórias articulatórias efetivamente medidas na pauta em (b). A respeito das medições das trajetórias articulatórias, retoma-se o circuito da fala sob uma perspectiva multimodal proposto por Gick et al. (2013) apresentado na seção 2.1, Capítulo 2, através da Figura 2.2. Naquele circuito o dado acústico localizado na porção central do esquema é a matéria de estudo da Fonética Acústica. Para a Fonologia Articulatória, as matérias de estudo são os dados obtidos a partir das três dimensões de manifestação do circuito, ou seja, o dado acústico, os dados articulatórios e os dados de fluxo aéreo combinados.

Figura 3.5 – Pauta gestual da palavra “palm”



Fonte: Browman e Goldstein (1989a, p. 76).

Nota: Em (a) apenas as caixas representam a relação entre os gestos e em (b) as mesmas caixas foram sobrepostas às curvas de trajetórias dos gestos. *Wide*: largo; *narrow pharyngeal*: estreito faríngeo; *clo labial*: fechado labial.

A pauta gestual (b) apresentada na Figura 3.5 oferece uma ideia da temporalidade de ocorrência dos gestos, informação obtida pelas medidas do que equivaleria, no circuito da Figura 2.2, aos “movimentos faciais”, uma imprecisão, já que os movimentos monitorados nos estudos em Fonologia Articulatória extrapolam a face e envolvem as estruturas de língua, lábios, mandíbula, pregas vocais e véu palatino.

Ainda a respeito do circuito da fala sob perspectiva multimodal apresentado no Capítulo 2, cumpre observar que as estruturas envolvidas desde a perspectiva do “falante” são simétricas e análogas às estruturas envolvidas desde a perspectiva do “ouvinte”, ficando as manifestações, acústicas, motoras e aéreas, situadas na porção medial do circuito e servindo como elo entre as duas dimensões análogas, do falante e

do ouvinte. Essa analogia entre as duas dimensões corrobora a proposta da Fonologia Articulatória de que a unidade de análise, o Gesto Articulatório, se presta tanto à percepção quanto à produção, uma vez que serve às duas dimensões.

Para a Fonologia Articulatória, uma estrutura fonética não é simplesmente uma sequência linear de parâmetros ou traços, mas também deve ser descrita como uma espécie particular de organização dinâmica que os valores desses parâmetros servem para modular. As mudanças sucessivas em tais valores podem ser relacionadas a pontos particulares na organização dinâmica. Essa é uma diferença fundamental da teoria em relação às representações fonéticas tradicionais, que não oferecem nenhuma forma explícita de gerar as trajetórias articulatórias a partir de uma sequência de especificações de parâmetros.

A fim de tratar dessa questão, Browman e Goldstein (1990a) apresentam o modelo gestual computacional, representado na Figura 3.6 a seguir⁵¹. De acordo com a proposta, tem-se como entrada um texto escrito ou expresso em ARPABET⁵² (ver Anexo A) e como saída, fala sintética. Após a entrada, os parâmetros dinâmicos são originados pelo *modelo linguístico gestual* (BROWMAN E GOLDSTEIN, 1986, 1992) e determinam a *tarefa global do modelo dinâmico* (SALTZMAN 1986; SALTZMAN E KELSO, 1987) através da *pauta gestual*. As *trajetórias articulatórias* resultantes, alimentam o sintetizador articulatório (modelo do trato vocal) (RUBIN ET AL, 1981; ISKAROUS ET AL., 2003⁵³) que produz o resultado acústico.

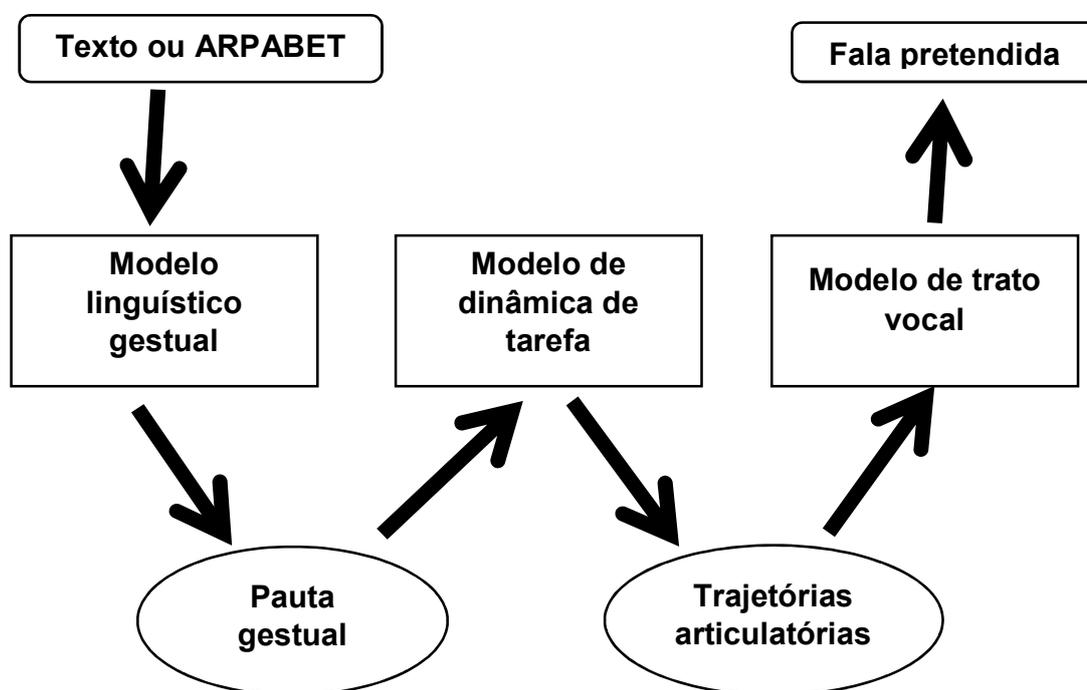
O esquema do processo completo envolvido no modelo computacional está representado na Figura 3.6 a seguir.

⁵¹ O modelo gestual computacional possui uma versão implementada, o Task Dynamic Application (Browman, Goldstein, Nam et al., 2003), disponível para *download* na página do Laboratório Haskins.

⁵² ARPABET é um código de transcrição fonética desenvolvido pela Agência de Projetos de Pesquisa Avançados, nos EUA, em que cada fonema da língua inglesa é representado por uma sequência de caracteres ASCII, permitindo sua interpretação por sistemas computacionais. O alfabeto consta no Anexo A.

⁵³ O sintetizador foi inicialmente idealizado por Rubin et al. (1981), porém o modelo atualmente utilizado é o CASY (ISKAROUS ET AL., 2003).

Figura 3.6 –Modelo gestual computacional.



Fonte: Browman e Goldstein (1990a, p. 300); Nam et al. (2012, 3982)⁵⁴.

Além do processo de sintetização de fala, o modelo da Fonologia Articulatória serve de fundamentação para as pesquisas no campo de inversão de fala, em que representações físicas do processo de fala são elaboradas a partir do sinal acústico da fala, seja através de pautas gestuais dedicadas a servirem ao reconhecimento de fala (NAM ET AL., 2010; NAM ET AL., 2012), seja através de modelos geométricos do trato oral (YEHIA, 1997; PANCHAPAGESAN E ALWAN, 2011) dedicados a representarem espacialmente o trato durante a fala. Os estudos no campo da análise/síntese podem oferecer elementos importantes para entendimento da organização dos gestos e em que medida eles são fenômenos que estão relacionados a aspectos do indivíduo ou da comunidade de fala e servirem, assim, a aplicações forenses. Mitra et al. (2010) e Mitra et al. (2011) já obtiveram resultados satisfatórios em pesquisas com objetivo de realizar o reconhecimento automático de falantes através da obtenção de variáveis do trato, que reforçam a ideia de aplicação no campo forense.

⁵⁴ O modelo foi originalmente proposto em 1990, mas a figura específica é do artigo de 2012, conforme referido na fonte da Figura 2.6. A única diferença é na entrada, que mais recentemente se definiu por texto ou ARPABET, mais um indicativo da evolução da implementação baseada na teoria da Fonologia Gestual.

A representação da vogal na pauta gestual da palavra *camper*, conforme apresentada na Figura 3.4, recebe uma definição genérica, diferentemente do que ocorre com todos os gestos relacionados à articulação das consoantes. Essa definição genérica e a representação vocálica no modelo computacional suscitam considerações específicas que serão apresentadas na seção a seguir para posteriormente serem discutidas na seção 5.3.2, Capítulo 5.

3.3 AS VOGAIS NA FONOLOGIA ARTICULATÓRIA

O conhecimento consagrado em Fonética Acústica e em Fonética Articulatória já deixou evidente que as vogais e consoantes são eventos articulatórios e acústicos de natureza claramente diferentes (como foi abordado no Capítulo 2) e a Fonologia Articulatória adota uma perspectiva que corrobora tal entendimento.

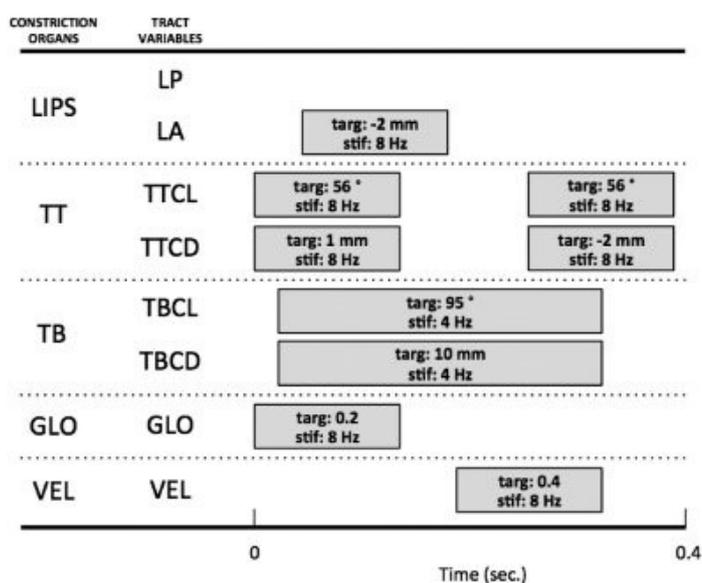
Quando considerado o movimento oscilatório do gesto, os gestos vocálicos são claramente mais lentos do que os gestos consonantais. As vogais são realizadas fundamentalmente pelo movimento do corpo da língua, que altera o formato da cavidade oral, enquanto os consoantes são obstruções rápidas do trato vocal. Browman e Goldstein (1990c, p. 354), ao definirem a organização da pauta gestual em camadas articulatórias independentes, expõem com precisão essa diferença ao afirmarem que

“Contiguity operates differently on the C and V tiers. This captures a functional difference between vowels and consonants, where vowels act as a kind of background to the “figure” of the consonants. On the one hand, gestural contiguity on the V tier is completely independent of the C tier. This reflects the fact [...] that vowel articulations are contiguous (or partially overlapping), regardless of the number of consonants intervening. On the other hand, contiguity on the C tier is sensitive to intervening vocalic gestures, in the sense that consonantal gestures overlap considerably less, if at all, when a vocalic gesture intervenes. A related effect of this functional difference has to do with the change in stiffness as more consonants are inserted between vowels: the vowel gesture decrease in stiffness while the consonantal gestures increase their stiffness (BROWMAN E GOLDSTEIN, 1990c, p. 354).⁵⁵”

⁵⁵ A contiguidade funciona de forma diferente nas camadas C e V. Isso capta uma diferença funcional entre vogais e consoantes, em que as vogais atuam como um tipo de fundo para a "figura" das consoantes. Por um lado, a contiguidade gestual na camada V é completamente independente da camada C. Isso reflete o fato de que [...] as articulações da vogal são contíguas (ou parcialmente sobrepostas), independentemente

Nam et al. (2012) atribuem valores padrão para rigidez e estabelecem como referenciais 8Hz para as consoantes e 4Hz para as vogais, alegando que esse parâmetro serve justamente para diferenciar as categorias de consoantes e vogais (NAM ET AL., 2012, p. 3981). A Figura 3.7 a seguir apresenta os parâmetros numéricos atribuídos à pauta gestual para a palavra “span”.

Figura 3.7 – Pauta gestual para a palavra *span*: modelo gestual computacional.



Fonte: Nam et al. (2012, p. 3981).

Como fica evidente nas definições dos parâmetros das variáveis, as vogais são definidas com uma rigidez menor do que as consoantes, 4Hz e 8Hz respectivamente. A frequência estipulada para as variáveis de TB como a metade daquela estipulada para as demais variáveis resulta em gestos consonantais com o dobro da velocidade relativa aos gestos vocálicos. Na pauta gestual apresentada na Figura 3.7 observa-se essa relação através da dimensão horizontal das caixas, de forma que aquelas referentes aos gestos vocálicos têm o dobro do comprimento em relação aquelas referentes aos gestos consonantais. A ideia, apresentada por Fowler et al. (1980, pg. 410-411), de que as

do número de consoantes intervenientes. Por outro lado, a contiguidade na camada C é sensível aos gestos vocálicos intervenientes, no sentido de que os gestos consonantais se sobrepõem consideravelmente menos, se for o caso, quando um gesto vocalico intervém. Um efeito relacionado desta diferença funcional tem a ver com a mudança de rigidez à medida que mais consoantes são inseridas entre as vogais: o gesto da vogal diminui a rigidez enquanto os gestos consonantais aumentam a rigidez.

vogais são mais lentas do que as consoantes, fica preservada na implementação do modelo.

As vogais apresentam propriedades essenciais que são livres de contexto, porém não são movimentos invariantes, o que caracteriza tais propriedades como dinâmicas e não como alvos canônicos a serem atingidos (FOWLER ET AL., 1980, pg. 410-411). Assim, não há um formato fixo do trato oral que corresponda ao alvo canônico para cada vogal, mas há uma propriedade dinâmica que as caracterizam e que pode ser tomada como algo semelhante a um alvo espacial. A princípio, isso contrariaria a afirmação de Browman e Goldstein (1986), já referida neste capítulo, de que o gesto não possui um estado de estabilidade, o que foi ilustrado pela Figura 3.1. As vogais, porém, ocupam uma posição peculiar na Fonologia Articulatória, como será abordado a seguir.

Na evolução da descrição teórica de Browman e Goldstein (1986), ao considerar-se a pauta gestual apresentada na Figura 3.4, percebe-se o desenvolvimento do raciocínio teórico sobre os gestos vocálicos, inicialmente definidos como a primeira camada daquela pauta e a respeito dos quais não se estabelece maiores descrições. Browman e Goldstein (1988) limitam-se a apenas indicar que a primeira linha da constelação é ocupada pelas vogais expressas genericamente. A descrição das vogais também é deixada de lado no experimento proposto na mesma obra, a respeito do que os autores justificam que não foram capazes de, com o mesmo *led* – recurso utilizado para traçar as trajetórias articulatórias – que acompanha a função das consoantes, verificar o comportamento das vogais, restringindo sua pesquisa apenas a considerações a respeito dos padrões consonantais (BROWMAN E GOLDSTEIN, 1988, p. 88) e deixando de lado a descrição das vogais segundo o modelo.

A respeito dessa imprecisão na definição das vogais, em trabalho seguinte, Browman e Goldstein (1989a) afirmam que a descrição para as vogais é explicada em termos dos parâmetros TBCD e TBCL (ver Figura 3.2). Porém eles reconhecem que, para algumas distinções entre vogais, é possível que os parâmetros descritos não sejam suficientes e que sejam necessários outros parâmetros para estabelecer-se a distinção linguística. Essa indistinção fica explícita ao comparar-se as definições das consoantes

e das vogais expressas na Tabela I.2⁵⁶ (ver Anexo B), que consta no apêndice do manual do modelo da Aplicação de Dinâmica de Tarefa – Task Dynamic Application, doravante TADA – (Browman, Goldstein, Nam et al., 2006). Na Tabela I.2, as vogais têm atribuídas a si apenas descritores de TBCL, [palatal], [faríngeo] e [uvular]. O descritor TBCD fica definido genericamente como “v” e as vogais do inglês são distinguidas apenas pelos valores numéricos do *target*, o que sugere que, pelo menos no que se refere à implementação, a vogal é uma classe que dificilmente pode ser categorizada.

As definições dos descritores dos gestos articulatórios no TADA funcionam categoricamente principalmente para as consoantes, como se pode constatar ao analisar na Tabela I.2. A Figura 3.8 a seguir ilustra a porção inicial da Tabela I.2, com alguns exemplos para as oclusivas.

Figura 3.8 – Parcial da Tabela I.2 de “gestos atribuídos a segmentos”: Oclusivas

Stops						
ARPA	Organ	Osc	TV	Constr	Target	Stiff
B	Lips	clo	LA	CLO	.	.
	Lips	rel	LA	REL	.	.
	Velum	clo	VEL	CLO	.	.
P	Lips	clo	LA	CLO	.	.
	Lips	rel	LA	REL	.	.
	Glottis	h	GLO	WIDE	.	.
	Velum	clo	VEL	CLO	.	.
M	Lips	clo	LA	CLO	.	.
	Lips	rel	LA	REL	.	.
	Velum	n	VEL	WIDE	.	.
D	TT	clo	TTCL	ALV	.	.
	TT	clo	TTCD	CLO	.	.
	TT	rel	TTCL	REL	.	.
	TT	rel	TTCD	REL	.	.
	Velum	clo	VEL	CLO	.	.
T	TT	clo	TTCL	ALV	.	.
	TT	clo	TTCD	CLO	.	.
	TT	rel	TTCL	REL	.	.
	TT	rel	TTCD	REL	.	.
	Glottis	h	GLO	WIDE	.	.
	Velum	clo	VEL	CLO	.	.

Fonte: Adaptado de Browman, Goldstein, Nam et al. (2003).

Como se observa na tabela da Figura 3.8, e ocorre de forma semelhante para todas as consoantes apresentadas, os descritores são suficientes para representar a distinção entre segmentos propostos, sem a necessidade de haver uma atribuição de valores para alvo (Target) ou rigidez (Stiff).

⁵⁶ No Anexo B, apresenta-se a Tabela I.2 com a listagem integral das definições de parâmetros para todos os sons do inglês americano.

No que diz respeito às vogais, os descritores apenas não são suficientes para categorizar todas as vogais, havendo casos em que a distinção se dá pela atribuição dos valores de alvo e/ou de rigidez, como está exemplificado na Figura 3.9⁵⁷ a seguir.

Figura 3.9 – Parciais da tabela de “gestos atribuídos a segmentos”: Vogais.

Vowels

ARPA	Organ	Osc	TV	Constr	Target	Stiff
IH	TB	v	TBCL	PAL	.	3
	TB	v	TBCD	V	8	3
EY	TB	v	TBCL	PAL	.	.
	TB	v	TBCD	V	8.5	.
EH	TB	v	TBCL	PAL	.	3
	TB	v	TBCD	V	11.5	.
AE	TB	v	TBCL	PHAR	.	.
	TB	v	TBCD	V	17	.
AA	TB	v	TBCL	PHAR	.	.
	TB	v	TBCD	V	11	.

Fonte: Adaptado de Browman, Goldstein, Nam et al. (2003).

De acordo com a Figura 3.9, as vogais “IH”, “EY” e “EH” possuem os mesmos descritores para as variáveis TBCL e TBCD, [palatal] e [v], respectivamente. De maneira análoga, as vogais “AE” e “AA” possuem, para as mesmas variáveis, os mesmos descritores [phar] e [v]. Para os dois casos, as distinções são estabelecidas por valores não categóricos de alvo, em milímetros, e/ou rigidez, em Hz.

Sobre a questão do poder distintivo do modelo teórico para as vogais, Byrd (2003, p. 90) aponta a necessidade de estudos para entender com mais profundidade as categorias relacionadas às vogais e afirma ser necessário pesquisar representações para dar conta do contraste de tenso/solto e de ATR, por exemplo.

Comparando-se as pautas da Figura 3.5 e da Figura 3.7, percebe-se a relação entre os descritores dos gestos, em nível representacional, e as dimensões numéricas que atuam na implementação computacional do modelo. A respeito das categorias

⁵⁷ A Figura 2.9 foi montada com partes recortadas da tabela original, para dispor todos os elementos – título, cabeçalho das colunas e os dois grupos de vogais – na mesma imagem.

estabelecidas pelos descritores e sua relação com a implementação das vogais no modelo serão propostos questionamentos no Capítulo 5, seção 5.3.

As bases da Fonologia Articulatória foram apresentadas, com particular foco sobre as vogais, uma vez que esses sons são objeto precioso no exame de Comparação de Locutores e material fundamental presente nesta pesquisa, juntamente com as consoantes nasais. Apesar de o capítulo anterior ser aquele que contém a fundamentação teórica principal a ser aplicada sobre os dados desta pesquisa, o capítulo atual ofereceu uma introdução aos conceitos da Fonologia Articulatória, alguns deles relacionados intimamente com conceitos da Fonética Acústica e que foram referidos na tentativa de estreitar o relacionamento entre os campos de pesquisa. A estratégia de mensuração da consistência dos parâmetros propostos, sejam os tradicionais formantes vocálicos ou os aqui pleiteados como consistentes, será apresentada no capítulo seguinte, juntamente como o delineamento metodológico e instrumental da pesquisa em tela.

4 METODOLOGIA

O presente trabalho caracteriza-se como uma pesquisa prospectiva de caráter quantitativo-descritivo, cujo desenvolvimento é experimental e que tem por tema a variabilidade individual de aspectos acústicos e articulatórios específicos da fala. Espera-se que o número de participantes estabelecido valorize os achados sobre a variabilidade interindividual e que a repetição de coletas para cada participante permita oferecer material de estudo consistente sobre variabilidade intraindividual. As coletas deram-se com risco ou constrangimento mínimos para os participantes e foram realizadas através da aplicação de instrumento, a ser descrita no item 4.2, elaborado especificamente para este trabalho.

Primeiramente neste capítulo, na seção 4.1, será apresentada a amostra consolidada para a elaboração da pesquisa, através da descrição de seu perfil e dos critérios de seleção dos participantes. O procedimento de tratamento e análise acústica dos dados será descrito na seção 4.3 e, na seção 4.4 será apresentado o conjunto de variáveis estabelecidas. A análise estatística conduzida terá seu delineamento explanado ao final do capítulo, na seção 4.5.

4.1 AMOSTRA E SELEÇÃO DE PARTICIPANTES

A população alvo de investigação refere-se aos indivíduos masculinos adultos da cidade de Porto Alegre – RS. O fato de se ter optado exclusivamente por indivíduos masculinos, bem como por não se contemplar a faixa etária como variável de pesquisa, deve-se à demanda efetiva de perícias de CL, realizada majoritariamente em material proveniente da fala de homens adultos. Apoio-me no argumento de Foulkes et al. (2010, p. 712-3) de que, para se entender a variação na fala, é mais importante investigar as diferenças dentro dos grupos de gêneros do que entre gêneros. Adicionalmente, pesquisas que tratam da normalização de medidas acústicas entre gêneros dão suporte para uma comparação entre os dados desta pesquisa e de eventuais pesquisas com indivíduos femininos (MIRANDA, 2012; MIRANDA E MEIRELES, 2012).

Com relação ao número de participantes, tomou-se como base os estudos de mesma natureza citados na seção 2.4, Capítulo 2 (GLENN E KLEINER, 1968; WOLF, 1972, HÖFKER 1977, 1983; HEUVEL, 1996), que se configuraram na dimensão entre

dez e trinta participantes. Conforme Kinoshita e Ishihara (2015), que realizou estudo a respeito da relação entre tamanho de amostra linguística e consistência de resultados voltados para o uso forense, vinte indivíduos seria o número mínimo aceitável para uma pesquisa nessa área. Assim, e também considerando o volume resultante de dados, a amostra foi composta pela fala de vinte indivíduos, os quais não apresentam histórico de alterações orgânicas que comprometam a voz ou a fala, são residentes na cidade de Porto Alegre, monolíngues, filhos de pais também monolíngues. Apresenta-se no Quadro 4.1 a seguir as idades dos indivíduos no momento da coleta e os respectivos graus de escolaridade.

Quadro 4.1 – Idades e níveis de escolaridade dos participantes..

Informante	Idade	Escolaridade	Informante	Idade	Escolaridade	Informante	Idade	Escolaridade
ACRR	46	Médio	EGS	38	Superior	PMP	50	Médio
AFM	30	Superior	FTC	48	Médio	PROF	48	Superior
AFRB	48	Superior	IAB	44	Superior	RCS	25	Superior
APG	41	Médio	MFS	33	Médio	RFS	50	Médio
CAPM	53	Médio	MPG	51	Médio	TBS	25	Médio
CDK	58	Médio	MRR	34	Superior	VJP	59	Médio
CRD	36	Médio	MZ	35	Superior			

Fonte: O Autor (2017).

Todos os participantes deste estudo são colegas de trabalho, plantonistas, com carga horária de 40 horas semanais. A distribuição de carga horária na semana varia conforme a função, porém cada um dos indivíduos realiza pelo menos um plantão de 24 horas em dia fixo da semana. Assim sendo, foram realizadas as coletas em dias de plantão dos participantes. Essa estratégia foi escolhida tendo em vista que a amostra foi projetada para ser composta por vinte indivíduos que foram gravados repetidamente em turnos e dias diferentes, como será detalhado no item 4.2 a seguir.

4.2 COLETA DE DADOS

Como é intuito deste trabalho a pesquisa da variabilidade individual, definiu-se como objetivo de coleta a obtenção da fala monitorada através de instrumento de coleta para eliciar as frases-veículo. Cada indivíduo foi gravado em três momentos diferentes no mesmo dia, a fim de se contemplar eventuais diferenças que pudessem ocorrer em função do estado de uso/cansaço vocal. Assim, realizaram-se sessões de coleta

individuais em cada turno do dia (manhã, tarde e noite) por sete dias diferentes. Dessa forma, a coleta através do instrumento foi estimada para ser realizada vinte e uma vezes (3 turnos x 7 dias) por participante⁵⁸. A combinação dos elementos considerados para a pesquisa, a saber, o número de indivíduos (em vinte), os turnos do dia (em três) e as repetições semanais das sessões de gravação (sete⁵⁹), resultou em uma estimativa de 420 sessões de gravação. O número de gravações repetidas estimado, vinte e um, foi estabelecido conforme o número de participantes, uma vez que para testar a variabilidade intraindividual necessita-se de um número razoável de repetições.

Antes da primeira gravação, foi apresentado, para cada participante, o Termo de Consentimento Livre e Esclarecido (apresentado no Apêndice A) e o Termo de Compromisso para Utilização de Dados (apresentado no Apêndice B), momento em que foi explicado como se daria a participação na pesquisa.

Iniciou-se cada coleta em uma manhã com a intenção de realizar as gravações adicionais à tarde e à noite do mesmo dia. Para cada turno, foi estabelecido um intervalo de horário dentro do qual seriam efetuadas todas as coletas – entre 9h e 10h para a manhã, entre 15h e 16h para a tarde e entre 21h e 22h para a noite. Porém a natureza da atividade profissional dos participantes (pronto-atendimento) não permitiu garantir que essas três gravações ocorressem sempre no mesmo dia. Assim, quando não foi possível efetuar uma gravação em um determinado turno, foram realizadas gravações extraordinárias para os turnos eventualmente faltantes.

Ao iniciarmos cada gravação, foi oferecida uma cópia do instrumento (Apêndice C) para que o participante se guiasse, de forma que a ordem de produção das frases sempre foi a mesma.

Todas as gravações foram acompanhadas pelo autor do trabalho e, antes de encerrar cada gravação, foi solicitada a repetição de itens específicos nos casos de intercorrência durante a gravação (ruídos indesejados ou equívocos de leitura⁶⁰, por exemplo), que, nesse caso, era editada, conforme está explanado no item 4.3.2. Após o

⁵⁸ Como será descrito seguir, o número de gravações não foi exatamente o previsto e a quantidade efetiva de gravações por informante está apresentada no Quadro 4.1, na seção 4.4.

⁵⁹ Nem sempre foi possível gravar os três turnos no mesmo dia, uma vez que os indivíduos eventualmente saíam para atendimento. Se, ao final do sétimo dia de gravação houvesse turnos faltantes, prosseguia-se com a gravação em dias adicionais até complementar sete repetições para cada turno.

⁶⁰ Considerei equívoco de leitura quando uma vogal foi produzida no lugar de outra.

término de cada sessão de coleta, foi realizada a oitiva do registro de áudio e, caso se apresentasse comprometida após a saída do participante, a gravação era descartada e a sessão repetida em gravação extraordinária, da mesma forma que foi feito nos casos em que o participante estava em atendimento externo durante o intervalo do respectivo turno de gravação. Assim, não foram efetuadas gravações fora das faixas de horários estabelecidas.

Somente após a coleta de sete gravações por turno para cada participante foi considerado encerrado o conjunto de gravações de fala monitorada⁶¹.

O instrumento de coleta para fala monitorada foi formado pela combinação de seis contextos consonantais precedentes ([f s ʃ m n ɲ] às oito possibilidades de realização vocálica ([i e ε ə a ɔ o u]), totalizando 48 sílabas, ao que foram acrescentadas mais cinco em que a consoante nasal bilabial fechava a sílaba iniciada pela fricativa lábio-dental desvozeada (itens “FAM”, “FEM”, “FIM”, “FOM” e “FUM” do instrumento, apresentado no Apêndice C). Como não foi previsto realizar medições nos segmentos vocálicos nasais, tampouco nas consoantes palatais, as sílabas formadas por esses segmentos serviram como distratores⁶². Cada sílaba foi eliciada pela frase-veículo “Ele dirá ___ calmamente”. Um exemplo de espectrograma com a segmentação e etiquetagem da frase veículo, processos a serem explicados na seção 4.3.2 adiante, é apresentado na Figura 2.4, no Capítulo 2.

4.2.1 Equipamentos e ambiente de gravação

Os equipamentos disponíveis para a gravação foram um microfone de baixa impedância marca **Shure**, modelo **Beta 58**, ligado a uma placa de captura de áudio marca **Edirol**, modelo **UA-25**, conectada a um computador portátil marca **Dell**, modelo **Inspiron 1545**, com processador **intel® Core™2 Duo 2,20 GHz**, com 4 Gb de memória RAM.

O ambiente de gravação foi a sala do pesquisador na própria instituição onde trabalhavam os participantes por ocasião da coleta. Não houve condições de realizar os procedimentos em ambiente com isolamento acústico, assim buscou-se, dentro do possível, controlar as fontes de ruído nos arredores da sala.

⁶¹ Um dos participantes forneceu apenas 19 gravações devido ao seu afastamento por licença.

⁶² Além desses cinco itens, os oito iniciados pela nasal palatal também foram considerados distratores, totalizando treze itens.

Cada participante foi posicionado sentado em frente ao microfone, mantendo a distância de 10 a 15 cm entre sua boca e o microfone.

4.3 ANÁLISE ACÚSTICA: INSTRUMENTO E PROCEDIMENTO

Para gravação, edição, etiquetagem e análise acústica foi utilizado o software Praat (BOERSMA E WEENICK, 2011, doravante Praat), obtido gratuitamente através do site <http://www.fon.hum.uva.nl/praat/> e atualizado periodicamente com a instalação da versão mais recente.⁶³ A taxa de amostragem utilizada para a gravação do sinal de fala foi 44,1 KHz e a quantização foi em 16 bits. Os áudios foram salvos e formato WAVE PCM em um canal apenas.

4.3.1 Preparação, denominação e etiquetagem dos arquivos

Após encerrada cada sessão de gravação, caso houvesse itens repetidos por intercorrências durante sua produção, conforme já dito, o arquivo de áudio era editado. Recortava-se a frase-veículo inteira, que era então substituída pela frase-veículo contendo o item repetido, preservando assim a ordem de ocorrência da sílaba em questão e, conseqüentemente, facilitando o procedimento de etiquetagem. Além disso, foram recortados todos os segmentos que não interessavam à pesquisa, como conversas anteriores ou posteriores à leitura do instrumento, pausas para aguardar que algum ruído temporário cessasse ou devido a interrupções eventuais. Logo após o encerramento da sessão de gravação, o arquivo de áudio era salvo em formato WAVE e era nomeado obedecendo o critério “NOME-número de ordem por diaTURNO”, onde “NOME” está indicado pelas iniciais do indivíduo em caixa alta, “número de ordem por dia” indica o dia em que foi feita a coleta⁶⁴ e “TURNO” é a indicação do turno por um dos caracteres “M, T ou N”, para manhã, tarde ou noite, respectivamente. Seguem exemplos de arquivos para ilustração: ACRR-1M.wav; EGS-7T.wav; VJP-15N.wav.

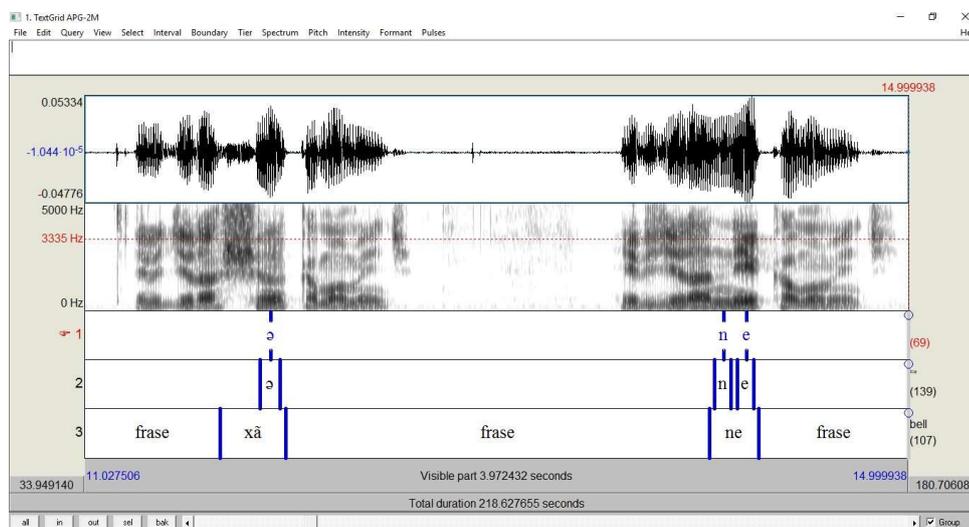
⁶³ A Versão 6.0.28 era a mais atual, em maio de 2017, devendo-se levar em conta que a atualização de versões não compromete as análises oferecidas pelo *software*.

⁶⁴ O esperado seriam dígitos de 1 a 7, pois foram projetados sete dias diferentes, mas como houve a necessidade de completar turnos específicos em dias extras, a numeração acabou sendo mais alta, dependendo do número de vezes em que não se conseguiu realizar uma gravação com determinado informante.

Após cada arquivo ter sido editado para eliminação das intercorrências, passou-se ao seu etiquetamento, através do recurso de anotação do Praat, que permite associar um arquivo de texto ao arquivo de áudio. O arquivo de texto contém as informações temporais, que visualmente são percebidas como barras verticais no editor do *software*⁶⁵, e de anotação definidas pelo usuário e que podem ser visualizadas e operadas a partir da janela de edição do Praat, bem como podem ser operadas a partir de um outro arquivo contendo conjunto de comandos quaisquer do Praat, que é chamado de *script*⁶⁶.

O recurso de etiquetagem permite a criação de diferentes camadas de marcação de instantes. Assim, foram criadas três camadas para cada arquivo gravado, uma para o ponto de medição⁶⁷, outra para o trecho de medição dos parâmetros acústicos, onde estava indicado o som medido, vogal oral ou consoante nasal, e outra para a indicação da sílaba em questão. A Figura 4.1 a seguir apresenta uma janela do Praat como exemplo da etiquetagem.

Figura 4.1 – Exemplo de etiquetagem utilizada. Janela de edição do Praat.



Fonte: O Autor (2017).

⁶⁵ Em realidade, o arquivo TextGrid não passa de um arquivo com extensão “.txt” que contém os instantes das marcações e as anotações efetuadas pelo usuário.

⁶⁶ O *script* também é um arquivo de texto que contém sequências de comandos em sintaxe de programação semelhante às linguagens computacionais mais simples. O *script* pode ser executado de dentro do ambiente do Praat e permite a automatização de teoricamente qualquer procedimento realizado sobre os áudios e outros objetos relativos ao programa.

⁶⁷ Inicialmente, a etiquetagem foi realizada com marcações de pontos de medição na primeira camada e com a indicação da sílaba na segunda camada. O *script* utilizado inicialmente previa o ponto de marcação. O *script* efetivamente utilizado, mais adequado à proposta, foi elaborado após a etiquetagem e previa trecho para medição, não ponto. Foi elaborado um *script* intermediário para inserir o trecho em torno do ponto original e a camada de ponto foi preservada.

O trecho de medição para as vogais que seguem as fricativas refere-se à porção medial da vogal e para as vogais que seguem as nasais tomou-se uma porção deslocada mais à direita, com o intuito de evitar o efeito coarticulatório proveniente da nasalidade consonantal, uma vez que as medidas serão comparadas. Para as consoantes nasais, o trecho de medição também é o medial. Essas informações servem de referência para a varredura efetuada pelo *script* de medição, a ser descrito a seguir.

4.3.2 Script para medição

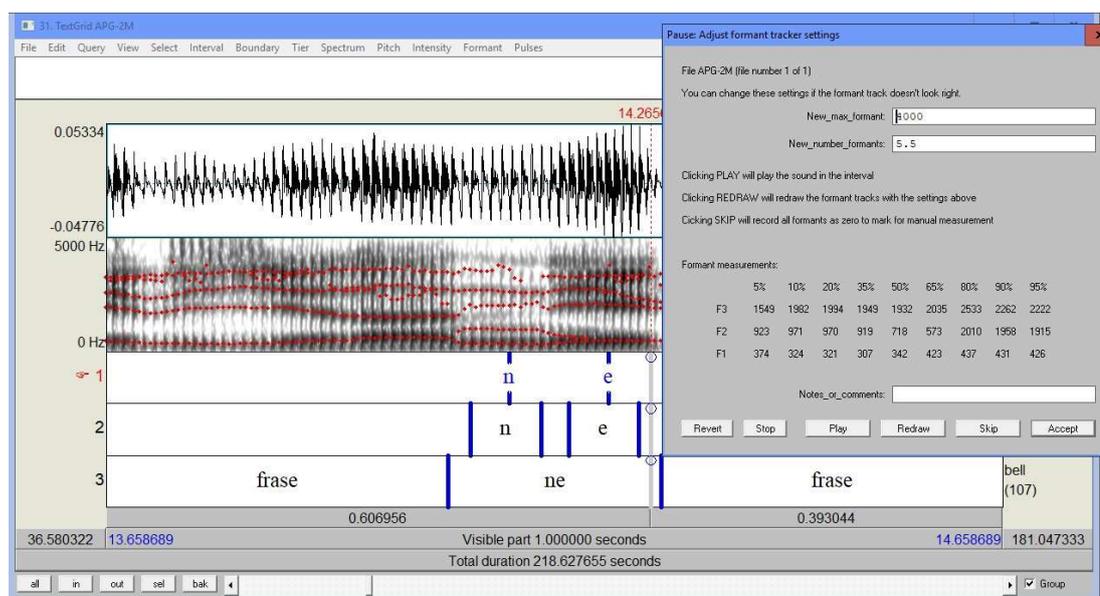
Foi utilizada a versão 0.3 do *script* de McCloy e McGrath (2012), que permite a medição dos parâmetros acústicos com monitoramento do segmento que está sendo medido através da janela de edição do Praat, o que torna a tarefa mais lenta, uma vez que a rotina para a cada medição permite que se visualize o espectrograma com o histórico de formantes e se decidam as configurações de medição do parâmetro. Apesar de tornar o procedimento mais lento, esse monitoramento permite controlar o acoplamento do LPC ao respectivo formante do espectrograma⁶⁸ (THOMAS, 2011, p. 43). A Figura 4.2 a seguir apresenta a janela de operação do *script* em primeiro plano e a janela de edição em segundo, onde se monitorava o acoplamento dos formantes.

O *script* teve seu conjunto de comandos alterado para que ficasse configurado conforme a necessidade do presente estudo (ver o código do *script* no Anexo C). Como a definição do número de coeficientes para determinar o número de formantes calculados é diferente entre vogais orais e consoantes nasais, foram feitas adaptações para a medição em cada um desses grupos, de forma que o *script* varreu, primeiramente, todas as vogais e, em um segundo momento, através de alteração de seu código, as nasais. Dessa forma, foram geradas duas planilhas de dados, uma para as vogais, como resultado da primeira varredura, e outra para as nasais, resultado da segunda varredura.

⁶⁸Com relação ao cálculo do LPC, apesar da previsão do número de formantes ser em torno de um a cada 1000Hz para falantes masculinos adultos, o número de coeficientes pode variar de acordo com características individuais, forçando que se determine números de coeficientes diferentes para cada indivíduo e para vogais com qualidades diferentes de um mesmo indivíduo. Por exemplo, as vogais anteriores podem necessitar de um número de coeficientes diferente em relação às vogais posteriores para o mesmo indivíduo.

Eventualmente, para corrigir equívocos durante o processo, o *script* era alterado para permitir a busca por som específico.

Figura 4.2 – Exemplo de janela do *script* utilizado sobre a janela de edição do Praat..



Fonte: O Autor (2017).

4.4 VARIÁVEIS

As variáveis acústicas consideradas para a pesquisa foram os três primeiros formantes⁶⁹ e a frequência fundamental, que foram obtidos através do programa computacional Praat. Além dos parâmetros acústicos, definimos como variáveis o som medido (vogais e consoantes nasais), o contexto próximo ao som (três fricativas desvozeadas e duas consoantes nasais como contexto anterior ao som vocálico e oito sons vocálicos como contexto seguinte aos sons nasais), o turno em que ocorreu a gravação (manhã, tarde e noite) e o participante (indicados pelas iniciais). Algumas gravações extras foram efetuadas para fins de reposição de algum turno faltante, resultando em 436 gravações. O Quadro 4.2 a seguir apresenta a distribuição de gravações por participante.

⁶⁹ Apesar de a maioria dos registros recebidos para perícia não apresentar qualidade acústica para se obter formantes acima do terceiro, eventualmente recebemos registros em que é possível sua visualização.

Quadro 4.2 – Número de sessões de gravação por participante.

Informante	Gravações	Informante	Gravações	Informante	Gravações
ACRR	23	EGS	22	PMP	21
AFM	23	FTC	21	PROF	23
AFRB	21	IAB	23	RCS	21
APG	22	MFS	22	RFS	22
CAPM	23	MPG	21	TBS	21
CDK	22	MRR	23	VJP	19
CRD	23	MZ	21		

Fonte: O Autor (2017).

Em cada sessão, das 53 sílabas pronunciadas, 40 trechos de medição referiam-se às vogais (8 sons vocálicos X 5 contextos anteriores) e 16 trechos referiam-se às nasais (2 nasais X 8 sons vocálicos como contexto posterior). Como resultado, obteve-se a medição de 17.440 sons vocálicos e 6.976 sons consonantais. As medidas acústicas foram tomadas em nove pontos (não foram todos utilizados para a obtenção da média, conforme será explicado a seguir) de cada trecho sonoro etiquetado, aos 5, 10, 20, 35, 50, 65, 80, 90 e 95 por cento da duração do trecho, valores que podem ser vistos na janela do *script*, como se observa na Figura 4.2, resultando em uma planilha com 156.960 entradas para as vogais e em outra planilha com 62.784 entradas para as nasais. Essas duas planilhas com 219.744 entradas serviram de base para a composição das planilhas efetivamente utilizadas para a análise estatística a ser descrita.

4.5 RECURSOS PARA A ANÁLISE ESTATÍSTICA

O recurso de tabulação e disposição dos dados para análise estatística foi montado com base em uma série de planilhas: as duas primeiras, uma para as vogais e outra para as nasais, produzidas pelas saídas do *script* rodado sobre o Praat, e as demais elaboradas a partir de processamentos recursivos de dados. Todas as planilhas são compatíveis com o programa computacional Excel, que oferece as funções necessárias para a estatística descritiva bem como recursos de otimização para organização de dados, conforme será descrito a seguir.

A otimização de dados mencionada foi possível através da utilização da função “SUBTOTAL” e da ferramenta de filtro do Excel. A função permite que, ao se definir uma planilha intermediária, as medidas de estatística descritiva sejam automaticamente

atualizadas quando aplicado o filtro na planilha original. As medidas escolhidas inicialmente foram a média, o desvio-padrão, o coeficiente de variação, a amplitude amostral, o valor mínimo e o valor máximo como descritores do conjunto de dados. Definida essa planilha intermediária, associou-se outra planilha, que contém o conjunto completo dos dados. Esse procedimento foi feito separadamente para as vogais e para as nasais, a partir das duas planilhas originais referidas na seção anterior. Dessa maneira, utilizando o recurso de filtro do Excel, ao selecionar os parâmetros que determinaram cada subconjunto analisado, através da associação entre as duas planilhas, é possível obter instantaneamente os valores descritivos, que eram copiados e colados em uma terceira planilha de destino.

Esse procedimento foi repetido para cada planilha e resultou nos quadros que estão apresentados no Capítulo 5, sendo que para cada conjunto uma combinação própria de critérios foi selecionada para a filtragem dos dados componentes da respectiva planilha. Para todas as combinações de critérios utilizou-se como critério constante a seleção dos três pontos centrais dos trechos, ou seja, os pontos em 35, 50 e 65 por cento do trecho marcado, o que se refere à zona de estabilidade do trecho, lembrando que no caso das vogais antecederidas por nasais essa região era mais à direita, conforme foi explicado na seção 4.3.1, Capítulo 4.

Os referenciais estatísticos escolhidos para conduzir as análises de variabilidade individual presentes neste estudo são baseados na noção de Coeficiente de Variação. Para respeitar a noção de variabilidade individual proposta por Heuvel (1996) e descrita na seção 2.5.4, Capítulo 2, serão estabelecidos, para cada variável analisada, seja ela um parâmetro acústico (f_0 , F_1 , F_2 ou F_3) ou um som (vocálico ou consonantal nasal): (i) um indicador da variabilidade intraindividual; (ii) um indicador da variabilidade interindividual e (iii) a razão matemática entre esses indicadores (indicador da variabilidade intraindividual / indicador da variabilidade interindividual), que servirá como critério para determinar a consistência do parâmetro para a CL. Dessa maneira, o valor do coeficiente de variação médio individual, calculado a partir da média dos coeficientes de variação de cada participante, será considerado o referente de variabilidade intraindividual, doravante CV_{intra} . O coeficiente de variação de um conjunto de dados da amostra inteira, calculado a partir dos valores de média e desvio padrão obtidos, será

considerado como o referente de variabilidade interindividual, doravante CV_{inter} . A razão entre o CV_{intra} e o CV_{inter} , que servirá de referente comum de variabilidade de cada variável analisada, será denominado Índice de Variabilidade, doravante IV. O IV servirá como indicador para o ordenamento dos sons e parâmetros. Os valores de CV serão expressos em porcentagem e o de IV, em absoluto, com três casas decimais, para facilitar a diferenciação entre IV e CV durante a leitura do próximo capítulo, em que os resultados serão apresentados e discutidos.

Cumprе ressaltar que o CV_{intra} deverá ser tanto maior quanto menor for a amostra investigada, uma vez que o desvio padrão tenderá a variar mais quanto menor for o tamanho da amostra. Contudo, o CV_{intra} é calculado sobre amostras dependentes e assim seu comportamento tenderá a ser menor uma vez que existe uma correlação maior entre amostras do que o CV_{inter} , que é obtido sobre amostras independentes.

Feita a ressalva, é importante manter a noção de que o fundamental para a aplicação do conceito da variabilidade à evidência forense é a relação entre variabilidade intra e intersujeitos. Gonçalves (2014), por exemplo, utilizou o coeficiente de correlação intraclasse, que é uma razão de variâncias independentes, para verificar essa relação e Nolan (1983) e Rose (2002) referem a razão R (a partir de *F-ratio*) como representante da relação de variabilidades.

A escolha por se trabalhar com a razão entre os CV se deve à forma como os dados foram organizados e a noção intuitiva que essa razão assume, bem como à intenção de propor uma solução alternativa à questão, ainda que se assuma um risco em consequência dessa escolha. Os dados ficam à disposição para futuras pesquisas que pretendam testar a relação segundo outras ferramentas estatísticas.

5 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

A pesquisa aqui apresentada busca a proposição de indicadores de desempenho de parâmetros acústicos utilizadas para o exame de CL. Os parâmetros referentes aos sons vocálicos são tradicionalmente utilizados, porém os formantes das consoantes nasais, o VSA e as DE, que não são tradicionalmente utilizados na CL, estão sendo apresentados como possíveis parâmetros para a CL⁷⁰. A estratégia estabelecida para a avaliação de desempenho tem por base a relação entre a variabilidade intraindividual e a variabilidade interindividual, resumida no IV, e todos os parâmetros apresentados serão submetidos à análise neste capítulo. Além de servirem à aplicação forense, os resultados são descritivos do PB e, como tal, serão comparados a trabalhos sobre o PB que levantaram os parâmetros acústicos das vogais e das consoantes nasais. Além da análise acústica com base na variabilidade individual, a Fonologia Articulatória, por ser uma teoria que se propõe como vantajosa na implementação computacional da fala, terá seus conceitos discutidos pensando-se na contribuição que pode oferecer ao campo forense.

Os resultados referentes às vogais e suas derivações, o VSA e as DE, serão apresentados na seção 5.1 e os resultados referentes às consoantes nasais, na seção 5.2. Em cada uma dessas duas seções, em primeiro lugar, os resultados serão confrontados, mediante uso de estatística descritiva, com resultados de natureza semelhante provindos de pesquisas que abordaram o PB. Ainda para cada uma das duas primeiras seções deste capítulo, em segundo lugar, serão descritos os valores estatísticos de variabilidade, principalmente os de coeficiente de variação, que servirão de base para a definição de um referencial numérico, originado pela relação de variabilidade intra e interindividual que permitirá estabelecer o ordenamento dos parâmetros acústicos e dos sons medidos, incluindo as medidas geométricas do VSA e das DE que o compõem.

⁷⁰ As larguras de banda dos formantes, vocálicos e nasais, apresentaram-se com valores de variabilidade inicialmente muito altos já nos primeiros ensaios. Por esse motivo, não foram considerados seus achados nos resultados apresentados. As diferenças de variabilidade entre os turnos estabelecidos, manhã, tarde e noite, foram claramente desprezíveis e, assim, foram consideradas simplesmente como repetições de gravação.

Na seção 5.3 será discutida a aplicabilidade, na área forense, dos resultados baseados em indicadores de variabilidade, sugeridos nas seções anteriores, das vogais de forma geral, dos espaços acústicos dos triângulos vocálicos e de distâncias euclidianas entre algumas dessas vogais, com o objetivo de determinar indicadores de variabilidade para esses parâmetros, analogamente aos estabelecidos para as vogais e nasais. Por fim, encerrando a seção 5.3, os princípios da Fonologia Articulatória serão discutidos com relação a sua possível contribuição para a atividade forense.

5.1 VOGAIS

O foco deste trabalho recai sobre a variabilidade individual de parâmetros acústicos medidos em sons específicos. O percurso para a obtenção de indicadores de variabilidade que permitam a comparação entre os parâmetros, assim como entre os diferentes sons, passa inicialmente pela mensuração acústica de curto termo de tais sons e o levantamento da estatística descritiva dessas medidas. A comparação com outros trabalhos do PB que contemplam estatísticas descritivas em seu bojo será conduzida, apesar de não atender diretamente ao foco principal desta pesquisa, a variabilidade individual, por oferecer um elemento importante para o procedimento de CL.

Considerações a respeito do sistema linguístico a partir do qual os dados foram obtidos podem ser úteis à CL, tendo em vista que o exame não se baseia apenas em medidas acústicas, pois adota outras referências para o processo de cotejo das falas, o que significa que outros aspectos além dos parâmetros acústicos de curto termo são levados em consideração para determinar as semelhanças entre as amostras questionada e padrão. De fato, os peritos procuram incluir nos laudos periciais uma caracterização que permita situar a fala do indivíduo sob análise em relação à comunidade de fala na qual está inserido. Caso se encontrem particularidades nessa relação, mais um vestígio a respeito do indivíduo está sendo levantado. Extrapolando os limites da CL, a descrição de características e fenômenos regionais de língua ainda pode servir para traçar um perfil linguístico que colabore com aquelas investigações em que não se têm suspeito definido, na medida em que oferecem elementos para delimitar o escopo da investigação. Frente aos argumentos colocados, será exposto a seguir um discreto, porém intencional, desvio do foco principal da tese.

Usualmente, nas pesquisas linguísticas, os resultados médios dos valores de F1 e F2 são apresentados paralelamente a outros trabalhos que envolvam, de uma forma ou de outra, medidas acústicas. Os trabalhos que serão alinhados comparativamente estão entre aqueles referidos na seção 2.3, Capítulo 2.

A primeira comparação alinha apenas os valores médios dos dois primeiros formantes dos dados obtidos por: 1) Moraes, Callou e Leite (1996) para os três participantes masculinos de Porto Alegre, dados esses que são de fala espontânea; 2) Rauber (2006) para os seis participantes masculinos, dois de cada estado da Região Sul; 3) Escudero et al. (2009) para os dez participantes masculinos de São Paulo; 4) Barbosa e Madureira (2015) para os quatro participantes masculinos de estados distintos. Os trabalhos indicados por 2, 3 e 4 utilizaram de coleta de fala através de instrumentos, conforme foi esclarecido na seção 2.3, Capítulo 2. A estratégia de tomar os subconjuntos de participantes masculinos e do PB visa minimizar as eventuais discrepâncias consequentes de metodologias diferentes em cada trabalho. Tal minimização pretendida leva em consideração o possível de ser efetuado, uma vez que o cenário ideal pressuporia delineamentos previamente estabelecidos de forma análoga, situação difícil de encontrar no universo da linguística, argumento ilustrado pelo Quadro 2.2, da seção 2.3, Capítulo 2, que resume os delineamentos metodológicos de dez pesquisas envolvendo dados do PB.

O Quadro 5.1, a seguir, permite a comparação dos valores médios de F1 e F2 dos quatro trabalhos recém listados com aqueles valores obtidos nesta pesquisa. Os valores das vogais cardeais estão expressos na primeira linha e servem para fixar um ponto comum de comparação para cada trabalho.

Na comparação proposta, que pode ser visualmente acompanhada através dos respectivos traçados vocálicos expostos na Figura 5.1 adiante, os valores obtidos neste trabalho se aproximam, de forma geral, mais daqueles obtidos pelos três trabalhos mais recentes, de Rauber (2006), Escudero et al. (2009) e Barbosa e Madureira (2015), do que do trabalho mais antigo, de Moraes, Callou e Leite (1996). Essa conclusão torna-se mais facilmente observável ao avaliar os resultados de F1 para as extremidades do traçado vocálico, em que as vogais fechadas apresentam valores mais altos e a vogal aberta valor mais baixo no trabalho de Moraes, Callou e Leite (1996). Quando comparados os

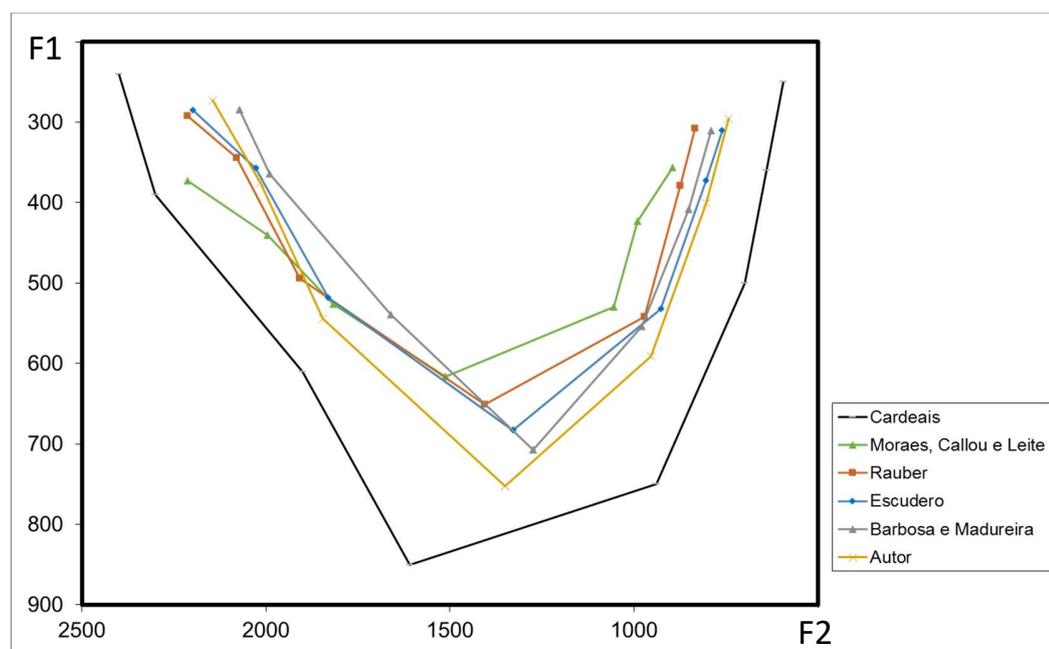
valores de F2, essa relação de aproximação e afastamento deste trabalho em relação aos demais se verifica novamente para os valores da vogal aberta e das vogais posteriores. Os trabalhos mais recentes apresentam valores menores de F2 do que o trabalho de Moraes, Callou e Leite (1996).

Quadro 5.1 – Comparação com trabalhos que mediram F1 e F2.

Autor	Fmt	i	e	ɛ	a	ɔ	o	u
Cardeais	F1	240	390	610	850 750	500	360	250
	F2	2400	2300	1900	1610 940	700	640	595
Moraes, Callou e Leite	F1	373	440	526	616	530	423	356
Porto Alegre 1996	F2	2213	1996	1816	1513	1056	990	896
Rauber	F1	292	344	494	651	542	379	308
	2006	F2	2212	2080	1908	1405	971	874
Escudero	F1	285	357	518	683	532	372	310
	2009	F2	2198	2028	1831	1329	927	804
Barbosa e Madureira	F1	284	364	539	707	554	408	310
	2015	F2	2072	1991	1660	1274	979	852
O Autor	F1	273	375	544	752	591	400	296
	2017	F2	2145	2014	1845	1350	954	804

Fonte: O Autor (2017), adaptado a partir de Moraes, Callou e Leite (1996), Rauber (2006), Escudero et al. (2009) e Barbosa e Madureira (2015). Os valores das Cardeais foram obtidos em Catford (1977).

Figura 5.1 – Comparação de pesquisas do PB.



Fonte: O Autor (2017), adaptado a partir de Moraes, Callou e Leite (1996), Rauber (2006), Escudero et al. (2009) e Barbosa e Madureira (2015).

Relativamente aos valores das vogais cardeais, os valores de F1 indicam que, para o conjunto de trabalhos mais recentes, incluindo o presente trabalho, as vogais fechadas são mais fechadas e a aberta mais aberta do que no trabalho de Moraes, Callou e Leite (1996). F2 indica que as vogais posteriores são mais posteriores quando feito o mesmo agrupamento dos trabalhos.

Apesar da possibilidade de o agrupamento ter se estabelecido por cronologia, uma explicação possível para esse efeito é de ordem metodológica, relativa à estratégia de obtenção da produção de fala, como será abordada na seção 5.3.

No que diz respeito a parâmetros de variabilidade, apenas os trabalhos, entre aqueles expostos no Quadro 5.1, de Rauber (2006) e de Escudero et al. (2009), apresentam dados de desvio padrão de forma a serem pareados com os resultados do presente trabalho. Além disso, as duas pesquisas dedicaram-se a medir valores de f0 e de F3, além dos tradicionalmente medidos F1 e F2. No Quadro 5.2, a seguir, repetimos os valores médios desses dois trabalhos juntamente com os obtidos na presente pesquisa, acrescentando os valores de f0 e F3. Os valores médios estão dispostos na metade superior do conjunto e os valores de desvio padrão estão dispostos na metade inferior do conjunto que forma o Quadro 5.2.

Quadro 5.2 – Comparação com trabalhos que incluíram f0 e F3, além de F1 e F2.

	Rauber (2006)				Escudero <i>et al.</i> (2009)				O Autor (2017)			
Média												
	f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3
i	151	292	2212	2950	137	285	2198	2952	139	273	2145	2867
e	146	344	2080	2755	131	357	2028	2719	130	375	2014	2628
ɛ	136	494	1908	2614	124	518	1831	2572	124	544	1845	2490
a	136	651	1405	2346	122	683	1329	2324	122	752	1350	2329
ɔ	135	542	971	2334	122	532	927	2335	124	591	954	2344
o	146	379	874	2466	132	372	804	2380	133	400	804	2447
u	157	308	834	2526	140	310	761	2309	142	296	743	2350
Desvio Padrão												
	f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3
i	41	23	130	214	27	22	171	195	25	33	153	228
e	26	35	170	200	24	27	154	209	23	36	139	188
ɛ	30	64	133	156	23	46	132	129	22	59	131	168
a	28	109	157	224	24	65	117	195	22	102	94	217
ɔ	30	69	105	266	22	85	100	161	22	70	95	250
o	35	37	128	186	26	37	74	143	24	39	94	207
u	35	21	174	261	31	22	76	180	26	33	114	203

Fonte: O Autor (2017), adaptado a partir de Rauber (2006) e Escudero et al. (2009).

Retomando a comparação dos valores médios, agora para os quatro parâmetros acústicos, f0, F1, F2 e F3, e para os três estudos referidos, verifica-se a compatibilidade dos valores entre os estudos. O comportamento de f0 nos três trabalhos confirmou a tendência de as vogais fechadas apresentarem valores de média maiores (151 para [i] e 157 para [u]) do que aqueles das vogais abertas (146 para [e]; 136 para [ɛ] e [a]; 135 para [ɔ] e 146 para [o]), conforme Ohala e Eukel (1987) já afirmaram, para o inglês americano, em uma proposta de explicação, sem conclusão determinante, para tal comportamento.

Com relação ao comportamento global dos valores médios, apesar da pesquisa de Rauber (2006) ter tomado medidas de indivíduos da Região Sul e, assim, ocasionar uma expectativa de maior proximidade de seus valores com os do presente estudo, exclusivamente com participantes de Porto Alegre/RS, chama principalmente a atenção a proximidade entre os valores médios obtidos pelo presente estudo e aqueles de Escudero et al. (2009). Para estimar a proximidade, calculou-se as diferenças percentuais entre os resultados da pesquisa em tela e os resultados de Rauber (2006) e os de Escudero et al. (2009). As diferenças estão apresentadas no Quadro 5.3 a seguir e foram obtidas a partir do Quadro 5.2, através do seguinte procedimento de cálculo: seja “A” o valor da célula referente ao resultado obtido pelo Autor (2017) e “C” o valor obtido pelo estudo a ser comparado, Rauber (2006) ou Escudero (2009), sempre tomados relativamente ao mesmo parâmetro da mesma vogal, a diferença será igual a $((A-C)/A)*100$. Resultados negativos indicam que o valor correspondente a este estudo é menor do que o comparado e resultados positivos, o contrário.

Quadro 5.3 – Diferença em percentuais deste estudo para aqueles de Rauber (2006) e Escudero et al. (2009).

	Proximidade entre as medidas do Autor (2017) e							
	Rauber (2006)				Escudero et al. (2009)			
	f0	F1	F2	F3	f0	F1	F2	F3
i	-8%	-7%	-3%	-3%	1%	-4%	-2%	-3%
e	-11%	9%	-3%	-5%	-1%	5%	-1%	-3%
ɛ	-9%	10%	-3%	-5%	0%	5%	1%	-3%
a	-10%	16%	-4%	-1%	0%	10%	2%	0%
ɔ	-8%	9%	-2%	0%	2%	11%	3%	0%
o	-9%	6%	-8%	-1%	1%	8%	0%	3%
u	-10%	-4%	-11%	-7%	1%	-5%	-2%	2%

Fonte: O Autor (2017).

Tomando-se os valores absolutos percentuais expresso no Quadro 5.3 e calculando-se a média de todos os valores absolutos das diferenças de cada conjunto, observa-se uma proximidade maior dos valores da pesquisa em tela com os de Escudero et al. (2009), já que a diferença média para os valores de Rauber (2006) resultou em 6% e para Escudero et al. (2009), em 3%. Uma possibilidade de explicação para essa proximidade seria o número de participantes da amostra, questão que merecerá consideração na seção 5.3.

A respeito dos resultados de desvio padrão para f_0 , expressos no Quadro 5.2, não se observa uma relação evidente de variabilidade entre as vogais, o que pode ser observado para F_1 , nos três trabalhos, em que as vogais abertas ([a, ε, ɔ] apresentam valores maiores dos que as vogais fechadas ([i, e, o, u]). Para F_2 , os valores de desvio padrão são maiores para as vogais anteriores ([i, e, ε]) do que para as posteriores ([u, o, ɔ]), com a exceção do valor de [u] em Rauber (2006). Para F_3 , percebe-se em [ε] o menor valor (156 para Rauber (2006); 129 para Escudero et al. (2009) e 168 para esta pesquisa) em comparação aos das demais vogais para as três pesquisas, indicando uma possível vantagem desse som sobre os demais com relação à variabilidade, o que pode estar relacionado ao resultado de Gonçalves et al. (1995) para vogal sustentada, conforme discutido na seção 2.3, Capítulo 2, em que esse som foi indicado como aquele em que a conformação do trato tem menor influência de ressonância sobre a fonte sonora das pregas vocais.

Ao encerrar a comparação entre pesquisas, retoma-se aqui o percurso em busca de um indicador comum de variabilidade. Assim, no que concerne aos valores médios e de variabilidade específicos obtidos neste trabalho, serão apresentados os resultados partindo do panorama mais genérico e seguir-se-á com a estratificação dos agrupamentos em mais conjuntos, com cada vez menos ocorrências em cada conjunto. Desse modo, em primeiro lugar, serão apresentados os valores das vogais sem considerar nenhuma estratificação por contexto anterior, depois serão apresentados os resultados estratificando-os conforme a classe do contexto anterior à vogal, consoantes nasais ou fricativas desvozeadas, e por fim estratificando-os por som do contexto anterior à vogal, [m, n, f, s, ʃ].

O Quadro 5.4, a seguir, estratifica os resultados simplesmente por vogais, sem considerar o som antecedente, e apresenta os valores de média; de desvio padrão; valor mínimo; valor máximo; de amplitude da amostra – resultado da subtração de máximo por mínimo; de CV_{inter} , calculado a partir dos valores de média e desvio padrão que constam no quadro, e que será considerado como o referente de variabilidade interindividual; de CV_{intra} , o valor do coeficiente de variação médio individual, calculado a partir da média dos coeficientes de variação de cada participante, e que será considerado como o referente de variabilidade intraindividual; e, ao final do Quadro 5.4, o de IV, a razão entre o CV_{intra} e o CV_{inter} , que servirá de referente comum de variabilidade de cada variável analisada⁷¹.

⁷¹ A respeito de ressalvas sobre o uso do IV, ver comentário ao final da seção 4.5, Capítulo 4.

Quadro 5.4⁷² – Resultados gerais de média e variabilidade (desvio padrão, mínimo, máximo, amplitude, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas⁷³ aparece na segunda coluna.

	N	Média				Desv				Mínimo				Máximo				
		f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3	
í	6540	139	273	2145	2867	25	33	153	228	58	183	250	1985	283	485	2712	3532	í
e	6540	130	375	2014	2628	23	36	139	188	53	157	390	1879	212	627	2588	3457	e
ε	6540	124	544	1845	2490	22	59	131	168	53	196	584	1767	200	1068	2447	3165	ε
a	6540	122	752	1350	2329	22	102	94	217	50	173	759	1284	207	1328	1694	3465	a
ɔ	6540	124	591	954	2344	22	70	95	250	51	209	513	1180	195	904	2218	3199	ɔ
o	6540	133	400	804	2447	24	39	94	207	51	251	520	1059	268	604	2333	3168	o
u	6540	142	296	743	2350	26	33	114	203	58	166	295	1134	231	562	2303	3107	u
ə	6540	131	438	1403	2476	23	58	141	182	52	129	583	1347	212	892	1851	3153	ə
		Amplitude				CV geral (interindiv)				CV médio (intraindiv)				Índice de Variabilidade				
		f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3	f0	F1	F2	F3	
í	6540	225	302	2462	1547	18.1%	11.9%	7.2%	8.0%	8.2%	5.9%	3.5%	4.4%	0.453	0.495	0.485	0.549	í
e	6540	159	469	2198	1578	17.4%	9.6%	6.9%	7.1%	7.9%	7.1%	4.3%	3.6%	0.452	0.737	0.616	0.509	e
ε	6540	147	873	1863	1398	17.3%	10.8%	7.1%	6.7%	7.8%	6.2%	2.8%	3.5%	0.447	0.580	0.389	0.512	ε
a	6540	157	1155	935	2181	17.7%	13.6%	7.0%	9.3%	7.6%	6.2%	3.1%	3.6%	0.429	0.452	0.438	0.389	a
ɔ	6540	143	696	1706	2018	18.1%	11.8%	10.0%	10.7%	8.6%	6.8%	2.8%	3.5%	0.478	0.574	0.280	0.325	ɔ
o	6540	217	353	1813	2110	17.9%	9.8%	11.7%	8.5%	7.8%	6.5%	6.3%	3.6%	0.434	0.658	0.541	0.427	o
u	6540	173	396	2008	1973	18.6%	11.0%	15.3%	8.7%	8.1%	5.9%	5.0%	4.4%	0.433	0.531	0.324	0.506	u
ə	6540	160	763	1268	1807	17.3%	13.2%	10.0%	7.4%	8.5%	6.8%	8.4%	4.4%	0.493	0.511	0.832	0.595	ə

Fonte: O Autor (2017).

⁷² Os valores decimais aparecerão separados por ponto neste e nos demais quadros, enquanto que, no texto, apresentados separados por vírgula.

⁷³ O número de medidas neste quadro e nos quadros 5.9, 5.13 e 5.14 será referido como “N”, sempre posicionado na segunda coluna, após a indicação do som medido.

Em um primeiro momento, poderia parecer intuitivo utilizar o coeficiente de variação médio por indivíduo, o CV_{intra} , ou mesmo o coeficiente geral médio, o CV_{inter} , como critérios para definir o melhor parâmetro ou som para a CL, já que esse referencial estatístico, o CV, tem a característica de, por ser a divisão do desvio padrão pela média, normalizar as diferenças de grandeza entre as unidades métricas envolvidas, e, por ser uma média da amostra, seja individual ou geral, representar o conjunto dos indivíduos. Porém, uma vez que a relação entre as variabilidades intraindividual e interindividual é o foco da pesquisa aqui apresentada, o IV foi definido, conforme descrito na seção, 4.5, Capítulo 4, como critério de ordenamento para os sons e parâmetros. Seguindo essa orientação serão expostos os resultados a seguir.

Os valores expressos no Quadro 5.5 a seguir foram retirados do Quadro 5.4, do quadrante relativo ao IV, e ordenados, em cada coluna, ou seja, para cada parâmetro, crescentemente de cima para baixo. À esquerda de cada valor consta a respectiva vogal, pois os ordenamentos vocálicos são independentes em cada coluna de parâmetros, e foram acrescentadas uma coluna, à direita, com os valores médios dos parâmetros acústicos para cada vogal, e uma linha, abaixo, com os valores médios das vogais para cada parâmetro acústico.

Quadro 5.5 – Ordenamento dos sons vocálicos por parâmetro acústico.

Ordenamento das vogais conforme IV para cada parâmetro									
	f0		F1		F2		F3		Por vogal
a	0.429	a	0.452	ɔ	0.280	ɔ	0.325	ɔ	0.381
u	0.433	i	0.495	u	0.324	a	0.389	a	0.426
o	0.434	ə	0.511	ɛ	0.389	o	0.427	u	0.432
ɛ	0.447	u	0.531	a	0.438	u	0.506	ɛ	0.472
e	0.452	ɔ	0.574	i	0.485	e	0.509	i	0.493
i	0.453	ɛ	0.580	o	0.541	ɛ	0.512	o	0.499
ɔ	0.478	o	0.658	e	0.616	i	0.549	e	0.559
ə	0.493	e	0.737	ə	0.832	ə	0.595	ə	0.583
f0 médio	0.452	F1 médio	0.555	F2 médio	0.438	F3 médio	0.460	Média vogais	0.472

Fonte: O Autor (2017).

O Quadro 5.5 resume os valores médios por vogais e por parâmetro acústico, nas células periféricas, na linha inferior e na coluna à direita, apresentando como resultado global dos dados o IV médio de todo conjunto, na célula inferior direita, com o valor de 0,472, que pode servir de referencial para julgar se um IV específico é um parâmetro aceitável ou não⁷⁴.

A expectativa inicial, baseada em pesquisas que buscaram ordenar sons mais apropriados para o reconhecimento de falantes (HÖFKER, 1977; GLENN E KLEINER, 1968; WOLF, 1972; HEUVEL, 1996; NOLAN, 2005; ROSE, 2007; AMINO et al., 2012), era de encontrar ordenamentos que mostrassem alguma homogeneidade, por som, entre os parâmetros acústicos estudados, uma vez que essas pesquisas apresentaram seus resultados por som. Assim, a coluna direita do Quadro 5.5, que apresenta os valores médios dos parâmetros por vogal, indica o ordenamento e mostra que as abertas [ɔ, a, ε] (respectivamente, 0,381, 0,426 e 0,472) e a posterior alta [u] (0,432) apresentaram valores abaixo ou igual à média global. Esse resultado, porém, será comentado ao final desta seção.

Com base no IV global, de 0,472, ao considerar os valores, expressos na linha inferior do Quadro 5.5, de 0,452 para f0, de 0,555 para F1, de 0,438 para F2 e de 0,460 para F3, fica evidente que F1 destoa, genericamente, como parâmetro aconselhável para a CL e que os outros três são parâmetros mais consistentes. Essa consideração permite avaliar os valores de F1 e determinar que para esse parâmetro apenas a vogal aberta, com valor de 0,452, deve ser tomada como elemento de comparação, uma vez que é a única vogal que possui valor inferior ao global para esse parâmetro.

Como a proposta deste estudo foi medir e observar também o comportamento de parâmetros acústicos isoladamente, os ordenamentos dos sons encontrados mostraram uma relação acústico-articulatória própria para cada parâmetro, a saber, f0, F1 e F2, os quais correlacionam-se, respectivamente, à fonte glótica, à dimensão vertical do trato e à dimensão horizontal do trato. Assim, a partir dos valores de IV já expostos no Quadro

⁷⁴ Supondo um IV de 0,500, isso significaria que o CV_{intra} é a metade do CV_{inter} e poderia intuitivamente ser considerado um valor de referência, o que se torna desnecessário uma vez que se calculou o IV global. Essa referência de 0,500 para o IV, porém será útil para discussões a serem apresentadas na seção 5.3, deste capítulo.

5.5, apresentam-se os dados ordenados para os três formantes e por f_0 separadamente, a partir do som vocálico, que é o elemento ordenado.

Quando considerado o ordenamento por f_0 no Quadro 5.5, não há uma vogal que se destaque em relação às demais, o que pode ser constatado pela proximidade de todos os valores em torno da média do parâmetro (0,452). Esse resultado é esperado, uma vez que o parâmetro refere-se à fonte glótica e não oferece correlato articulatório em relação à configuração do trato.

Em relação ao ordenamento dos sons para F1, as vogais das extremidades [a] e [i] ordenam-se acima das demais, seguidas de [ə], com um valor de IV que já se considera alto (0,511) como parâmetro para a CL (0,472), conforme se comentará a seguir. O resultado é esperado tendo em vista a relação do parâmetro com a dimensão vertical do trato. Na discussão retomaremos esse comportamento, principalmente no que se refere à vogal aberta, uma vez que não se entende essa relação como numérica simplesmente, mas antes como uma relação acústico-articulatória.

Para F2, ao se observar o ordenamento obtido, as posteriores [ɔ] (0,280) e [u] (0,324) e as baixas [ɛ] (0,389) e [a] (0,438) tendem a apresentar menor IV do que as demais.

Considerando o ordenamento por F3, as médias posteriores ([ɔ] com 0,325 e [o] com 0,427) e a aberta (0,389) tendem a apresentar menor IV do que as demais vogais. Esse parâmetro acústico representa uma relação mais complexa entre os elementos de ressonância e, assim, não se percebe uma relação acústico-articulatória evidente, assim como o obtido para F2.

Por fim, a estratégia de cálculo que culminou no IV como referencial para a consideração da relação de variabilidade individual ainda permite o ordenamento por vogal, observado na última coluna do Quadro 5.5, que apresenta o valor médio dos quatro parâmetros por vogal. Nesse caso, as vogais abertas ([ɔ] com 0,381; [a] com 0,426; [ɛ] com 0,472) e a alta posterior (0,432) apresentam IV abaixo ou igual ao global. Esse resultado, porém, deve ser observado com cuidado, pois carrega uma “homogeneização” dos contextos anteriores às vogais, ou seja, os valores apresentados carregam os efeitos coarticulatórios dos diferentes contextos anteriores à vogal medida. A seguir, será

descrita a estratégia para escrutinar os contextos anteriores com o intuito de se definir valores de referência que apresentem mais consistência para a utilização em CL.

5.1.1 Consideração sobre os efeitos do contexto anterior à vogal

Uma vez que os resultados para as vogais apresentam duas classes distintas de ambientes precedentes, fricativas desvozeadas e consoantes nasais, foi efetuada estratificação dos resultados conforme cada uma dessas classes, o que se denominará *primeiro passo*, e posteriormente pelas cinco consoantes do ambiente precedente, o que se denominará *segundo passo*, na tentativa de se observar alterações nos resultados de ordenamento causadas por efeitos coarticulatórios. Alterações são esperadas uma vez que a influência coarticulatória perseverativa da consoante sobre a vogal (SU ET AL. 1974; FOWLER 1992; FARNETANI, 1999) é variável de uma classe, as nasais, para outra, as fricativas desvozeadas.

Primeiramente serão apresentados os resultados de cada estratificação. Após a apresentação dos resultados, será explicada a estratégia de comparação dos ordenamentos entre cada um dos passos.

O resultado do primeiro passo está apresentado no Quadro 5.6 a seguir, onde se apresentam os valores dos IV, por classe, nasais ou fricativas desvozeadas, que antecede a vogal medida. Ao observar os ordenamentos resultantes da estratificação das medidas vocálicas segundo a classe do contexto anterior, percebe-se que os ordenamentos não se mantêm iguais ao que se observa no Quadro 5.5.

No que diz respeito aos valores médios por parâmetro, f_0 e F_3 seguem apresentando índices menores do que F_1 e F_2 , tanto para o conjunto dos contextos anteriores nasais (porção superior do Quadro 5.6) como para o conjunto dos contextos anteriores fricativos (porção inferior do Quadro 5.6).

Quadro 5.6 – Ordenamento por classe do contexto anterior

Índice de Variabilidade por classes do segmento anterior									
	f0		F1		F2		F3		Média Parametros
Nε	0.443	Na	0.496	Ne	0.479	Nɔ	0.467	Ni	0.514
No	0.445	Nɔ	0.571	Ni	0.486	No	0.501	Nɔ	0.538
Nɔ	0.459	Ni	0.574	Nε	0.511	Ni	0.541	Nε	0.544
Ni	0.467	Nə	0.581	Na	0.706	Ne	0.548	Ne	0.544
Nə	0.474	Nu	0.632	Nɔ	0.737	Na	0.568	Na	0.555
Nu	0.483	Nε	0.638	Nə	0.770	Nu	0.595	No	0.580
Na	0.497	Ne	0.693	No	0.794	Nə	0.599	Nə	0.588
Ne	0.501	No	0.720	Nu	0.811	Nε	0.637	Nu	0.609
f0 médio	0.470	F1 médio	0.605	F2 médio	0.632	F3 médio	0.552	Média	0.558
Fe	0.433	Fa	0.490	Fi	0.444	Fɔ	0.534	Fe	0.515
Fo	0.450	Fə	0.616	Fe	0.448	Fe	0.553	Fε	0.541
Fɔ	0.466	Fɔ	0.645	Fε	0.462	Fo	0.561	Fi	0.545
Fa	0.473	Fu	0.670	Fɔ	0.759	Fε	0.573	Fa	0.570
Fu	0.473	Fi	0.674	Fo	0.772	Fi	0.579	Fɔ	0.581
Fε	0.477	Fe	0.705	Fa	0.785	Fa	0.627	Fo	0.596
Fə	0.487	Fo	0.709	Fə	0.840	Fu	0.642	Fu	0.630
Fi	0.533	Fε	0.720	Fu	0.840	Fə	0.678	Fə	0.631
f0 médio	0.472	F1 médio	0.645	F2 médio	0.619	F3 médio	0.590	Média	0.573

Fonte: O Autor (2017).

Para prosseguir com o processo de análise dos efeitos dos contextos anteriores às vogais, o segundo passo é a estratificação de cada uma das duas classes apresentadas no Quadro 5.6, a das nasais na sua metade superior e a das fricativas na metade inferior. O resultado dessa nova estratificação será apresentado nos dois próximos quadros. O Quadro 5.7 a seguir estratifica os valores do contexto anterior da classe nasal pelas duas consoantes que compõem a classe ([m] e [n]).

Quadro 5.7 – Ordenamento por consoante contexto anterior - nasais

Índice de Variabilidade por segmento anterior - nasais									
	f0		F1		F2		F3		Média Parametros
mɔ	0.419	ma	0.488	me	0.413	mɔ	0.427	mi	0.491
mɛ	0.428	mə	0.549	mi	0.478	ma	0.462	ma	0.500
mə	0.436	mi	0.557	mɛ	0.480	mi	0.472	mɔ	0.501
mo	0.441	mɔ	0.567	ma	0.538	mo	0.491	me	0.503
mu	0.466	mɛ	0.613	mə	0.556	me	0.507	mə	0.508
mi	0.468	mu	0.618	mu	0.647	mə	0.512	mɛ	0.517
me	0.475	me	0.690	mɔ	0.671	mu	0.586	mo	0.570
ma	0.517	mo	0.703	mo	0.775	mɛ	0.596	mu	0.570
f0 médio	0.454	F1 médio	0.590	F2 médio	0.549	F3 médio	0.501	Média	0.519
no	0.433	na	0.451	ni	0.446	nɔ	0.423	nɔ	0.486
ni	0.435	nɔ	0.504	ne	0.494	no	0.444	ni	0.488
na	0.443	ni	0.557	nɛ	0.500	ni	0.539	na	0.511
nɛ	0.447	nə	0.562	nɔ	0.554	nu	0.545	no	0.530
nu	0.481	nu	0.585	nə	0.572	ne	0.555	nɛ	0.535
nɔ	0.482	nɛ	0.597	na	0.594	nə	0.571	nə	0.546
nə	0.489	ne	0.653	no	0.651	na	0.602	ne	0.547
ne	0.511	no	0.691	nu	0.693	nɛ	0.636	nu	0.566
f0 médio	0.464	F1 médio	0.566	F2 médio	0.552	F3 médio	0.530	Média	0.525

Fonte: O Autor (2017).

Ainda compondo o segundo passo de estratificação, o Quadro 5.8 a seguir apresenta os valores estratificados da classe das fricativas desvozeadas em seus três sons componentes ([f], [s] e [ʃ]).

A questão que se coloca é como comparar os efeitos coarticulatórios dos agrupamentos prévios a cada um dos passos e os resultados das respectivas estratificações. Uma maneira de se observar os efeitos coarticulatórios pode ser através da avaliação do quanto os ordenamentos se alteram num primeiro momento, quando se estratifica o primeiro agrupamento, apresentado no Quadro 5.5, em duas classes, apresentadas no Quadro 5.6, e, num segundo momento, quando se estratifica cada classe em seus sons componentes, estratificações apresentadas nos Quadros 5.7 para as nasais e no Quadro 5.8, para as fricativas.

Quadro 5.8 – Ordenamento por consoante contexto anterior – fricativas desvozeadas

Índice de Variabilidade por segmento anterior - fricativas									
	f0		F1		F2		F3		Média Parametros
fe	0.420	fa	0.480	fe	0.389	fɔ	0.408	fe	0.470
fɛ	0.426	fə	0.583	fɛ	0.390	fa	0.430	fa	0.477
fa	0.434	fu	0.629	fi	0.428	fi	0.446	fɛ	0.483
fə	0.448	fɔ	0.638	fo	0.472	fo	0.474	fi	0.495
fɔ	0.457	fi	0.640	fɔ	0.573	fe	0.479	fɔ	0.503
fu	0.486	fo	0.664	fa	0.600	fɛ	0.510	fo	0.517
fo	0.498	fe	0.682	fə	0.612	fə	0.512	fə	0.531
fi	0.518	fɛ	0.709	fu	0.684	fu	0.523	fu	0.569
f0 médio	0.458	F1 médio	0.620	F2 médio	0.497	F3 médio	0.469	Média	0.504
se	0.410	sa	0.483	si	0.359	so	0.422	si	0.464
so	0.420	sə	0.565	se	0.411	sɔ	0.470	se	0.474
sɔ	0.423	sɔ	0.600	sɛ	0.424	si	0.473	sɔ	0.512
sə	0.441	sɪ	0.637	sɔ	0.605	su	0.481	so	0.512
su	0.456	su	0.648	sə	0.652	se	0.491	sɛ	0.513
si	0.468	so	0.649	sa	0.654	sɛ	0.516	sa	0.530
sa	0.495	se	0.649	so	0.659	sa	0.518	sə	0.551
sɛ	0.496	sɛ	0.677	su	0.722	sə	0.593	su	0.556
f0 médio	0.449	F1 médio	0.607	F2 médio	0.526	F3 médio	0.491	Média	0.512
ʃo	0.394	ʃa	0.436	ʃi	0.378	ʃi	0.436	ʃi	0.471
ʃu	0.410	ʃɔ	0.537	ʃe	0.425	ʃo	0.439	ʃe	0.487
ʃɛ	0.418	ʃə	0.597	ʃɛ	0.431	ʃɔ	0.445	ʃɛ	0.488
ʃe	0.437	ʃu	0.607	ʃɔ	0.569	ʃe	0.482	ʃɔ	0.500
ʃa	0.455	ʃi	0.617	ʃu	0.636	ʃɛ	0.500	ʃo	0.508
ʃɔ	0.467	ʃo	0.645	ʃo	0.666	ʃə	0.512	ʃa	0.511
ʃə	0.483	ʃe	0.671	ʃa	0.691	ʃa	0.531	ʃu	0.537
ʃi	0.516	ʃɛ	0.672	ʃə	0.732	ʃu	0.556	ʃə	0.566
f0 médio	0.444	F1 médio	0.587	F2 médio	0.535	F3 médio	0.484	Média	0.507

Fonte: O Autor (2017).

Para avaliar a coarticulação e a conseqüente alteração nos ordenamentos das vogais, adota-se uma estratégia não-paramétrica em que são atribuídos postos, de 1 a 8, às vogais apresentadas nos quadros, conforme suas ordens, do menor para o maior, ou de cima para baixo. Assim, como exemplo, para o Quadro 5.5 já apresentado, [a] recebeu os valores 1 (primeiro posto) para F1, 4 (quarto posto) para F2 e 2 (segundo posto) para F3; [ɛ] recebeu 6 para F1, 3, para F2 e 6 para F3; [e] recebeu 8 para F1, 7 para F2 e 5 para F3 e assim sucessivamente para todos os sons e para todos os quadros. Não foram considerados os valores da primeira coluna, f_0 , nem o da última, valor médio dos parâmetros, pois não são afetados diretamente pela coarticulação, como ocorre com F1, F2 e F3. A estratégia de comparação consiste em efetuar a diferença entre os postos para cada vogal, ou seja, 24 diferenças para cada comparação, e se obter a média dos valores absolutos das diferenças, que representam o quanto um som mudou de posto de um quadro para o outro. Esse valor é a mudança média de postos ocasionada pela estratificação e pode ser interpretado como indicador do efeito coarticulatório. Quanto maior o valor, mais mudanças de postos e, conseqüentemente, mais coarticulação ocasionada por aquele agrupamento.

Para a comparação do primeiro passo de estratificação, são tomados os valores de ordem referentes ao Quadro 5.5 e cotejados com os valores de ordem do Quadro 5.6, que apresenta as estratificações discriminadas por classe do contexto anterior. Para a comparação do segundo passo, são tomados os valores de ordem do Quadro 5.6 para cada classe, e comparados com os valores de ordem por consoante do contexto anterior, nos Quadros 5.7 e 5.8.

No primeiro passo, ao se comparar o conjunto não estratificado com as classes estratificadas, obtêm-se para a classe nasal o resultado médio de 1,83 alterações de postos e para a classe fricativa desvozeada, de 1,92 alterações. O resultado torna-se relevante com comparação desses valores com àqueles obtidos no segundo passo. No segundo passo, a estratificação da classe nasal para a nasal bilabial resultou na alteração média de 0,83 e para a nasal alveolar em 0,50. Com relação às fricativas desvozeadas, a alteração média foi de 1 para a labiodental; 0,91 para a alveolar e de 1 para a palatal.

Os valores podem ser interpretados no sentido de que o ordenamento das vogais muda consideravelmente no primeiro passo em relação ao segundo e que no segundo

passo a mudança é menor para as nasais do que para as fricativas. A nasal alveolar sofre menos coarticulação do que a nasal bilabial.

Encerrando a seção envolvendo os IV das vogais, resta evidente que os IV devem ser tomados com atenção para se definir ordenamentos, sendo tanto mais confiáveis quanto mais homogêneos forem os agrupamentos no que dizem respeito aos contextos anteriores. Essa noção será discutida mais profundamente na seção 5.3.1. O valor de referência para o IV em torno de 0,500 (ver nota 71) como marco superior parece ser adequado quando se observam os valores médios dos quadros mais estratificados.

5.1.2 Espaço acústico

Uma das propostas da pesquisa em tela é avaliar o espaço acústico, além das medidas formânticas vocálicas. Na busca por uma referência concreta sobre fenômenos acústico-articulatórios, a partir dos dados considerados nesta pesquisa, as informações de espaço acústico como definidas por Kent e Kim (2003), conforme seção 2.3.1, Capítulo 2, pensadas para dar conta da análise de dados de pacientes disártricos, bem como as distâncias euclidianas entre os pontos extremos do triângulo vocálico, {ai}, {au} e {ui}, podem ser avaliadas analogamente ao que já foi feito para os formantes vocálicos e frequência fundamental.

Os dados de espaço acústico, isto é, os valores de área e de distâncias entre os vértices de cada triângulo vocálico, podem ser pareados, uma vez que se obteve os CV_{inter} e CV_{intra} e, conseqüentemente, o IV.

Os valores de média, desvio padrão, CV_{inter} , CV_{intra} e IV estão dispostos em cinco quadrantes, um para cada valor, e três agrupamentos de linhas, expostos no Quadro 5.9 a seguir. Na primeira linha aparecem os valores globais encontrados. Na segunda e terceira linhas, aparecem os valores estratificados conforme a classe do contexto anterior, nasais ou fricativas, às vogais vértices, [i, a, u], do triângulo. Nas cinco linhas inferiores aparecem os valores estratificados conforme os sons dos contextos anteriores à vogal, [f, s, ʃ, m, n]. Nas primeiras colunas de cada quadrante aparecem os valores referentes ao VSA, seguidos dos valores das DE de {ai}, {iu} e {au}.

Quadro 5.9: Resultados gerais de média e variabilidade do VSA e distâncias euclidianas (desvio padrão, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas aparece na segunda coluna.

Contexto	N	Média				Desv				CV Interindividual			
		VSA	ai	iu	au	VSA	ai	iu	au	VSA	ai	iu	au
TODOS	2180	335226	940	1404	765	115791	157	201	156	34.54%	16.68%	14.32%	20.41%
Fric	1308	307875	923	1352	715	97117	152	194	135	31.54%	16.46%	14.34%	18.91%
Nasal	872	376222	965	1483	841	128760	161	186	155	34.22%	16.64%	12.52%	18.42%
f_	436	328873	975	1442	741	90887	142	157	125	27.64%	14.54%	10.89%	16.83%
s_	436	300119	916	1306	690	95827	142	193	132	31.93%	15.48%	14.74%	19.14%
ʃ_	436	294614	879	1308	713	101075	156	198	143	34.31%	17.79%	15.11%	20.10%
m_	436	385377	995	1519	843	127139	144	162	154	32.99%	14.43%	10.64%	18.32%
n_	436	367067	936	1447	839	129862	171	201	156	35.38%	18.29%	13.87%	18.55%
Contexto	N	CV Intraindividual				Índice de Variabilidade				VSA	ai	iu	au
		VSA	ai	iu	au	VSA	ai	iu	au				
TODOS	2180	12.99%	6.94%	6.31%	9.89%	0.376	0.416	0.440	0.485				
Fric	1308	11.93%	6.49%	6.25%	9.81%	0.378	0.394	0.436	0.519				
Nasal	872	15.00%	7.75%	6.39%	10.01%	0.438	0.466	0.511	0.543				
f_	436	11.45%	6.29%	5.66%	9.09%	0.414	0.432	0.519	0.540				
s_	436	12.11%	6.11%	6.56%	10.04%	0.379	0.395	0.445	0.525				
ʃ_	436	12.26%	7.15%	6.64%	10.40%	0.357	0.402	0.439	0.517				
m_	436	15.27%	7.03%	5.71%	10.37%	0.463	0.487	0.537	0.566				
n_	436	14.73%	8.63%	7.26%	9.67%	0.416	0.472	0.523	0.521				

Fonte: O Autor (2017).

É possível adotar-se como referência o valor de IV estabelecido na seção anterior, de 0,500 como marco superior para o IV. Assim, considera-se para a definição de consistência dos quatro parâmetros geométricos aqui expostos a média das cinco categorias de sons, definidas pelo contexto anterior à vogal, ou seja, as cinco linhas inferiores do quadrante do IV no Quadro 5.9. Como resultado obteve-se o valor médio de 0,406 para o VSA e o valor médio de 0,438 para a DE_{ai} como aparentemente adequados para a aplicação na CL. Os valores médios para DE_{iu} e para DE_{au} , de 0,493 e 0,534 respectivamente, poderiam ser considerados como menos consistentes por estarem muito próximos ao valor de referência.

Finalizada a descrição dos resultados obtidos para as vogais, passa-se, na seção 5.2, à descrição dos resultados obtidos para as consoantes nasais.

5.2 CONSOANTES NASAIS

A seguir serão apresentados os resultados obtidos para as nasais [m] e [n], seguidas pelas vogais [a], [i] e [u], confrontando-os com os resultados obtidos por Seara (2000). O Quadro 5.10 a seguir apresenta um sumário dos resultados obtidos no presente estudo, a partir dos valores de média, desvio padrão CV do corpus relativos a f0, F1 e F2. O objetivo do sumário dos resultados é restringir os valores globais obtidos àqueles que serão objeto de comparação com os resultados de Seara (2000).

Quadro 5.10 – Estatística descritiva resumida das consoantes nasais.

	Média			Desv			CV do Corpus		
	f0	F1	F2	f0	F1	F2	f0	F1	F2
ma	104	235	1055	14	46	145	13.2%	19.4%	13.8%
mi	115	237	1330	16	22	222	14.1%	9.1%	16.7%
mu	110	232	1011	17	26	210	15.7%	11.3%	20.8%
na	104	267	1315	17	41	196	16.3%	15.5%	14.9%
ni	118	264	1142	20	31	265	16.9%	11.6%	23.2%
nu	114	253	1273	17	34	197	14.4%	13.3%	15.5%

Fonte: O Autor (2017)

Os resultados relativos a f0 apresentam comportamento esperado, pois a coarticulação da consoante nasal com a vogal seguinte ocasionou valores mais altos para os sons seguidos das vogais altas do que aqueles seguidos pela vogal baixa. Dos indicadores de variabilidade, o desvio padrão não apresenta padrão fora do esperado e o CV chama a atenção por apresentar valores, com exceção da sílaba [ma], maiores do que os análogos para F1.

Com relação aos valores médios de F1, percebe-se a maior consistência dos resultados, que permanecem próximos das faixas que a literatura refere, com valores maiores para [n] do que para [m]. Como já foi referido, para F1 chama a atenção os valores de CV menores do que os de f0. Ambos indicadores de variabilidade apontam valores maiores para [m] e [n] quando seguidos por [a] do que quando seguidos pelas vogais fechadas.

Para os valores de F2, os resultados médios não apontam para uma faixa de frequência preferencial e os indicadores de variabilidade, particularmente o CV, apontam para maior variabilidade desse parâmetro em relação a f0 e F1.

O Quadro 5.11 a seguir apresenta as composições dos achados da pesquisa em tela com os resultados de Seara (2000), de maneira que os valores apresentados nas linhas adjacentes, formem um par para comparação. Assim, nas linhas sombreadas, na posição superior do par em cada quadrante, aparecem os valores de média, desvio padrão e coeficiente de variação encontrados no trabalho aqui apresentado e, nas linhas sem somreamento, na posição inferior do par, aparecem os valores análogos obtidos por Seara (2000). Os parâmetros em cada quadro são os mesmos apresentados por Seara (2000) com exceção de F3, uma vez que no levantamento aqui realizado esse parâmetro foi excluído, como será explicado posteriormente. Os dados estão separados de forma análoga ao apresentado por Seara (2000), organizados segundo a consoante analisada – [m] ou [n] – e a vogal que segue – [a], [i] ou [u].

Quadro 5.11 – Comparação com os achados de Seara (2000).

[m] antes de [a]				[n] antes de [a]			
	f0	F1	F2		f0	F1	F2
Média (Autor)	104	235	1055	Média (Autor)	104	267	1315
Média (Seara, 2000)	127	261	840	Média (Seara, 2000)	124	260	1265
Desv. Padr. (Autor)	14	46	145	Desv. Padr. (Autor)	17	41	196
Desv. Padr. (Seara, 2000)	16	24	81	Desv. Padr. (Seara, 2000)	15	24	133
Coefic. Var. (Autor)	13.18%	19.40%	13.79%	Coefic. Var. (Autor)	16.27%	15.45%	14.92%
Coefic. Var. (Seara, 2000)	12.30%	9.38%	9.65%	Coefic. Var. (Seara, 2000)	11.47%	9.24%	10.51%
[m] antes de [i]				[n] antes de [i]			
	f0	F1	F2		f0	F1	F2
Média (Autor)	115	237	1330	Média (Autor)	118	264	1142
Média (Seara, 2000)	123	258	919	Média (Seara, 2000)	134	247	1094
Desv. Padr. (Autor)	16	22	222	Desv. Padr. (Autor)	20	31	265
Desv. Padr. (Seara, 2000)	10	23	161	Desv. Padr. (Seara, 2000)	26	37	253
Coefic. Var. (Autor)	14.07%	9.13%	16.69%	Coefic. Var. (Autor)	16.90%	11.64%	23.24%
Coefic. Var. (Seara, 2000)	8.47%	8.73%	17.51%	Coefic. Var. (Seara, 2000)	19.17%	14.94%	23.10%
[m] antes de [u]				[n] antes de [u]			
	f0	F1	F2		f0	F1	F2
Média (Autor)	110	232	1011	Média (Autor)	114	253	1273
Média (Seara, 2000)	130	262	678	Média (Seara, 2000)	132	273	981
Desv. Padr. (Autor)	17	26	210	Desv. Padr. (Autor)	17	34	197
Desv. Padr. (Seara, 2000)	13	31	97	Desv. Padr. (Seara, 2000)	14	22	141
Coefic. Var. (Autor)	15.65%	11.25%	20.81%	Coefic. Var. (Autor)	14.43%	13.27%	15.45%
Coefic. Var. (Seara, 2000)	9.70%	11.79%	14.26%	Coefic. Var. (Seara, 2000)	10.79%	8.04%	14.38%

Fonte: O Autor (2017). Adaptação a partir de Seara (2000).

A comparação dos resultados por formante aponta que para os valores de F1 as diferenças entre os valores médios indicam consistência entre os achados deste trabalho e os de Seara (2000) no que se refere a [n], uma vez que em ambos trabalhos as médias estão na casa dos 250-300 Hz referidos pela literatura (STEVENS, 1998, p. 489, FANT, 1970), com exceção da sílaba [ni] em Seara (2000), com valor de 247 Hz, próximo da faixa mencionada. Para [m], chama a atenção que em Seara (2000) os valores médios foram próximos aos obtidos para [n], enquanto que no estudo em tela os valores de [m] foram menores do que os de [n].

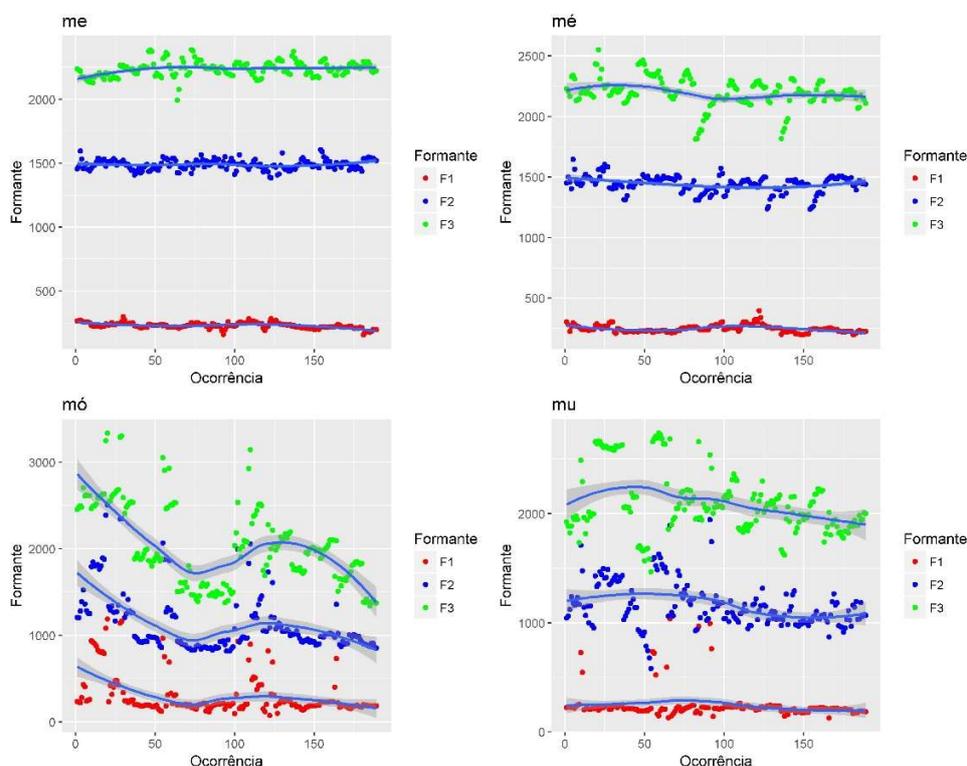
Para F2 não há correspondência evidente entre os seis pares comparados. A própria variabilidade interindividual grande, conforme aponta Seara (2000, p. 147-148), pode servir de explicação, apoiada no fato de que a cavidade nasal confere uma característica muito individual, específica do falante. Para f0, as seis médias encontradas por Seara são mais altas do que as obtidas neste estudo. Os valores de desvio padrão e CV referentes a f0 não apontam discrepâncias na variabilidade das respectivas amostras, o que poderia explicar a diferença entre os valores médios. Assim, pode-se concluir que a diferença entre as médias seja resultante da composição das amostras comparadas, já que Seara (2000) considera quatro participantes para seu estudo e esta pesquisa considera vinte participantes.

Inicialmente os valores globais de variabilidade para as consoantes nasais foram altos a ponto de evidentemente não apresentarem valores de CV favoráveis em relação aos obtidos para as vogais. Essa variabilidade aparentemente está vinculada a uma característica de ênfase articulatória particular dos indivíduos, o que ocasiona diferenças de intensidade e conseqüentemente prejudica a estratégia adotada para a leitura dos formantes nasais. Este estudo, apesar de ter-se baseado em estratégia semelhante à proposta por Seara (2000), no que diz respeito a extrapolação do número de coeficientes do LPC, conforme descrito na seção 2.4, diferiu, em função do número maior de indivíduos e de dados medidos, no estabelecimento dos formantes nasais. Naquele estudo, para cada ocorrência da nasal, era dedicada uma análise minuciosa do dado acústico através da curva de LPC, que era confrontada manualmente com o espectro FFT para a determinação do valor dos formantes. No caso do presente estudo, não era viável tal análise minuciosa em função do tamanho da amostra e a estratégia adotada

aqui foi a medição através do monitoramento pelo espectrograma, conforme descrito em 4.3.3, na tentativa de acelerar o processo de medição, o que, em ensaio prévio realizado com um dos sujeitos, havia apresentado resultados satisfatórios.

O efeito da estratégia de medição foi de que, em algumas ocorrências, uma faixa de frequência era interpretada como um formante e, em outras, era interpretada como um formante vizinho, superior ou inferior, apesar de sempre se obter o acoplamento do LPC ao espectrograma. Como exemplo do fenômeno ocorrido, na Figura 5.2 a seguir apresentam-se os resultados em diagramas de dispersão⁷⁵ de quatro agrupamentos, referentes às sílabas [me, mɛ, mɔ, mu] do participante EGS.

Figura 5.2 – Dispersão das medidas de consoantes nasais [me, mɛ, mɔ⁷⁶, mu] – Participante EGS



Fonte: O Autor (2017).

⁷⁵ Os diagramas de dispersão foram produzidos em ambiente estatístico do pacote R, através do R studio, com o auxílio do professor Ms. Filipe Zabala (PUCRS). Cada gráfico foi composto pelo alinhamento de todas as ocorrências, nos diversos turnos e dias, de uma mesma sílaba para o sujeito indicado. Assim, sua abscissa indica apenas a ordem em que ocorreram as medidas.

⁷⁶ O pacote R não lida adequadamente com caracteres diferenciados, como o são os símbolos do IPA, por isso, nos gráficos produzidos nesse ambiente utilizou-se caracteres ortográficos acentuados para representar as vogais médias abertas e o *schwa*.

A Figura 5.2 apresenta as distribuições das medidas para F1, em vermelho, para F2, em azul, e para F3, em verde. As cores permitem perceber que, por vezes, a identificação dos formantes não ocorreu de maneira clara, o que se pode constatar pela presença de ocorrências de uma cor na faixa que seria esperada para outra cor. Cada ponto colorido no gráfico refere-se a um ponto de medição do som.

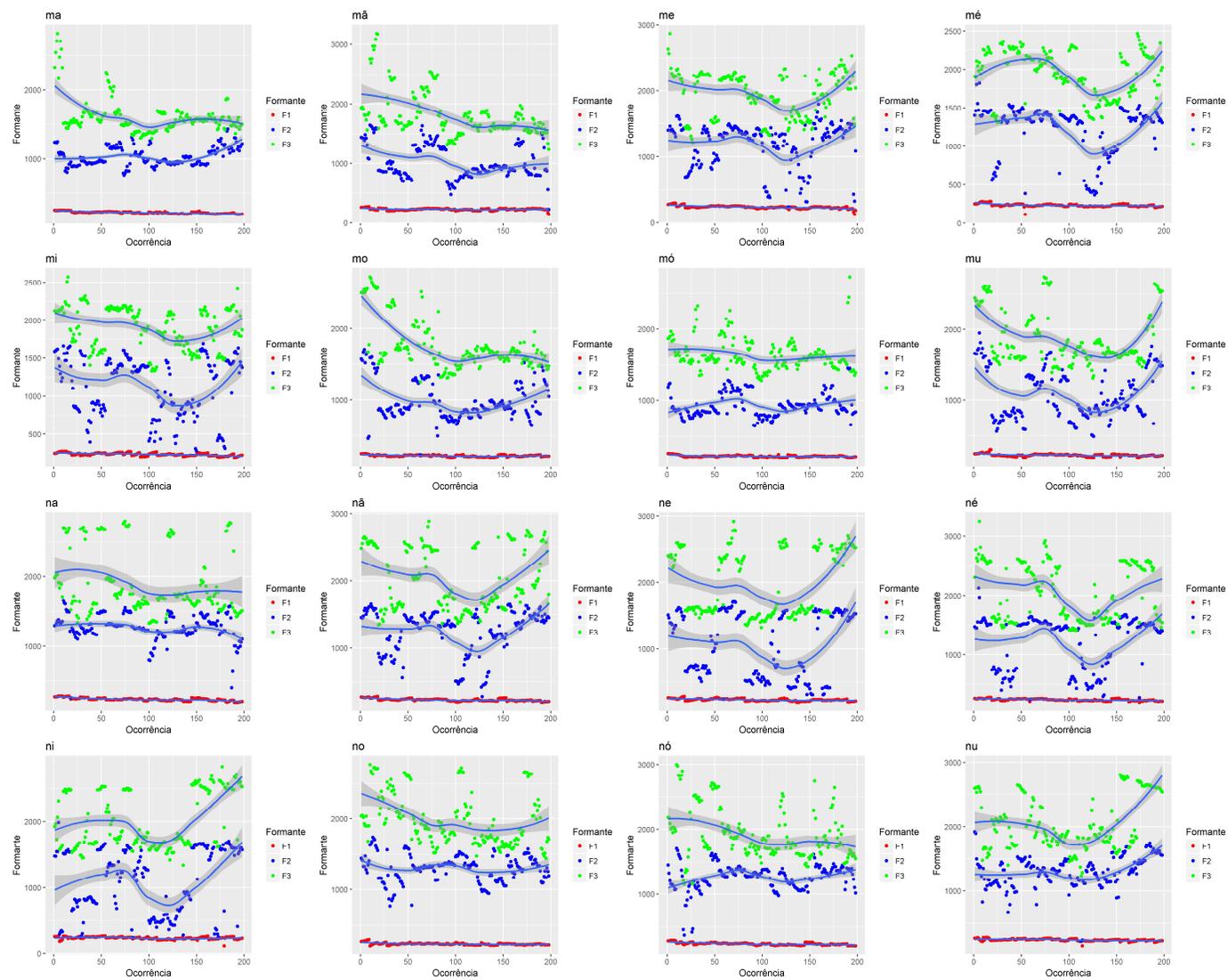
Os dois gráficos na porção superior da Figura 5.2 são exemplos de contextos de menor variabilidade, o que se pretendia obter com a estratégia de medição adotada e já descrita na seção 4.3.2, Capítulo 4. Os dois gráficos na porção inferior da Figura 5.2 são exemplos de dispersões excessivas e interpretações equivocadas dos formantes, notadas pela intrusão de pontos de uma cor na faixa concernente à cor de um formante superior ou inferior ao esperado. Observou-se que a imprecisão nas leituras do LPC ocasionada pela variabilidade individual da ressonância nasal foi excessiva em muitos casos. Como tentativa de minimizar os efeitos das imprecisões das leituras, buscou-se selecionar os contextos em que a variabilidade foi menor. A seleção tomou por referência justamente as distribuições recém descritas. Em uma primeira triagem, adotou-se um critério perceptivo para descartar aqueles contextos e participantes que evidentemente não apresentavam condições de variabilidade para compor o conjunto de dados bem como selecionar aqueles contextos em que evidentemente a variabilidade permitia a composição do conjunto de dados. A Figura 5.3 adiante apresenta o conjunto das dezesseis sílabas ([ma, mə, me, mɛ, mi, mo, mɔ, mu, na, nə, ne, nɛ, ni, no, nɔ, nu]) produzidas pelo participante MFS, que ilustra o pior dos casos no que se refere à dispersão excessiva dos valores, a saber, aquela em que nenhum dos contextos silábicos foi selecionado para a composição de um novo conjunto de dados.

Em seguida apresenta-se, na Figura 5.4, as dezesseis sílabas do sujeito AFRB, considerado, para as consoantes nasais, como o melhor dos casos, ou seja, em que todos os contextos apresentaram dispersão reduzida, servindo assim a composição de um novo conjunto. Para esse participante, todos os contextos silábicos foram selecionados para a composição do novo conjunto dos dados.

Os casos restantes⁷⁷ foram selecionados tomando-se uma linha de corte em 20% referenciada no CV da primeira amostragem. O valor para linha de corte baseia-se nas faixas consideradas por Seara (2000), porém revisadas neste estudo a partir de um critério mais rígido, em função da natureza diferenciada da pesquisa em tela, a qual busca justamente a invariabilidade individual, ao passo que a autora busca a invariabilidade linguística, que permite um critério menos rígido. Além disso, excluíram-se os valores de F3 que invariavelmente se apresentaram excessivamente dispersos.

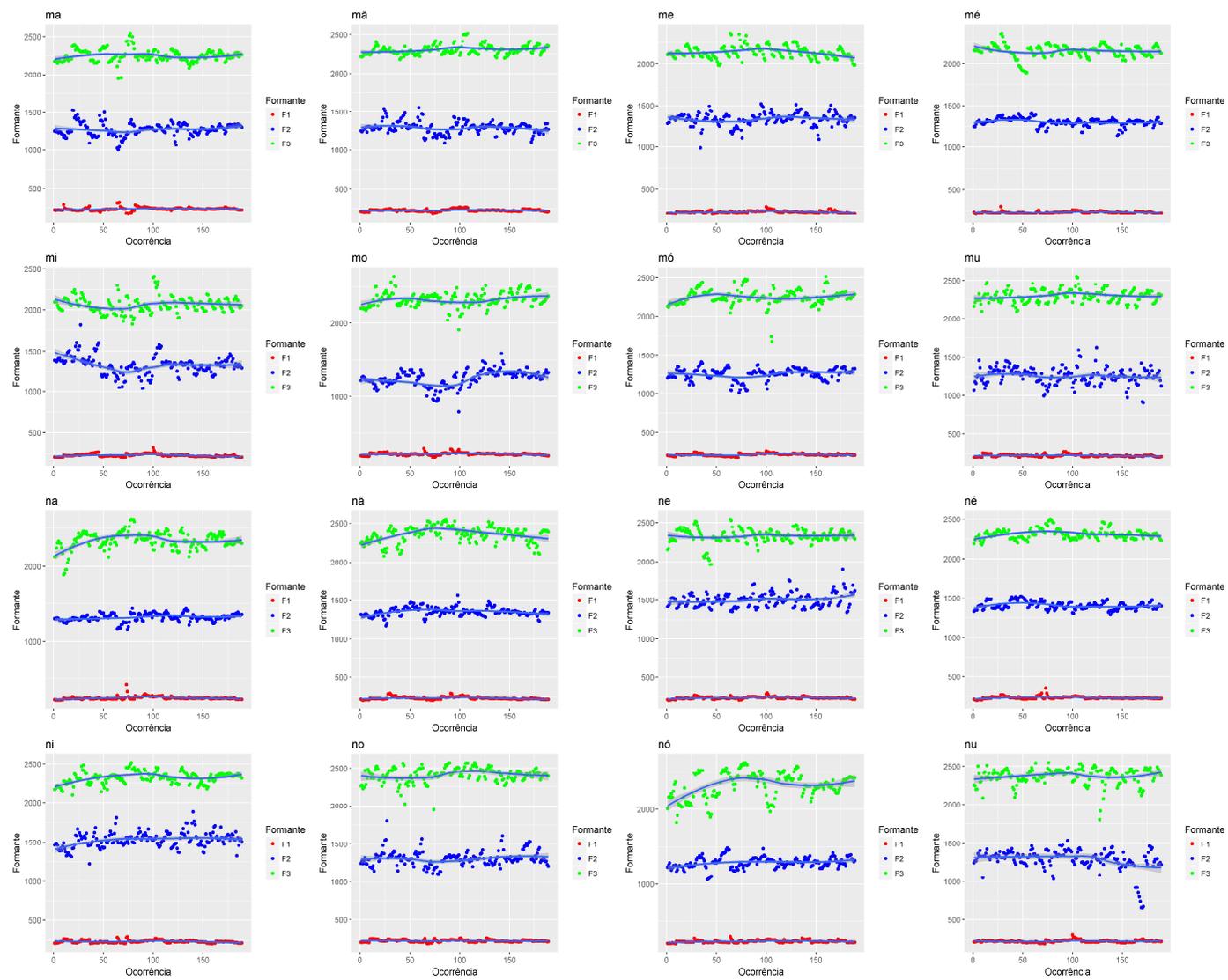
⁷⁷ Os gráficos de dispersão dos demais participantes da pesquisa em tela encontram-se no Apêndice D.

Figura 5.3 – Dispersão das medidas de consoantes nasais – Participante MFS



Fonte: O Autor (2017).

Figura 5.4 – Dispersão das medidas de consoantes nasais – Participante AFRB



Fonte: O Autor (2017).

Após os procedimentos de triagem em busca dos contextos com menores valores de variabilidade, o corpus resultante sobre o qual calculou-se novamente a estatística descritiva das consoantes nasais permaneceu com o número de “participantes por sílaba” conforme apresentado no Quadro 5.12 a seguir. As sílabas aparecem ordenadas por número de participantes em ordem decrescente de cima para baixo e da esquerda para a direita.

Quadro 5.12 – Número de participantes por contexto selecionado

Sílaba	N inform	Sílaba	N inform
ma	18	nu	14
nó	17	mu	10
mé	16	nã	10
mã	15	mi	9
mó	15	ne	8
mo	14	me	7
na	14	né	6
no	14	ni	4

Fonte: O Autor (2017).

Inicialmente apresentam-se os resultados em analogia aos expostos para as vogais. O Quadro 5.13 a seguir apresenta o primeiro resultado, expresso pelos valores de tendência central e de variabilidade, para as consoantes nasais, a partir da consideração de todos os participantes.

Quadro 5.13 – Resultados gerais para as consoantes nasais, estratificadas em bilabial e alveolar, expressos pela média e variabilidade (desvio padrão, amplitude, mínimo, máximo, coeficiente de variação interindividual, coeficiente de variação intraindividual e Índice de Variabilidade). O número de medidas aparece na segunda coluna.

	N	Média			Desv			Mínimo			Máximo		
		f0	F1	F2	f0	F1	F2	f0	F1	F2	f0	F1	F2
m	12558	110	246	1210	16	33	204	50	66	264	167	1160	2413
n	12558	111	247	1221	16	32	208	51	84	220	166	548	2608
		Amplitude			CV do Corpus			Coeficiente Var			Índice de Variabilidade		
		f0	F1	F2	f0	F1	F2	f0	F1	F2			
m	12558	116	1094	2149	14.8%	13.3%	16.9%	7.3%	9.3%	14.0%	0.496	0.704	0.830
n	12558	114	464	2388	14.9%	12.8%	17.1%	7.5%	8.6%	16.3%	0.504	0.670	0.954

Fonte: O Autor (2017).

No Quadro 5.14 a seguir apresentam-se os resultados para os agrupamentos de dados determinados pela vogal que segue a consoante.

Quadro 5.14 – Resultados gerais para as consoantes nasais, estratificadas por sílaba, expressos pela média e variabilidade (desvio padrão, mínimo, máximo, amplitude, coeficiente de variação interindividual, coeficiente de variação intraindividual e a índice de variabilidade).

	N	Média			Desv			Mínimo			Máximo			
		f0	F1	F2	f0	F1	F2	f0	F1	F2	f0	F1	F2	
ma	1182	104	235	1055	14	46	145	56	125	396	145	1160	2275	ma
mã	987	107	230	1084	16	31	162	52	95	504	155	323	1692	mã
me	459	114	240	1317	14	22	228	92	176	351	155	304	1880	me
mé	1050	107	245	1303	17	32	175	51	169	264	147	396	1917	mé
mi	579	115	237	1330	16	22	222	61	186	632	167	334	2314	mi
mo	918	109	229	1004	17	27	172	50	139	685	164	369	1635	mo
mó	990	105	227	1013	17	40	147	52	78	711	147	737	2016	mó
mu	663	110	232	1011	17	26	210	82	176	342	165	365	1757	mu
na	921	104	267	1315	17	41	196	52	152	379	151	403	1955	na
nã	660	114	252	1321	16	33	208	86	153	577	161	401	1811	nã
ne	522	111	256	1342	16	33	286	59	140	292	156	358	2036	ne
né	402	112	262	1305	15	33	291	58	205	604	146	356	1914	né
ni	264	118	264	1142	20	31	265	86	212	248	160	347	1882	ni
no	921	112	252	1279	16	32	184	52	182	453	152	489	1861	no
nó	1119	107	258	1273	16	41	174	51	137	288	150	546	1840	nó
nu	921	114	253	1273	17	34	197	61	176	482	160	416	1886	nu
		Amplitude			CV do Corpus			CV Médio individual			Índice de Variabilidade			
		f0	F1	F2	f0	F1	F2	f0	F1	F2	f0	F1	F2	
ma	1182	89	1035	1879	13.2%	19.4%	13.8%	7.3%	10.7%	8.3%	0.554	0.550	0.601	ma
mã	987	104	228	1188	15.4%	13.3%	14.9%	7.5%	7.1%	9.5%	0.486	0.531	0.633	mã
me	459	63	128	1529	12.6%	9.2%	17.3%	6.3%	6.7%	8.9%	0.496	0.737	0.512	me
mé	1050	97	227	1653	15.6%	13.1%	13.5%	8.4%	7.0%	9.2%	0.537	0.535	0.681	mé
mi	579	105	148	1682	14.1%	9.1%	16.7%	7.7%	6.9%	9.3%	0.547	0.758	0.555	mi
mo	918	113	230	950	15.7%	11.7%	17.2%	8.0%	6.7%	10.5%	0.512	0.567	0.613	mo
mó	990	95	659	1305	16.0%	17.5%	14.5%	7.5%	10.0%	9.0%	0.469	0.571	0.618	mó
mu	663	83	189	1415	15.7%	11.3%	20.8%	5.7%	6.1%	10.7%	0.361	0.540	0.516	mu
na	921	100	250	1576	16.3%	15.5%	14.9%	9.3%	9.1%	7.8%	0.571	0.587	0.526	na
nã	660	74	248	1234	13.7%	13.1%	15.7%	6.9%	7.7%	8.1%	0.505	0.586	0.517	nã
ne	522	97	219	1744	14.6%	13.0%	21.3%	6.5%	6.9%	10.0%	0.447	0.529	0.467	ne
né	402	89	151	1310	13.5%	12.6%	22.3%	6.4%	6.5%	10.7%	0.474	0.518	0.480	né
ni	264	74	135	1634	16.9%	11.6%	23.2%	5.7%	5.8%	8.6%	0.339	0.500	0.372	ni
no	921	100	307	1408	14.7%	12.8%	14.4%	7.9%	7.4%	9.1%	0.540	0.574	0.635	no
nó	1119	98	408	1551	14.7%	15.7%	13.7%	7.2%	8.8%	8.8%	0.490	0.561	0.642	nó
nu	921	100	241	1405	14.4%	13.3%	15.5%	7.1%	8.3%	10.0%	0.494	0.623	0.647	nu

Fonte: O Autor (2017).

Apesar da estratégia adotada para a minimização dos efeitos de dispersão dos valores lidos para as consoantes nasais, a maior parte dos valores de IV de F1 e F2 ficaram acima do referencial genérico de 0,500 (ver nota 72). As três exceções referem-se aos valores de IV de F2 (ne: 0,467; né: 0,480 e ni: 0,372) e, apesar de estarem abaixo do referencial genérico, comportam-se dessa maneira devido aos valores altos de seus CV_{inter} (ne: 21,3%; né: 22,3% e ni: 23,2%). Diante do exposto, conclui-se que os parâmetros não se apresentam com indicativos de variabilidade individual para a utilização na CL. Esse achado será retomado na discussão.

5.3 DISCUSSÃO

Para o início da seção de discussão dos resultados, serão feitas algumas considerações sobre as comparações realizadas com os demais trabalhos descritivos do PB, tanto para as vogais quanto para as nasais. Com relação aos trabalhos a respeito de vogais foi constatada maior proximidade entre valores aqui obtidos e os dos três trabalhos mais recentes entre aqueles quatro descritos na seção 5.1, Capítulo 5, a saber, de Rauber (2006), de Barbosa (2015) e principalmente de Escudero (2009), que foi o que apresentou resultados mais próximos.

Uma explicação para a confluência de resultados pode ser metodológica: os trabalhos acima apontados coletaram fala monitorada, não espontânea; o trabalho de Moraes, Callou e Leite (1996) considerou dados obtidos através de coleta de fala espontânea. Além disso, a maior proximidade com o trabalho de Escudero (2009) também pode indicar a influência do número de indivíduos das amostras, uma vez que, considerando apenas os homens falantes do PB, a amostra de Barbosa e Madureira (2015) contou com quatro sujeitos; Rauber (2006) contou com seis e Escudero (2009) contou com dez, podendo ser considerada a mais representativa entre os três, se considerarmos o número mínimo sugerido por Kinoshita e Ishihara (2015) de vinte indivíduos, conforme já referido na seção 4.2, Capítulo 4.

A respeito do número de participantes, salienta-se, na comparação do presente trabalho com o de Seara (2000), o conjunto de dados referente à sílaba [ni], que se apresentou mais semelhante entre os seis conjuntos mencionados, a saber, [ma, mi, mu, na, ni, nu]. Tal conjunto foi justamente aquele que teve um número mais próximo de participantes, em função da seleção efetuada sobre a amostra inicial, descrita em 5.2. Após retirados os contextos de dispersão extrema, os valores foram obtidos sobre apenas quatro participantes, enquanto a pesquisa de Seara (2000), para as consoantes nasais, foi realizada sobre cinco participantes. Apesar dessa particularidade, os trabalhos descritivos do PB, principalmente os quantitativos, de aspectos acústicos ou relativos a outros fenômenos linguísticos, são fundamentais para o trabalho forense, uma vez que permitem que se produza conhecimento a respeito da tipicidade dos fenômenos

da língua, tema que foi introduzido na seção 2.5.4, Capítulo 2 e que será discutido na seção 5.3.1 a seguir.

O ordenamento mencionado no Quadro 5.2, no que diz respeito ao desvio padrão de F1, indica uma relação articulatória de maior variabilidade para as vogais abertas em relação às fechadas, porém há de se considerar que o desvio padrão é proporcional ao tamanho médio da variável medida, e as vogais abertas sempre apresentam F1 médio maior do que as fechadas. Confirma-se essa característica a respeito do desvio padrão ao se comparar os CV_{inter} e CV_{intra} expressos no Quadro 5.4 para F1 das vogais. Não se percebe comportamento evidentemente diferente entre as fechadas e a aberta, indicando a propriedade intrínseca do CV de permitir a comparação de grandezas diferentes.

Na sequência da observação anterior e com o objetivo de introduzir o tema da próxima seção, a aplicabilidade do IV como critério de comparação entre os parâmetros acústicos dos sons estudados, observa-se, também no Quadro 5.4, o IV de F1 da vogal aberta, de 0,452, valor abaixo do das outras vogais, fato que aponta para sua relevância para a CL em relação aos demais valores presentes no Quadro 5.4. Esse resultado é corroborado no Quadro 5.15, a ser apresentado na seção 5.3.1, a seguir, em que para todas as estratificações apresentadas, definidas pelo som do contexto anterior à vogal, a aberta, para F1, ocupa as cinco primeiras posições, de maneira que se torna a mais indicada para a situação forense. O conceito que permite essa inferência é o tema da presente pesquisa e será discutido na próxima seção.

5.3.1 Aplicação forense

Antes de discutir especificamente a utilidade do IV para a CL, levanta-se uma particularidade em relação ao trabalho de Bosch (2003), que se utilizou de MANOVA para avaliar os resultados das medições acústicas. A análise estatística conduzida indica a significância de f0 e F3 como melhores parâmetros para CL devido à interação dessas variáveis com a variável falante. Há uma diferença de proposta entre esse estudo e o estudo em tela, pois através do IV busca-se determinar um indicador não apenas a respeito da utilidade do parâmetro acústico, mas, além disso, estabelecer um ordenamento entre os parâmetros. Há outro ponto diferencial em relação ao trabalho de Bosch (2003)

no que diz respeito à estratégia de medição através de *script*. Naquela pesquisa, o autor define a configuração para o cálculo de LPC, número de formantes e valor máximo de formantes, de maneira fixa, o que permitiu a leitura de forma automática. Na pesquisa em tela, o *script* permitiu que esses valores fossem determinados em cada leitura dos sons, através da apresentação de uma tela, ilustrada na Figura 4.2, Capítulo 4. Tal recurso tornou o processo obviamente mais lento, porém permitiu o acoplamento adequado do LPC ao espectrograma para todas as vogais, conforme recomendado por Thomas (2011, p. 41), o que garante consistência maior dos dados aqui obtidos em relação aos de Bosch (2003). Nesse aspecto reside justamente uma explicação para o insucesso em se obter indicador de variabilidade para as consoantes nasais, já que a leitura de formantes nasais foi obtida pela extrapolação do número de formantes na configuração do LPC, conforme indicado por Seara (2000) e Thomas (2011). A variabilidade individual e a menor energia dos formantes nasais, porém, fez com que se obtivessem leituras diferenciadas desses formantes, com muitos formantes espúrios presentes nas leituras. Seara (2000), para evitar o problema, realizou leitura do LPC associada ao FFT. Essa estratégia, porém, não era viável com a quantidade de dados envolvidos na pesquisa em tela. Apesar do insucesso das medições, a variabilidade individual das consoantes nasais sugere que, através de outras estratégias de análise, seja possível que esses sons ofereçam contribuição à CL.

A ideia inicial da pesquisa em tela parte das noções de Rose (2001) e Aitken e Gold (2013) de similaridade e tipicidade. A similaridade diz respeito à semelhança entre as evidências cotejadas. A tipicidade diz respeito da posição da evidência em relação à língua compartilhada pela comunidade estudada. Eriksson (2012) propõe a mesma ideia utilizando outros termos, consistência e distintividade, que são considerados pelo autor como paralelos à similaridade e tipicidade. A relação entre esses conceitos definirá o peso da evidência sob análise.

Heuvel (1996) dá nome ao peso da evidência, que chama de especificidade de falante, e o define como a razão entre variabilidade intraindividual de falante e a variabilidade interindividual de falante. Essa relação entre termos é a que se adotou na presente pesquisa e tais termos estão

representados numericamente pelo CV_{intra} , para similaridade ou consistência; CV_{inter} , para tipicidade ou distintividade, e IV para a especificidade ou peso da evidência.

Adotando-se as noções referidas para analisar os resultados do trabalho em tela no que toca à sua aplicação ao âmbito forense, percebe-se o conceito da especificidade de falante em relação aos parâmetros propostos, na observação do Quadro 5.15 a seguir, que apresenta os valores de IV ordenadamente, de maneira crescente para cada parâmetro. Da esquerda para a direita, tem-se f_0 , F1, F2, F3 e, na coluna mais à direita, o valor médio dos quatro parâmetros para a sílaba indicada, o que permite uma avaliação do desempenho por sílaba.

Como análise genérica do Quadro 5.15, sempre mantendo a referência de 0,500 como o marco superior para os IV, fica clara, e já foi comentada, a posição de vantagem da vogal aberta [a] em relação às demais para os valores referentes a F1, pois as cinco primeiras posições são ocupadas por sílabas com essa vogal ([[a, na, fa, sa, ma]). Retoma-se aqui a menção, feita na seção 5.1 deste capítulo, de que há uma relação acústico-articulatória para o IV de F1 de [a] ser o menor independente do contexto anterior. Como o F1 é o parâmetro relacionado à abertura e [a] é o som mais aberto, a pista articulatória é mais evidente para o indivíduo, oferecendo um referencial mais consistente para sua produção e ocasionando uma variabilidade intraindividual menor. A variabilidade interindividual, por sua vez, tenderá a seguir mais consistentemente a dimensão física dos indivíduos, ou seja, é mais sensível às dimensões da cavidade oral. Para F2, observa-se no Quadro 5.15 uma melhor colocação das vogais anteriores e altas ([si, ji, fe, fe, se, me, se, je, fi, je, ni, fo, mi, me, ne, ne]), exceção feita à sílaba [fo] e, para F3, as posteriores médias assumem de forma geral uma melhor colocação ([fo, so, no, mo, fa, ji, jo, no, jo]).

Quadro 5.15 – Ordenamento das sílabas medidas por IV.

Sílabas ordenadas pelo Índice de variabilidade									
f0		F1		F2		F3		Média dos Parâmetros	
jo	0.394	ja	0.436	si	0.359	fɔ	0.408	si	0.464
se	0.410	na	0.451	ji	0.378	so	0.422	fe	0.470
ju	0.410	fa	0.480	fe	0.389	nɔ	0.423	ji	0.471
jɛ	0.418	sa	0.483	fɛ	0.390	mɔ	0.427	se	0.474
mɔ	0.419	ma	0.488	se	0.411	fa	0.430	fa	0.477
so	0.420	nɔ	0.504	me	0.413	ji	0.436	fɛ	0.483
fe	0.420	ʃɔ	0.537	sɛ	0.424	ʃo	0.439	nɔ	0.486
sɔ	0.423	mə	0.549	je	0.425	no	0.444	je	0.487
fɛ	0.426	mi	0.557	fi	0.428	ʃɔ	0.445	jɛ	0.488
mɛ	0.428	ni	0.557	jɛ	0.431	fi	0.446	ni	0.488
no	0.433	nə	0.562	ni	0.446	ma	0.462	mi	0.491
fa	0.434	sə	0.565	fo	0.472	sɔ	0.470	fi	0.495
ni	0.435	mɔ	0.567	mi	0.478	mi	0.472	ʃɔ	0.500
mə	0.436	fə	0.583	mɛ	0.480	si	0.473	ma	0.500
je	0.437	nu	0.585	ne	0.494	fo	0.474	mɔ	0.501
sə	0.441	nɛ	0.597	nɛ	0.500	fe	0.479	fɔ	0.503
mo	0.441	ʃə	0.597	ma	0.538	su	0.481	me	0.503
na	0.443	sɔ	0.600	nɔ	0.554	je	0.482	ʃo	0.508
nɛ	0.447	ju	0.607	mə	0.556	mo	0.491	mə	0.508
fə	0.448	mɛ	0.613	ʃɔ	0.569	se	0.491	ja	0.511
ja	0.455	ji	0.617	nə	0.572	jɛ	0.500	na	0.511
su	0.456	mu	0.618	fɔ	0.573	me	0.507	sɔ	0.512
fɔ	0.457	fu	0.629	na	0.594	fɛ	0.510	so	0.512
mu	0.466	si	0.637	fa	0.600	mə	0.512	sɛ	0.513
ʃɔ	0.467	fɔ	0.638	sɔ	0.605	fə	0.512	fo	0.517
si	0.468	fi	0.640	fə	0.612	ʃə	0.512	mɛ	0.517
mi	0.468	ʃo	0.645	ju	0.636	sɛ	0.516	sa	0.530
me	0.475	su	0.648	mu	0.647	sa	0.518	no	0.530
nu	0.481	so	0.649	no	0.651	fu	0.523	fə	0.531
nɔ	0.482	se	0.649	sə	0.652	ja	0.531	nɛ	0.535
ʃə	0.483	ne	0.653	sa	0.654	ni	0.539	ju	0.537
fu	0.486	fo	0.664	so	0.659	nu	0.545	nə	0.546
nə	0.489	je	0.671	ʃo	0.666	ne	0.555	ne	0.547
sa	0.495	jɛ	0.672	mɔ	0.671	ju	0.556	sə	0.551
sɛ	0.496	sɛ	0.677	fu	0.684	nə	0.571	su	0.556
fo	0.498	fe	0.682	ja	0.691	mu	0.586	nu	0.566
ne	0.511	me	0.690	nu	0.693	sə	0.593	ʃə	0.566
ji	0.516	no	0.691	su	0.722	mɛ	0.596	fu	0.569
ma	0.517	mo	0.703	ʃə	0.732	na	0.602	mo	0.570
fi	0.518	fɛ	0.709	mo	0.775	nɛ	0.636	mu	0.570
f0 médio	0.454	F1 médio	0.593	F2 médio	0.531	F3 médio	0.494	Média geral	0.513

Fonte: O Autor (2017).

Além da apresentação dos valores de IV segundo as sílabas, que podem servir como referências ao exame de CL, os valores de IV encontrados concernentes ao VSA e as DE podem ser confrontados com aqueles presentes no Quadro 5.15. Como já foi dito, na seção 5.1.2, os valores de IV a serem considerados aqui, de VSA e DE, são aqueles presentes nas cinco linhas inferiores do Quadro 5.9, referentes às estratificações segundo os contextos sonoros anteriores aos sons vértices do triângulo vocálico. Tomando como referência as colunas, o VSA e a DE_{ai} apresentam valores de IV evidentemente

adequados à referência proposta, semelhantes aos melhores valores de F2 e de F3 das vogais. Além disso, tomando agora a orientação das linhas no Quadro 5.9, vê-se que os contextos anteriores de [s] e [ʃ] favorecem um IV menor, uma vez que para todas as colunas apresentam os menores valores.

Se o IV recém apresentado diz respeito à especificidade ou ao peso da evidência, o CV_{intra} diz respeito à similaridade. Assim, o CV_{intra} serve como referência para o momento da CL em que há a comparação envolvendo um parâmetro numérico específico, em que a comparação se dá através da diferença percentual, como será visto adiante. O Quadro 5.16 a seguir apresenta os valores de CV_{intra} ordenados segundo os sons do contexto anterior à vogal.

Este trabalho está estruturado para servir de instrumento de referência com relação aos achados de variabilidade. Assim, os resultados de CV_{intra} , para as vogais, agrupados em classes ou discriminados em suas subdivisões, podem permitir a avaliação da similaridade quando for realizado o confronto de vestígios acústicos em contextos linguísticos análogos, no momento da elaboração do laudo pericial, conforme será exemplificado a seguir.

A ilustração apresentada na Figura 2.26, na seção 2.5.4, Capítulo 2, trata de um exemplo real de estratégia utilizada na CL. As curvas LPC, padrão e questionada, de um ponto específico de uma vogal, em contexto de palavra idêntico e em ponto equivalente no momento articulatório, são cotejadas e as diferenças entre os picos das curvas são apresentadas em valores percentuais⁷⁸. Para exemplificar, imagine-se que aquela curva LPC fosse da palavra “foto” e que as diferenças percentuais fossem de 3% para F1 e 4% para F2. Seria possível consultar a similaridade da sílaba [f] no Quadro 5.16, de 5,5% para F1 e de 5,6% para F2, sendo então esse cotejo considerado como similar e com IV de 0,638 e 0,573, respectivamente. Tal procedimento permite estabelecer o peso dessa evidência frente a outras eventuais comparações efetuadas em que a similaridade indique a consistência da comparação.

⁷⁸ O procedimento de cálculo da diferença percentual é aquele apresentado na seção 5.1 e através do qual foram calculados os valores de proximidades entre os valores médios deste estudo e os valores médios de Rauber (2006) e Escudero et al. (2009), apresentados no Quadro 5.2.

Quadro 5.16 – Similaridade na CL – valores de CV_{intra}.

	f0		F1		F2		F3		Média Parâmetros	
	CV	IV	CV	IV	CV	IV	CV	IV	CV	IV
fa	8.1%	0.434	5.4%	0.480	3.8%	0.600	3.5%	0.430	4.7%	0.477
fɔ	8.2%	0.457	5.5%	0.638	5.6%	0.573	4.3%	0.408	5.6%	0.503
fe	7.5%	0.420	5.9%	0.682	2.6%	0.389	3.3%	0.479	4.0%	0.470
fə	8.2%	0.448	6.3%	0.583	5.5%	0.612	3.3%	0.512	5.2%	0.531
fɛ	7.3%	0.426	5.6%	0.709	2.8%	0.390	3.4%	0.510	4.1%	0.483
fi	8.7%	0.518	5.9%	0.640	2.6%	0.428	3.0%	0.446	4.0%	0.495
fo	8.0%	0.498	6.3%	0.664	6.1%	0.472	3.5%	0.474	5.5%	0.517
fu	8.9%	0.486	6.6%	0.629	7.4%	0.684	5.4%	0.523	6.8%	0.569
ma	8.1%	0.517	7.4%	0.488	3.8%	0.538	4.4%	0.462	5.3%	0.500
mɔ	7.8%	0.419	7.9%	0.567	6.0%	0.671	5.0%	0.427	6.4%	0.501
me	7.7%	0.475	6.7%	0.690	2.8%	0.413	3.4%	0.507	4.3%	0.503
mə	7.6%	0.436	9.3%	0.549	5.0%	0.556	3.6%	0.512	5.5%	0.508
mɛ	7.8%	0.428	7.4%	0.613	3.3%	0.480	3.7%	0.596	4.8%	0.517
mi	8.6%	0.468	7.2%	0.557	2.7%	0.478	3.1%	0.472	4.3%	0.491
mo	8.1%	0.441	7.3%	0.703	6.5%	0.775	3.8%	0.491	5.9%	0.570
mu	8.4%	0.466	7.2%	0.618	8.7%	0.647	4.5%	0.586	6.7%	0.570
na	7.8%	0.443	6.7%	0.451	3.2%	0.594	5.2%	0.602	5.1%	0.511
nɔ	8.4%	0.482	7.0%	0.504	5.4%	0.554	4.6%	0.423	6.0%	0.486
ne	8.1%	0.511	7.1%	0.653	3.0%	0.494	3.7%	0.555	4.6%	0.547
nə	8.4%	0.489	9.0%	0.562	3.8%	0.572	3.8%	0.571	5.3%	0.546
nɛ	7.6%	0.447	7.0%	0.597	3.4%	0.500	3.8%	0.636	4.8%	0.535
ni	8.1%	0.435	8.2%	0.557	3.7%	0.446	4.0%	0.539	5.2%	0.488
no	7.8%	0.433	7.6%	0.691	6.3%	0.651	3.3%	0.444	5.6%	0.530
nu	9.0%	0.481	7.2%	0.585	8.4%	0.693	3.9%	0.545	6.4%	0.566
sa	8.7%	0.495	5.3%	0.483	3.3%	0.654	4.4%	0.518	4.8%	0.530
sɔ	7.8%	0.423	5.0%	0.600	4.2%	0.605	4.1%	0.470	4.9%	0.512
se	7.9%	0.410	5.7%	0.649	2.7%	0.411	3.1%	0.491	4.0%	0.474
sə	7.5%	0.441	6.0%	0.565	3.3%	0.652	3.6%	0.593	4.5%	0.551
sɛ	7.5%	0.496	5.2%	0.677	3.0%	0.424	3.5%	0.516	4.2%	0.513
si	8.8%	0.468	6.5%	0.637	2.7%	0.359	3.6%	0.473	4.3%	0.464
so	7.6%	0.420	5.7%	0.649	6.5%	0.659	3.5%	0.422	5.4%	0.512
su	8.7%	0.456	6.4%	0.648	8.8%	0.722	4.2%	0.481	6.4%	0.556
ja	8.5%	0.455	5.2%	0.436	3.4%	0.691	4.7%	0.531	4.9%	0.511
ɔ	8.3%	0.467	5.0%	0.537	4.2%	0.569	4.1%	0.445	5.0%	0.500
je	7.5%	0.437	6.0%	0.671	2.8%	0.425	3.9%	0.482	4.4%	0.487
jə	7.8%	0.483	6.2%	0.597	4.5%	0.732	4.0%	0.512	5.3%	0.566
jɛ	7.7%	0.418	6.1%	0.672	3.0%	0.431	3.8%	0.500	4.5%	0.488
ji	9.1%	0.516	6.4%	0.617	2.5%	0.378	3.9%	0.436	4.3%	0.471
jo	7.3%	0.394	5.9%	0.645	6.3%	0.666	4.0%	0.439	5.6%	0.508
ju	7.9%	0.410	6.5%	0.607	8.6%	0.636	4.2%	0.556	6.3%	0.537

Fonte: O Autor (2017).

Com relação à alteração dos IVs nas estratificações apresentadas em 5.1.1 para as vogais, ressalta-se que houve um aumento nos valores médios desse índice à medida que se avançou na estratificação. Há de se levar em conta que no agrupamento inicial, em que não havia estratificação, estava inserida uma variabilidade no conjunto de dados que é inerente à língua e não aos indivíduos, pois as classes sonoras, nasais e fricativas, apresentam efeitos coarticulatórios diferenciados sobre o comportamento das vogais. Assim, ao analisar-se as

vogais como um todo, é possível obter-se uma variabilidade interindividual maior, que não é própria das diferenças dos indivíduos, mas ocasionada por característica inerente ao sistema linguístico e que enviesa a mensuração do IV.

Outro fato que clama pela atenção diz respeito ao IV da consoante nasal para a sílaba [ni]. Observando-se o Quadro 5.12, encontra-se o valor de IV de 0,372 para essa sílaba, o que levaria a acreditar que seria um bom elemento para a CL. Porém, como já foi referido no início da seção 5.3, para [ni] apenas quatro participantes compuseram o cálculo do respectivo IV. Confrontando o Quadro 5.10, que apresenta o número de participantes envolvidos no cálculo para cada sílaba, com os valores de IV, percebe-se uma certa relação entre o número menor de indivíduos e os valores mais baixos de IV. A lógica dessa hipótese pode ser explicada uma vez que o CV_{inter} foi alto para as nasais, a ponto de ocasionar o descarte desses sons, e a tendência natural é a de que, para um número menor de indivíduos, o CV_{intra} seja forçadamente menor. Daí se infere que o IV será tanto mais consistente quanto maior for o tamanho da amostra estudada.

Essas duas considerações levantadas obrigam o pesquisador e o perito, ao realizar a CL, a manter a atenção sobre o conjunto dos dados que for utilizado para estabelecer esses indicadores. A segunda consideração leva a pensar que o conjunto de dados para obter os indicadores deve conter um número considerável de indivíduos e a primeira consideração indica que a variabilidade do conjunto deve ser proveniente das diferenças entre as características particulares dos indivíduos e não inerente a aspectos do sistema linguístico, como foi o caso mostrado na estratificação efetuada na seção 5.1.1, em que o CV_{inter} apresenta-se alto pelo agrupamento de contextos anteriores muito diferenciados ao se considerar influências coarticulatórias nasais e fricativas. Nesse caso, o que ocasiona a maior variabilidade é uma característica de diversidade linguística, a coarticulação gerada por diferentes classes, e não diferenças individuais entre os sujeitos pesquisados. Esse mesmo raciocínio pode ser aplicado ao observar o Quadro 5.9, na porção que apresenta os valores de VSA e DE. A primeira linha da porção do quadro referente ao IV apresenta os menores valores. As segunda e terceira linhas, valores um tanto maiores e as cinco linhas inferiores, os maiores valores. O valor aumenta à medida que se

estratificam os conjuntos, uma vez que a variabilidade é inerente ao sistema e não fruto exclusivo das diferenças individuais. Aqui reside uma característica que embasou a proposta do conjunto de dados tal como foi feita. O ambiente linguístico é favorável a uma variabilidade pequena e controlada na busca da variabilidade mínima.

Encerra-se a seção de aplicação forense, em que foram apresentados os indicadores e discutidas suas aplicações, principalmente no que diz respeito aos cotejos acústicos presentes na CL, com eventuais discussões de caráter linguístico mais geral. A seguir apresenta-se a discussão teórica envolvendo os conceitos da Fonologia Articulatória.

5.3.2 Tratamento teórico

Esta seção dedicar-se-á a considerações baseadas em alguns princípios da Fonologia Articulatória. Em um primeiro momento, será discutida a relação Fonética/Fonologia desde um ponto de vista influenciado pela aplicação dos conceitos da Fonética Acústica e Articulatória na Fonética Forense. Em seguida discutir-se-á a concepção de vogal dentro do modelo da Fonologia Articulatória, que pode ser considerada como um ponto questionável da teoria no que diz respeito a sua representação, apesar de não prejudicar o poder de implementação da fala através de seu modelamento computacional, que é considerado como um dos pontos fortes do modelo.

5.3.2.1 A comensurabilidade entre Fonética e Fonologia e a prática forense

A comensurabilidade entre Fonética e Fonologia é tratada por Albano (2001, pp. 26-31) na tentativa de mostrar que a aquisição da fonologia se dá também através da produção da fala, o que a autora denomina de “aprender fazendo”. A comensurabilidade é apresentada como uma noção que tradicionalmente sofre uma redução por simplificação, pois é tratada desde a perspectiva de o que é contínuo é fonético, e o que é discreto é fonológico, no que se refere ao universo dos sons da fala.

Albano (2001, p. 11) previne que “se a distinção entre as duas disciplinas se esgotasse no método, a Fonética e a Fonologia estariam claramente separadas”. Nesse mesmo sentido, Ohala (1990) distingue apropriadamente a

relação Fonética/Fonologia desde duas perspectivas diferentes, uma que toma as duas como disciplinas e outra, a qual ele se refere como parte do universo da fala, e que será referida como níveis da linguagem. Quando a proposta da Fonologia Articulatória por um objeto de análise que ao mesmo tempo dê conta da representação e da implementação remove uma distinção, ou um “fosso”, nos termos de Albano (2001), o paradigma que está sendo quebrado diz respeito aos níveis da linguagem e não às disciplinas.

Apesar de descrever minuciosamente os fenômenos acústicos e articulatórios, a Fonética Acústica e Articulatória limita-se ao campo da realidade fônica, o que não favorece a “transposição” da lacuna existente entre esse campo teórico com aquele abrangido pelos modelos fonológicos tradicionais. Assim, uma teoria que considere a realização e abstração sonoras como processos que têm por ator um objeto de análise unificado tem maiores chances de oferecer explicações para os fatos linguísticos que sirvam à perspectiva da criminalística, seja na CL, tanto para o procedimento de coleta de padrão como para as considerações técnicas dos laudos, seja na determinação de perfis linguísticos de indivíduos.

A decisão pela Fonologia Articulatória como referencial teórico da presente tese se deu pela postura frente à gênese da linguagem como um fato fisiológico. É justamente a comensurabilidade que fundamenta a decisão pela Fonologia Articulatória, pois tradicionalmente no âmbito da CL as teorias fonológicas não ofereciam explicações aos fatos linguísticos que indicassem uma possibilidade de combinação com a realidade forense, uma vez que o material é exclusivamente a manifestação acústica, tradicional objeto da Fonética. Não por acaso que a área se consagrou com a denominação de Fonética Forense. No exame de CL busca-se a descrição de uma manifestação linguística individual e não a explicação da origem dessa manifestação no sistema linguístico. De fato, é difícil imaginar como as teorias de representação fonológica que estejam no conjunto de Teorias de Tradução (FOWLER ET AL., 1980, p. 376) poderiam trazer algum acréscimo para o exame de CL, uma vez que as análises que o compõem recaem fundamentalmente sobre a realização fonética do indivíduo em estudo e, assim, os processos de representação, uma vez que são produtos de inferências e não de constatações diretas sobre o

material sonoro, normalmente não oferecem informações peculiares desse indivíduo. A CL baseia-se no fato fonético, ou implementação. Ao examinar o fato fonético e apenas o fato fonético, o exame pericial não tem como tirar algum proveito da fonologia tradicional, pois não há explicação a respeito do sistema linguístico que possa oferecer luz sobre o fato fonético individual.

A Fonologia Articulatória, ao tratar os dois níveis como um só, o fônico (ALBANO, 2001), é a teoria que se aproxima da natureza do trabalho pericial, pois a visão de linguagem como ação, e não como representação, é capaz de contemplar o simbólico e o dinâmico em um objeto. Browman e Goldstein (1989a,b) pleiteiam o grande poder de implementação computacional de seu modelo e nesse aspecto reside a possibilidade de acréscimo dessa teoria ao campo forense. Modelos computacionais que sintetizam a fala a partir de pautas gestuais (NAM ET AL., 2010; NAM ET AL., 2012) podem oferecer recursos para a simulação de condições específicas da fala, ao passo que modelos que representam espacialmente o trato vocal a partir do sinal acústico (YEHIA, 1997; PANCHAPAGESAN E ALWAN, 2011) e mesmo modelos que produzem pautas gestuais a partir do sinal acústico (MITRA ET AL., 2010), podem oferecer informações espaciais sobre o trato vocal e conseqüentemente sobre seus articuladores, que não são diretamente observáveis através de imagens espectrográficas. A capacidade de ir do sinal acústico para uma representação geométrica, a inversão de fala, e da representação por pauta gestual para o sinal acústico, a síntese de fala, pode oferecer estratégias de interpretação das manifestações linguísticas do indivíduo que colaborem com o processo de CL.

Tendo em vista que as vogais são o material de estudo principal da presente pesquisa, a seção seguinte dedicar-se-á a discutir sua posição singular dentro do modelo.

5.3.2.2 A situação peculiar das vogais em relação ao modelo teórico

Nesta seção discute-se o tratamento dispensado à vogal pela Fonologia Articulatória. Será apresentada uma análise com base nos dados vocálicos obtidos neste estudo que pretende sugerir um caminho de pesquisa em busca de uma abordagem para as vogais dentro do modelo.

A natureza diferenciada entre vogais e consoantes é uma das certezas tanto na Fonética quanto na Fonologia. A fonologia consegue estabelecer classes em função da dedicação à abstração e da “liberdade” em relação à Fonética. Essa, por sua vez, está dedicada ao fato, à implementação, ao mundo físico, o que dificulta que a Fonética se dedique à representação além daquelas provenientes de uma transcrição, espectrograma, palatograma, cinevideofluoroscopia, MRI, ou outros dispositivos que digam respeito à realidade da matéria fônica, seja acústica perceptual ou instrumental, articulatória por sensores ou por imagens. Por isso é tradicionalmente denominada como uma ciência que lida com o dado contínuo.

A grande conquista da Fonologia Articulatória é ser um modelo representacional fonológico que permite a implementação física, concretizada pelo sistema TADA, referido no Capítulo 3. Além disso, o modelo inova ao rever o primitivo de análise linguística. Isso funciona muito bem para as consoantes, pois, pelo que parece, elas são mais suscetíveis à classificação, são “menos contínuas” por natureza. As vogais são “mais contínuas”. A continuidade aqui referida não está vinculada à noção de traço fonológico, mas de característica física, no tempo. Vogal é facilmente produzida no tempo, enquanto a consoante não, excetuando-se alguns casos específicos, por suas naturezas articulatórias. Por isso a consoante é executada sobre a vogal. Ou como a Fonologia Articulatória define, a constrição se dá sobre a vogal. Assim, a constrição não é a vogal propriamente, mas uma espécie de gesto sobre o qual a constrição, característica da consoante, atua. Apesar dos descritores do modelo representarem a vogal, essa representação não é sempre categórica, havendo casos em que é necessária a complementação da representação através dos descritores contínuos dentro do TADA, alvo e rigidez (target e stiffnes), conforme foi apresentado no Capítulo 3.

Apesar da teoria considerar o gesto como unidade representacional, pode-se procurar tomar a unidade como o descritor, uma vez que são os descritores que definem as classes. Nesse sentido, os descritores assumem um papel fundamental, o que justificaria a dedicação ao estabelecimento de estratégias para a proposição de categorias mais definidas para as vogais.

Justamente porque a Fonologia Articulatória se dedica aos fenômenos de sobreposição (coprodução e coarticulação) que ocorrem pela combinação consoante-vogal, e principalmente porque são as consoantes que carregam as constrictões sobre as quais os fenômenos ocorrem, como se vê na maioria dos casos que se encontram na literatura, as consoantes são “privilegiadas” no tratamento pelo modelo, o que se percebe na evolução da teoria. Browman e Goldstein (1986), quando apresentam as primeiras pautas gestuais, reservam a primeira linha da pauta, ainda esboçada, para as vogais expressas genericamente, como pode ser observado na Figura 3.4, no Capítulo 3. Na sequência dos trabalhos, Browman e Goldstein (1988, p. 88) expressam de maneira clara que inicialmente tinham a intenção de abordar as vogais na relação consoante-vogal, mas que não foi possível, dado os mesmos pontos considerados nas medições com dispositivos sensoriais eletrônicos. As vogais normalmente têm uma representação que está principalmente associada às variáveis do articulador TB e mesmo Browman e Goldstein (1986, 1988 e 1989a) admitem representações que, por deterem-se em fatos consonantais, relevam a descrição da vogal a uma forma mais genérica.

A necessidade de estudos para que a Fonologia Articulatória contemple de forma mais adequada as categorias relacionadas às vogais já foi apontada por Byrd (2003, p. 90). Sobre a categorização de vogais nos sistemas linguísticos, apresentaram-se diferentes definições quanto ao número de posições possíveis de serem ocupadas por sons no espaço vocálico (LINDBLUM, 1986; LADEFOGED E MADDIESON, 2008; IPA, 2009), que expressam a dificuldade de se categorizar elementos em um espaço diminuto como é a cavidade oral.

Se considerarmos como verdadeira a afirmação de Koenig (2004), de que uma vogal classificada em uma língua segundo uma representação não necessariamente ocupa a mesma posição no espaço acústico do que a mesma representação vocálica em outra língua e que, conseqüentemente, não há de se imaginar um conjunto de vogais universais, ou símbolos vocálicos universais, mas sim há de se pensar em seres humanos que produzem e interpretam tais sons, então aí reside um argumento importante, principalmente no que diz

respeito à produção sonora por seres humanos, para o exercício que será proposto como encerramento da discussão.

Uma vez que o número de indivíduos que compõem a amostra do presente trabalho é adequado, segundo a definição de Kinoshita e Ishihara (2015), os valores obtidos serão representativos da produção vocálica, não com a pretensão de serem descritivos do PB, mas pelo menos com o objetivo de serem indicadores de variabilidade da ocupação do espaço vocálico. A noção a ser proposta baseia-se no descritor do grau de constrictão para a variável TBCD, o alvo (target), que é configurado de forma não categórica em milímetros e é justamente o elemento que determinará, associadamente à variável TBCL, qual a vogal, dentre aquelas ambíguas cujos exemplos foram apresentados na seção 3.3, Capítulo 3, que está sendo representada no modelo. Esse valor atribuído ao alvo de TBCD não é absoluto, apesar de ser um valor específico em milímetros, pois é relativo a uma dimensão de trato vocal que pode ser configurada no modelo e, assim, há uma proporção do alvo em relação às dimensões do trato estabelecidas.

Aproveitando a estrutura de dados que compõem o presente estudo, selecionaram-se os valores vocálicos médios das extremidades do triângulo que representa o espaço vocálico, apenas aqueles que se referem às vogais precedidas pelos sons fricativos, tendo em vista a menor coarticulação desses em relação aos consoantes nasais, e foram calculadas as DE do eixo anterior {ai} e seus intermediários {aε}, {εe} e {ei}. Esses valores de DE foram organizados conforme o Quadro 5.17, apresentado a seguir. Na primeira coluna indica-se o participante, na segunda, o contexto anterior à vogal, da terceira à sexta, as DE {ai}, {aε}, {εe} e {ei} e, nas últimas três, as proporções referentes às distâncias intermediárias e à distância total do eixo anterior. O eixo anterior foi escolhido em função de ter apresentado menor variabilidade em relação aos outros dois do triângulo vocálico. Estabeleceu-se, assim, a proporção entre cada uma das três distâncias intermediárias e a distância total entre [a] e [i] para os vinte sujeitos e para os três contextos fricativos anteriores à vogal.

Quadro 5.17 – DE total e parciais do eixo anterior e suas proporções.

Informante	Contexto Anterior	DE (ae, ee, ei)				Proporção das DE			
		ae	ee	ei	ai	ae/ai	ee/ai	ei/ai	
ACRR	f_	613	259	272	1,122	54.6%	23.3%	24.1%	
ACRR	s_	554	234	221	980	56.7%	24.1%	22.2%	
ACRR	x_	475	245	236	936	50.8%	26.4%	25.1%	
AFM	f_	755	200	150	1,043	72.5%	19.2%	14.3%	
AFM	s_	658	195	178	978	67.3%	19.9%	18.3%	
AFM	x_	612	209	157	951	64.4%	22.1%	16.4%	
AFRB	f_	452	246	126	804	56.3%	30.7%	15.6%	
AFRB	s_	404	244	123	750	53.8%	32.8%	16.4%	
AFRB	x_	343	231	110	645	53.2%	36.0%	17.1%	
APG	f_	442	260	152	828	53.4%	31.3%	18.5%	
APG	s_	335	265	188	773	43.2%	34.8%	24.1%	
APG	x_	314	283	187	774	40.6%	36.5%	24.1%	
CAPM	f_	533	181	160	846	63.2%	21.5%	18.8%	
CAPM	s_	481	160	172	785	61.6%	20.3%	21.9%	
CAPM	x_	447	178	172	783	57.1%	22.7%	21.9%	
CDK	f_	628	139	190	938	66.9%	14.8%	20.2%	
CDK	s_	517	151	191	839	61.7%	18.1%	22.6%	
CDK	x_	462	112	178	735	62.9%	15.3%	24.1%	
CRD	f_	531	263	207	982	54.2%	26.8%	20.9%	
CRD	s_	496	208	278	905	55.0%	22.9%	29.3%	
CRD	x_	451	256	196	882	51.1%	28.8%	22.3%	
EGS	f_	578	280	211	1,056	54.8%	26.6%	19.9%	
EGS	s_	524	256	213	975	53.8%	26.2%	21.8%	
EGS	x_	498	278	246	1,003	49.9%	27.8%	24.4%	
FTC	f_	715	262	101	1,005	71.3%	26.2%	10.1%	
FTC	s_	626	269	97	934	66.7%	29.0%	10.5%	
FTC	x_	572	192	119	840	68.1%	22.9%	14.2%	
IAB	f_	556	235	243	976	56.8%	24.2%	25.3%	
IAB	s_	548	223	215	931	58.4%	24.5%	23.5%	
IAB	x_	501	223	237	901	55.5%	25.2%	26.7%	
MFS	f_	766	202	220	1,172	65.5%	17.3%	18.6%	
MFS	s_	731	213	205	1,126	64.9%	18.8%	18.2%	
MFS	x_	673	226	228	1,117	60.3%	20.2%	20.4%	
MPG	f_	649	182	159	931	69.9%	19.6%	17.0%	
MPG	s_	460	201	178	784	58.8%	25.5%	22.7%	
MPG	x_	406	214	166	757	53.5%	28.4%	21.9%	
MRR	f_	827	285	152	1,222	67.6%	23.5%	12.3%	
MRR	s_	748	290	195	1,212	61.8%	24.0%	15.9%	
MRR	x_	784	285	185	1,231	63.7%	23.2%	14.9%	
MZ	f_	587	319	148	979	60.6%	33.0%	15.3%	
MZ	s_	532	274	209	984	54.1%	28.1%	21.2%	
MZ	x_	519	289	227	990	52.5%	29.2%	22.8%	
PMP	f_	617	257	234	1,083	56.9%	23.8%	21.5%	
PMP	s_	529	232	264	976	54.3%	23.7%	27.0%	
PMP	x_	431	231	247	889	48.4%	26.0%	28.0%	
PROF	f_	577	213	90	810	71.2%	26.5%	11.1%	
PROF	s_	521	207	120	783	66.6%	26.5%	15.3%	
PROF	x_	431	197	123	716	60.2%	27.7%	16.9%	
RCS	f_	637	214	187	976	65.3%	21.9%	19.2%	
RCS	s_	734	191	204	1,076	68.3%	17.7%	18.9%	
RCS	x_	637	214	187	976	65.3%	21.9%	19.2%	
RFS	f_	577	249	119	895	64.6%	27.9%	13.1%	
RFS	s_	468	253	154	858	54.9%	29.7%	17.5%	
RFS	x_	511	202	194	820	62.7%	24.8%	24.0%	
TBS	f_	533	234	150	884	60.3%	26.5%	16.9%	
TBS	s_	482	231	111	806	59.8%	28.8%	13.7%	
TBS	x_	435	224	115	760	57.3%	29.5%	15.1%	
VJP	f_	458	232	229	895	51.3%	26.1%	25.3%	
VJP	s_	436	197	223	828	52.6%	23.8%	26.8%	
VJP	x_	341	231	245	809	42.2%	28.6%	30.3%	
						Média das proporções	58.9%	25.2%	19.9%
						DP das proporções	7.2%	4.7%	4.8%
						CV das proporções	12.3%	18.8%	23.9%

Fonte: O Autor (2017).

Os valores médios das proporções ficaram em 58,9% para a relação $\{a\epsilon\}/\{ai\}$, 25,2% para a relação $\{\epsilon e\}/\{ai\}$ e para a relação $\{ei\}/\{ai\}$, 19,9%. Comparando-se as médias com as respectivas distâncias, pode-se notar uma certa consistência no comportamento dos valores.

Também calcularam-se os coeficientes de variação dos conjuntos, que ficaram em 12,3% para a relação $\{a\epsilon\}/\{ai\}$, 18,8% para a relação $\{\epsilon e\}/\{ai\}$ e para a relação $\{ei\}/\{ai\}$, 23,9%. Percebe-se a relação inversa entre a proporção das DE e o coeficiente de variação respectivo, indicando que o maior espaço acústico existente para a realização da vogal ocasiona uma regularidade maior. Nota-se portanto que conjuntos de dados como os aqui expostos podem oferecer pistas a respeito de caminhos a seguir para a resolução de uma questão como a que se coloca em relação às vogais para o modelo gestual. Outras distâncias entre os pontos vocálicos poderiam ser medidas para se realizar análises mais abrangentes do espaço vocálico e talvez colaborar no estabelecimento de categorias vocálicas mais representativas.

O aspecto que parece relevante é de que o sistema se comporta como um sistema dinâmico uma vez que traz inerentemente à sua organização uma certa invariância, ilustrada pela regularidade apresentada pelas DE intermediárias, ao mesmo tempo que a variabilidade está presente nas áreas ocupadas pelos sons e nas DE medidas.

Se há recuperabilidade de aspectos articulatórios através do sinal acústico (MITRA ET AL., 2011; ISKAROUS, 2010) e se a informação acústica compõe o gesto (ALBANO, 2001), então as medidas acústicas devem ser úteis para definir descritores mais adequados às vogais, imaginando a possibilidade de descritores diferenciados, ou estratégias descritivas diferenciadas, para consoantes e vogais. As distâncias euclidianas entre as vogais aparentemente são úteis, pois são medidas relativas e consistentes.

Para encerrar o capítulo, retoma-se a possibilidade de aplicação forense dos conceitos aqui expostos. Uma prática que pode ser sugerida para o procedimento de coleta de padrão envolvido na CL é a gravação das vogais extremas do triângulo vocálico, seja isoladamente ou em contextos semelhantes

aos aqui analisados, como recurso de interpretação articulatória da cavidade oral do indivíduo submetido ao exame. A consistência das proporções descritas anteriormente indica a possibilidade de se utilizar essas informações como elementos comparativos e descritivos do sistema vocálico dos indivíduos examinados na CL.

6 CONSIDERAÇÕES FINAIS

O trabalho apresentado procurou estabelecer indicador que permitisse avaliar o desempenho relativo entre os parâmetros acústicos utilizados na Perícia de Comparação de Locutores a partir de uma amostra de dados de sílabas tônicas, pronunciadas através de frase-veículo por vinte participantes porto-alegrenses. As bases da Fonética Acústica e Articulatória foram estabelecidas para o delineamento da pesquisa, medição dos dados e obtenção dos resultados. A amostra de dados obtida e a análise estatística conduzida foram suficientes para confirmar ou refutar as hipóteses levantadas.

Para avaliar o desempenho relativo dos parâmetros estudados, foi proposto o Índice de Variabilidade como referência, após ter sido descrita a sua obtenção através da razão entre o coeficiente de variação intraindividual e o coeficiente de variação interindividual. A utilidade do índice para a determinação do poder de evidência dos parâmetros acústicos estudados, guardadas as ressalvas a respeito da necessidade de consolidação desse indicador de variabilidade, ficou constatada.

Desse modo, retomam-se a seguir as hipóteses e seus desdobramentos.

1. *É possível determinar valores referenciais de variabilidade, com base em coleta de fala monitorada, que sirvam de referência para o exame de CL.*

Essa hipótese confirma-se uma vez que, através da utilização do Coeficiente de Variação e do Índice de Variabilidade como parâmetros comparativos, foi possível determinar valores de referência mínimos para cada agrupamento de dados.

2. *Há diferença na variabilidade dos parâmetros acústicos de curto termo de diferentes sons do PB, especificamente os vocálicos orais ([i e ε ə a ɔ o u]), que permita ordená-los. Dentre os vocálicos orais, os abertos apresentam menor variabilidade do que os fechados.*

Foi possível ordenar as vogais separadamente para cada formante, conforme apresentado nos Quadros 5.5 e 5.6. Para F2 e F3, as vogais anteriores apresentaram-se com os melhores índices.

3. *A variabilidade da área do espaço vocálico e das distâncias euclidianas vocálicas é análoga à dos sons vocálicos, o que permite seu uso na prática forense.*

A área do espaço vocálico e a distância euclidiana entre [a] e [i] apresentaram Índice de Variabilidade que se equiparam aos melhores índices encontrados para as vogais, o que indica a possibilidade de utilização dos parâmetros geométricos na Perícia de Comparação de Locutores.

4. *Há diferença na variabilidade dos parâmetros acústicos de curto termo de diferentes sons do PB, especificamente os consonantais nasais ([m n]), que permite ordená-los.*

Apesar de ser possível ordenar os valores, os efeitos de dispersão não permitiram que a estratégia estabelecida para a leitura de formantes estabelecesse um conjunto de dados que apresentasse os indicadores de variabilidade em níveis aceitáveis para que os valores formânticos das consoantes fossem considerados apropriados para a Perícia de Comparação de Locutores. Ressalta-se que foi justamente a variabilidade individual inerente às consoantes nasais que ocasionou o insucesso da estratégia adotada para a leitura dos formantes, o que sugere que outras estratégias de análise das consoantes nasais possam ser úteis para a Perícia de Comparação de Locutores. O fenômeno coarticulatório entre consoante nasal e vogal seguinte deve ser avaliado através de uma proposição nova de marcação dos dados, que considere as relações temporais entre os gestos envolvidos na sequência.

5. *Os sons consonantais nasais apresentam variabilidade menor do que os vocálicos orais, prestando-se com mais efetividade como parâmetros de comparação. Dentre as consoantes nasais, a bilabial é a que apresenta menor variabilidade.*

Não. As nasais apresentaram índices maiores do que a maioria das vogais para os três formantes.

6. *Dentre os parâmetros acústicos, f_0 e F_3 são os que apresentam menor variabilidade.*

F_3 e f_0 apresentaram os menores índices, sendo que F_1 e F_2 apresentaram índices semelhantes entre si. Além disso, a área do espaço vocálico e a distância euclidiana entre [a] e [i] também apresentaram índice de variabilidade em níveis adequados para a Perícia de Comparação de Locutores, inclusive menores do que f_0 e F_3 , em alguns casos.

A respeito da intenção, mais genérica, deste trabalho de lançar a proposta de que a Fonologia Articulatória pode colaborar com a análise da evidência forense de natureza linguística, a capacidade de modelamento do trato vocal proposto pela Fonologia Articulatória pode oferecer esclarecimentos sobre a configuração das estruturas componentes do trato relevantes ao exame de CL, não acessíveis através da análise acústica tradicional das evidências linguísticas. A existência de trabalhos de sintetização e, principalmente, de inversão de fala, com literatura referindo, inclusive, o reconhecimento automático de falantes com auxílio da definição de variáveis do trato vocal a partir do sinal acústico sugere que a Fonologia Articulatória oferece recursos que podem colaborar com a Perícia de Comparação de Locutores. Além disso, mesmo não estando fundamentados na Fonologia Articulatória, os resultados de Índice de Variabilidade da área do espaço acústico e da distância euclidiana entre [a] e [i] apontam para a utilidade forense de inferências geométricas e o desdobramento das medidas geométricas – as proporções do eixo anterior do triângulo vocálico – apontam para a possibilidade de existência elementos invariantes no espaço vocálico. A Fonologia Articulatória e a prática forense, particularmente a Perícia de Comparação de Locutores, estão intimamente ligadas ao tema da variabilidade individual.

As bases da Fonologia Articulatória foram apresentadas de maneira a se estabelecer a possibilidade de sua colaboração para a prática forense de CL, bem como foi proposta uma discussão da situação da representação vocálica no modelo, resultando em uma análise do papel da ocupação do espaço acústico pelos elementos vocálicos, pelo menos no que se refere ao eixo anterior do espaço vocálico do PB.

A tendência desta pesquisa foi a de medir a matéria sonora vocálica e consonantal nasal desde uma perspectiva segmental discreta e estática, pois a tradição em fonética aplicada assim determina. As medições foram feitas em uma “zona de estabilidade”, tanto da vogal quanto da consoante. Além desse fator, o conceito de segmento, consagrado pela literatura em Fonética e hábito do pesquisador imerso na realidade forense, induziu à tentativa de evitar os efeitos coarticulatórios entre consoante nasal e a vogal que a seguia, o que foi feito afastando a marcação da vogal em relação à consoante anterior, através da marcação mais à direita do trecho a ser medido. Além disso, a duração não foi considerada, em função de não ser uma medida tradicional na realidade forense. Tais elementos não permitiram, portanto, a observação de fenômenos de sobreposição dos gestos consonantais nasais e vocálicos. Uma revisão da marcação desse material sonoro levando-se em conta tal possibilidade de sobreposição de gestos é uma ideia a ser considerada para trabalhos futuros.

Um outro caminho natural para a continuidade da pesquisa é a inclusão de dados de fala espontânea, preferencialmente com os mesmos indivíduos pesquisados, uma vez que segue o contato do pesquisador com o grupo, o que permitirá determinar a propriedade de se generalizar para a fala espontânea os resultados obtidos através de fala monitorada, na expectativa de se colaborar com a Fonética Forense a partir da produção de material de consulta para a área da Criminalística.

REFERÊNCIAS

- AITKEN, C.; GOLD, E. Evidence evaluation for discrete data. **Forensic science international**, v. 230, n. 1–3, p. 147–55, 10 jul. 2013.
- ALBANO, E. C. **O Gesto e suas Bordas: por uma fonologia acústico-articulatória do português brasileiro**. Campinas: Mercado de Letras, 2001.
- ALBANO, E. C. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. **Revista da ABRALIN**. V. 11, n. 1, 2012, p. 1-30.
- ALEXANDER, A. **Forensic automatic speaker recognition using bayesian interpretation and statistical compensation for mismatched conditions**. 2005. Tese de Doutorado. Escola Politécnica Federal de Lausanne, Lausanne.
- AMINO, K. et al. Forensic Speaker Recognition. In A. Neustein & H. A. Patil (Eds.), **Forensic Speaker Recognition: Law Enforcement and Counterterrorism**. New York, NY: Springer New York. 2012, pp. 21–39.
- ATAL, B. S. Automatic recognition of speakers from their voices. **Proceedings of the IEEE** Vol. 64, nº 4, 1976, p. 460-5.
- BARBOSA, P. A.; MADUREIRA, S. **Manual de Fonética Acústica Experimental. Aplicações a Dados do Português**. São Paulo SP: Cortez Editora. 2015. 591 pp.
- BEIGI, H. **Fundamentals of Speaker Recognition**. New York, NY: Springer, 2011. 942 p.
- BELL-BERTI, F. Understanding velic motor control: studies of segmental context. In: Huffman, M. K.; Krakow, R. A. (eds.). **Nasals, nasalization, and the velum. Phonetics and Phonology Series, Vol. 5**. San Diego, CA: Academic Press, 1993, pp. 63-85.
- BOERSMA, P.; WEENICK, D. Praat: doing phonetics by computer (Version 5.5.03) [Computer program]. Obtido em 21/11/2011, em <<http://www.praat.org>>.
- BONASTRE, J. F. et al. Person authentication by voice: A need for caution. **Eurospeech 2003 - Interspeech 2003. Proceedings of the 8th European Conference on Speech Communication and Technology**, November 2001, p. 33–36, 2003.
- BOSCH, J. C. **Acoustic study of the vowel formant frequencies and F0: a contribution to Catalan forensic phonetics**. International Congress of Phonetic Sciences (ICPhS). Barcelona: 2003. p. 687-690.
- BRAID, A.C.M. **Fonética Forense**. Porto Alegre: Editora Sagra Luzzatto. 1999. 129 p.
- BROWMAN, C. P.; GOLDSTEIN, L. M. Towards an Articulatory Phonology. **Phonology Yearbook**, v. 3, p. 219–252, 1986.

BROWMAN, C. P.; GOLDSTEIN, L. M. Some Notes on Syllable Structure In Articulatory Phonology. **Phonetica**, v. 45, p. 140–155, 1988.

BROWMAN, C. P.; GOLDSTEIN, L. M. Articulatory gestures as phonological units. **Phonology**, v. 6, n. 2, p. 69–101, 1989a.

BROWMAN, C. P.; GOLDSTEIN, L. M. Gestural structures and phonological patterns. **Haskins Lab, Status Report, Speech**, p. 1–23, 1989b.

BROWMAN, C. P.; GOLDSTEIN, L. M. Gestural specification using dynamically-defined articulatory structures. **Journal of Phonetics**, v. 18, p. 299–320, 1990a.

BROWMAN, C. P.; GOLDSTEIN, L. M. Representation and reality: Physical systems and phonological structure. **Journal of Phonetics**, v. 18, p. 411–424, 1990b.

BROWMAN, C.; GOLDSTEIN, L. Tiers in Articulatory Phonology. In: Kingston, J.; Beckman, M. (Eds.). **Papers in Laboratory Phonology 1: between the grammar and physics of speech**. Cambridge: Cambridge University Press, 1990c. p. 341-376.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: an overview. **Phonetica**, n. 49, p. 155-180, 1992.

BROWMAN, C. P.; GOLDSTEIN, L. M. Dynamics and Articulatory Phonology. In: PORT, R. F.; GELDER, T. VAN (Eds.). . **Mind as Motion: Explorations in the Dynamics of Cognition**. Cambridge: MIT Press, 1995. p. 175–193.

BROWMAN, C.; GOLDSTEIN, L. Competing constraints on intergestural coordination and self-organization. **Les Cahiers de l'ICP**, n. 5, p. 25-34, 2000.

BROWMAN, C. P. et al. **TADA (TAsk Dynamics Application) manual**, 2006.

BYRD, D. Frontiers and challenges in Articulatory Phonology. **International Congress of Phonetic Sciences (ICPhS)**. Anais. Barcelona. 2003, p. 89-92.

CAGLIARI, L. C. **An Experimental Study of Nasality with Particular Reference to Brazilian Portuguese**. Ph.D. Dissertation, University of Edinburgh, 1977.

CALDAS NETTO, O. B. **A identificação humana pela voz: Uma realidade no Instituto de Criminalística**. Perícia Federal. Revista da Associação Nacional dos Peritos Criminais Federais, pp. 16, 18. 2003.

CATFORD, J. C. **Fundamental problems in phonetics**. Edinburgh: Edinburgh University, 1977. 278 p.

CATFORD, J. C. **A practical introduction to phonetics**. Oxford: Oxford University Press, 1988. 239 p.

CHAMPOD, C.; EVETT, I. W. Commentaries on “Some observations on the use of probability scales in forensic identification”. **Forensic Linguistics**, v. 7, n. 2, p. 238–243, 2000.

CHIBA, T.; KAJIYAMA, M. **The Vowel: Its Nature and Structure**, Tokyo: Tokyo-Kaiseikan Pub. 1941. 236 p.

CHOMSKY, N.; HALLE, M. **The sound pattern of english**, New York: Harper & Row Pub, 1968. 470 p.

CLEMENTS, N.; HUME, E. The internal organization of speech sounds. In: Goldsmith, J. (org.) **The handbook of phonological theory**. Cambridge: Blackwell, 1995, p. 245-306

CRISTÓFARO SILVA, T. **Fonética e fonologia do português**. 7ª ed., 2003. São Paulo: Contexto, 1998. 275 p.

CRYSTAL, D. **A Dictionary of Linguistics and Phonetics**. 6ª ed. Oxford, UK: Blackwell Publishing, 2008. 529 p.

DELGADO MARTINS, M.R. Análise acústica das vogais tónicas em português. **Boletim de Filologia**, tomo XXIII, 1973. Lisboa, pp. 303-314.

ELLIS. An Introduction to Signal Processing for Speech. In: Hardcastle, W. and Laver, J. (eds.), **The Handbook of Phonetic Sciences**. Oxford: Blackwell, 2011, p. 757-780.

ENGWALL, O. (n.d.). Interspeaker Variation in the Articulation of Nasal Vowels, (2003), 3–10.

ERIKSSON, A. Aural/Acoustic vs. Automatic Methods in Forensic Phonetic Case Work. In A. Neustein & H. A. Patil (Eds.), **Forensic Speaker Recognition: Law Enforcement and Counter-terrorism**. 2012. pp. 41–69.

ESCUADERO, P.; BOERSMA, P.; RAUBER, A. S.; BION, R. A. H. A Cross-Dialect Acoustic Description of Vowels: Brazilian and European Portuguese. **The Journal of the Acoustical Society of America**. 2009, volume 126(3), pp. 1379–93.

FANT, G. **Acoustic theory of speech production**. 2nd printing, 1970. Paris: Mouton. 1970[1960].

FARNETANI, E. Coarticulation and connected speech processes. In: Hardcastle, W. and Laver, J. (eds.), **The Handbook of Phonetic Sciences**. Oxford: Blackwell, 1999, pp. 371-404.

FEAGIN, C. Entering the community: fieldwork. In: CHAMBERS, J. K., TRUDGILL, P.; SCHILLING-ESTES, N. (eds.). **The handbook of language variation and change**. Malden: Blackwell, 2004, pp. 20-39.

FIGUEIREDO, R. M. **Identificação de falantes: aspectos teóricos e metodológicos**. 1994. f. Tese de doutorado (Doutorado em Ciências). Instituto de Estudos da Linguagem. UNICAMP. 1994.

FOULKES, P.; SCOBIE, J.M.; WATT, D. Sociophonetics. In: HARDCASTLE, W.; LAVER, J.; GIBBON, F.E. (Ed.). **The handbook of Phonetic Sciences**. 2^a ed., Oxford: Wiley-Blackwell, 2010.

FOWLER, C. A. Coarticulation and theories of extrinsic timing. **Journal of Phonetics**, v. 8, p. 113–133, 1980.

FOWLER, C. A. Phonological and articulatory characteristics of spoken language. **Haskins Lab. Status Rep. Speech Res., SR**, v. 109, n. 110, p. 1–12, 1992.

FOWLER, C. A., RUBIN, P. E., REMEZ, R. E., TURVEY, M. T. Implications for speech production of a general theory of action. In: BUTTERWORTH, B. (Ed.), **Language production**, 1980; p. 373-420.

FRENCH, P. An overview of forensic phonetics with particular reference to speaker identification. **International Journal of Speech, Language and the Law**, v. 1, n. 2, p. 169–181, 1994.

FRENCH, P.; NOLAN, F.; FOULKES, P.; HARRISON, P.; McDOUGALL, K. The UK position statement on forensic speaker comparison: a rejoinder to Rose and Morrison. **The International Journal of Speech, Language and the Law**, v. 17, n.1, p. 143–152, 2010.

FUJIMURA, O. **Analysys of nasal consonants**. The Journal of the Acoustical Society of America, 34(12), 1962, pp. 1865–1875.

GICK, B.; WILSON, I.; DERRICK, D. **Articulatory Phonetics**. West Sussex: Wiley-Blackwell, 2013. 272 p.

GLENN, J. W.; KLEINER, N. Speaker identification based on nasal phonation. **Journal of the Acoustical Society of America**, v. 43 n. 2, p. 368-372, 1968.

GONÇALVES, M. I. R. et al. Função de transferência das vogais orais do português brasileiro: análise acústica comparativa. **Brazilian Journal Of Otorhinolaryngology**, v. 75, n. 5, p. 680–684, 2009.

GONÇALVES, C. S. **Taxa de elocução e de articulação em corpus forense do português brasileiro**. Porto Alegre. PUCRS, 2013. 192 p.

GOLD, E., FRENCH, P. International Practices in Forensic Speaker Comparison. **International Journal of Speech Language and the Law**, 2011. 18(2). <http://doi.org/10.1558/ijssl.v18i2.293>

GOLDSTEIN, L. M.; FOWLER C. A. Articulatory Phonology: A Phonology for Public Language Use. In: SCHILLER, N. O., MEYER, A. S. (Eds.) **Phonetics and**

Phonology in Language Comprehension and Production. New York: Mouton de Gruyter, 2003, p. 159–207.

GOULD, S. J.; VRBA E. S. 1982. Exaptation - A Missing Term in the Science of Form. **Paleobiology** 8(1): 4–15.

HEUVEL, H. **Speaker variability in acoustic properties of Dutch phoneme realisations.** Nijmegen: Universidade Católica de Nijmegen. 1996. Obtido através do link "lands.let.kun.nl/literature/theses/heuvel_thesis.ps" em 20/10/2008.

HÖFKER, U. **Phoneme-ordering for speaker recognition.** Contributed Papers to the 9th International Congress on Acoustics, Madrid. Madrid: Spanish Acoustical Society. 1977.

HOLLIEN, H. **The Acoustics of Crime.** New York: Springer Science+Business Media, 1990. 370 p.

HOLLIEN, H. **Forensic voice identification.** London: Academic Press, 2002. 240 p.

HOLLIEN, H.; HUNTLEY BAH, R.; HARNSBERGER, J. D. Issues in forensic voice. **Journal of Voice**, v. 28, n. 2, p. 170–184, 2014.

HOLLIEN, H. An Approach to Speaker Identification. **Journal of Forensic Sciences**, v. 61, n. 2, p. 334–344, 2016.

HYMES, Dell. **Foundations in sociolinguistics: an ethnographic approach.** New Jersey: University of Pennsylvania Press, 1974. 247p.

INTERNATIONAL PHONETIC ASSOCIATION. **Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet.** 10th Edition, 2009. Cambridge: Cambridge University Press, 1999. 204 p.

ISKAROUS, K. **Vowel constrictions are recoverable from formants.** Journal of Phonetics, n.38, pp. 375–387. 2010.

ISKAROUS, K. et al. **CASY: The Haskins Configurable Articulatory Synthesizer.** 15th ICPHS. Barcelona: 2003, pp. 185-188.

JAKOBSON, R.; HALLE, M. **Fundamentals of language.** Leiden: Mouton & Co., 1956. 87p.

JOHNSON, K. **Acoustic and Auditory Phonetics.** Cambridge: Blackwell Publishers, 1^a ed. 1997. 169 p. 2003 (10^a ed.).

JESUS, M. de S. V. **Estudo fonético da nasalidade vocálica em falantes normais e com fissura de palato: enfoque acústico.** Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Belo Horizonte. 1999.

KENT, R. D.; READ, C. **The acoustic Analysis of speech**. San Diego: Singular, 1992.

KENT, R. D.; KIM, Y. Toward an acoustic typology of motor speech disorders. **Clinical linguistics & phonetics**, v. 17, n. 6, p. 427–445, 2003.

KERSTA, L.G. **Voiceprint identification**. *Nature*, v. 196, n. 1253-1257. 1962.

KINOSHITA, Y.; ISHIHARA, S. Background population: How does it affect LR-based forensic voice comparison? **International Journal of Speech Language and the Law**, v. 21, n. 2, p. 191–224, 2015.

KOENIG, L. L. **Towards a physical definition of the vowel systems of languages** *Hard-science linguistics*, 2004, p. 49-66.

KRAKOW, R. A.; HUFFMAN, M. K. Instruments and techniques for investigating nasalization and velopharyngeal function in the laboratory: an introduction. In: HUFFMAN, M. K.; KRAKOW, R. A. (Eds.). **Nasals, nasalization, and the velum. Phonetics and Phonology Series, Vol. 5**. San Diego, CA: Academic Press, 1993, p. 3-59.

KÜNZEL, H. Current approaches to forensic speaker recognition. **Automatic Speaker Recognition, Identification and Verification**, n. April, p. 135–141, 1994.

KUROWSKI, K. M.; BLUMSTEIN, S. E. Acoustic properties for the perception of nasal consonants. In: HUFFMAN, M. K.; KRAKOW, R. A. (Eds.). **Nasals, nasalization, and the velum. Phonetics and Phonology Series, Vol. 5**. San Diego, CA: Academic Press, 1993, p. 197-221.

LABOV, W. **Sociolinguistics Patterns**. Philadelphia: University of Pennsylvania, 1972.

LABOV, W.; YAGER, M.; STEINER, R. **A quantitative study of sound change in progress**. Philadelphia: U.S. Regional Survey, 1972. 337 p.

LADEFOGED, P. **A course in phonetics**. New York: Harcourt Brace Jovanovich, 1975. 300p.

LADEFOGED, P. **Elements of acoustic phonetics**. Chicago: University of Chicago, 1996. 216 p.

LADEFOGED, P.; MADDIESON, I. **The sounds of the world's languages**. 11th ed., 2008. Malden: Blackwell Publishing, 1996. 450 p.

LAVER, J. **Principles of phonetics**. Cambridge: Cambridge University Press, 1994. 707p.

LIBERMAN, P.; BLUMSTEIN, S. **Speech Physiology, Speech Perception, and Acoustic Phonetics**. Cambridge: Cambridge University Press, 1988. 249 p.

LIMA, R. **Análise acústica das vogais orais do português de Florianópolis - Santa Catarina**. Florianópolis. Universidade Federal de Santa Catarina, 1991.

LINDBLOM, B. Phonetic universals in vowel systems. In: Ohala, J. and Jaeger, J. (eds.). **Experimental phonology**, New York: Academic, 1986, pp 13-44.

MCCLOY, D.; MCGRATH, A. **Semi-auto formant extractor**. Praat script. Obtido no endereço eletrônico <<https://github.com/drammock/praat-semiauto/blob/master/SemiAutoFormantExtractor.praat>> em 09/08/2015.

MCDUGALL, K. **The Role of Formant Dynamics in Determining Speaker Identity**. Tese de doutorado. Universidade de Cambridge. 2005

MCMENAMIN, G. **Forensic linguistics: advances in forensic stylistics**. London: CRC Press. 2002. 341 p.

MIRANDA, I. **Análise acústico-comparativa de vogais do português brasileiro com vogais do inglês norte-americano**. Dissertação de mestrado. Universidade Federal do Espírito Santo, 2012. 147 p.

MIRANDA, I.; MEIRELES, A. Descrição acústica das vogais tônicas da fala capixaba. **Letras de Hoje**. V.47, n. 3. Porto Alegre. 2012, pp. 325-332

MITRA, V. et al. Retrieving Tract Variables From Acoustics: A Comparison of Different Machine Learning Strategies. **IEEE journal of selected topics in signal processing**, v. 4, n. 6, p. 1027–1045, 2010.

MITRA, V. et al. Articulatory information for noise robust speech recognition. **IEEE Transactions on Audio, Speech and Language Processing**, v. 19, n. 7, p. 1913–1924, 2011.

MORAES, J.; CALLOU, D.; LEITE Y. O sistema vocálico do português do Brasil: caracterização acústica. In: KATO, Mary A. (Org.). **Gramática do português falado**, Volume V; Convergências. Campinas: UNICAMP, 1996, pp. 33-53.

MORRISON, G. S. Distinguishing between forensic science and forensic pseudoscience: Testing of validity and reliability, and approaches to forensic voice comparison. **Science and Justice**, 2014. 54(3), 245–256. <http://doi.org/10.1016/j.scijus.2013.07.004>

NAM, H. et al. A procedure for estimating gestural scores from articulatory data. **The Journal of the Acoustical Society of America**, v. 127, n. 3, p. 1851, 2010.

NAM, H. et al. A procedure for estimating gestural scores from speech acoustics. **The Journal of the Acoustical Society of America**, v. 132, n. 6, p. 3980–9, 2012.

NOLAN, F. **The phonetic bases of speaker recognition**. Cambridge: Cambridge University Press, 1983. 221 p.

NOLAN, F. Forensic phonetics. **Journal of Linguistics**. Cambridge: Cambridge University Press, 1991, v.27, n. 2, p. 483–493.

NOLAN, F. Auditory and acoustic analysis in speaker recognition. In: J. Gibbons (ed.), **Language and the Law**. London: Longman., 1993, v.27, n. 2, p. 483–493.

NOLAN, F. Can the definition of each speaker be expected to come from the laboratory in the next decades? **International Congress of Phonetic Sciences (ICPhS). Anais**. Stockholm, 1995, p. 130-137.

NOLAN, F.; Speaker recognition and forensic phonetics. In: Hardcastle, W. and Laver, J. (eds.), **A Handbook of Phonetic Science**; Oxford: Blackwell, 1997.

NOLAN, F. The "telephone effect" on formants: a response. **Forensic Linguistics**. 2002, v.9, n.1, p. 74–82.

NOLAN, F.; GRIGORAS, C. (2005). A case for formant analysis in forensic speaker identification. **International Journal of Speech, Language and the Law**, 12(2), 143–173. <http://doi.org/10.1558/sll.2005.12.2.143>

NOLAN, F., MCDOUGALL, K., JONG, G.; HUDSON, T. (2006). A Forensic Phonetic Study of " Dynamic " Sources of Variability in Speech: The DyViS Project In: P. Warren & C. I. Watson (Eds.), **11th Australian International Conference on Speech Science and Technology** (pp. 13–18).

OHALA, J. J.; EUKEL, B. W. Explaining the Intrinsic Pitch of Vowels. In: CHANNON, R.; SHOCKEY, L. (Eds.). **In Honor of Ilse Lehiste**. Dordrecht, Holanda: Foris Publications, 1987, p. 207-215.

OHALA, J. J. There is no interface between phonology and phonetics: a personal view. **Journal of Phonetics**, n. 18, p. 153–171, 1990.

OHALA, J. J.; OHALA, M. The phonetics of nasal phonology: theorems and data. In: HUFFMAN, M. K.; KRAKOW, R. A. (Eds.). **Nasals, nasalization, and the velum. Phonetics and Phonology Series, Vol. 5**. San Diego, CA: Academic Press, 1993, p. 225-249.

PANCHAPAGESAN, S.; ALWAN, A. A study of acoustic-to-articulatory inversion of speech by analysis-by-synthesis using chain matrices and the Maeda articulatory model. **The Journal of the Acoustical Society of America**, v. 129, n. 4, p. 2144–2162, 2011.

PEREIRA, A. L. D. **Caracterização acústica do sistema vocálico tônico oral florianopolitano: alguns indícios de mudança**. Dissertação de mestrado. Florianópolis. Universidade Federal de Santa Catarina, 2001.

PETTINATO, M. et al. Vowel space area in later childhood and adolescence: Effects of age, sex and ease of communication. **Journal of Phonetics**, v. 54, p. 1–14, 2016.

RAUBER, A. S. **Perception and production of english vowels by brazilian efl speakers**. Florianópolis. Universidade Federal de Santa Catarina, Tese de doutorado. 2006. 157 p.

REETZ, H.; JONGMAN A. **Phonetics: transcription, production, acoustics and perception**. West Sussex. Blackwell. 2009. 316p.

ROSE, P. **Forensic speaker identification**. London: Taylor & Francis, 2002. 364 p.

ROSE, P. Forensic speaker discrimination with australian english vowel acoustics. International Congress of Phonetic Sciences (ICPhS). **Anais**. Saarbrücken: 2007.

RUBIN, P.; BAER, T.; MERMELSTEIN, P.. An articulatory synthesizer for perceptual research. **Journal of the Acoustical Society of America**, n.70, 1981. p. 321-328.

SALTZMAN, E. The task dynamic coordination of the speech articulators: a preliminary model. **Experimental Brain Research Series**. v.15, 1986. p. 129-144.

SALTZMAN, E., & KELSO, J. A. S. Skilled actions: A task dynamic approach. **Psychological Review**, 94. 1987, p. 84-106.

SANTOS, G. B. DOS. **Análise fonético-acústica das vogais orais e nasais do português: Brasil e Portugal**. Goiânia. Universidade Federal de Goiás, 2013.

SAPIR, S. et al. Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech. **J Speech Lang Hear Res**, v. 53, n. 1, p. 114–125, 2010.

SAUSSURE, F. **Curso de lingüística geral**. 20a ed. São Paulo: Cultrix, 1995[1970]. 279 p.

SEARA, I. C. **Estudo acústico-perceptual da nasalidade das vogais do português brasileiro**. Tese de doutorado. Universidade Federal de Santa Catarina. 2000.

SCHILLING-ESTES, N. Stylistic variation and the sociolinguistic interview: A reconsideration. **Proceedings from AESLA 25**. Ed. Rafael Monroy and Aquilino Sánchez. Murcia, Spain: Servicio de Publicaciones de la Universidad de Murcia, 2008.

SHANNON, C. A mathematical theory of communication. **Bell System Technical Journal**. 1948, v.27 n.3: p. 379–423.

SILVA, A. H. P. **As fronteiras entre Fonética e Fonologia e a alofonia dos róticos iniciais em PB: dados de dois informantes do sul do país**. Tese de doutorado. UNICAMP, 2002. 202p.

SILVA, A. H. P. Pela incorporação de informação fonética aos modelos fonológicos. **Revista Letras**, v. 60, jul./dez., p. 319–333, 2003.

SILVA, A. H. P. Primitivos fonológicos de tempo extrínseco vs primitivos de tempo intrínseco. **Fórum Linguístico**, v.5, n.1, pp. 1–12, 2008.

SOUSA, E. M. G. **Para a caracterização fonético-acústica da nasalidade no português do Brasil**. Dissertação de mestrado. Campinas. Universidade Estadual de Campinas, 1994. 170 p.

STEVENS, K. N. **Acoustic phonetics**. Cambridge: MIT Press, 1998. 607p.

STEVENS, K. N.; HOUSE, A. S. An acoustical theory of vowel production and some of its implications. In: Fry, D. B. (ed), **Acoustic phonetics**. Cambridge: Cambridge University, 1976, pp. 52-74.

STONE, M. Laboratory Techniques for Investigating Speech Articulation. In: Hardcastle, W. and Laver, J. (eds.), **The Handbook of Phonetic Sciences**. Oxford: Blackwell, 2011, p. 9-38.

SU, L., LI, K. P., FU, K. S., & JONES, D. Identification of speakers by use of nasal coarticulation. **Journal of Acoustic Society of America**. 1975, pp. 1876–1882.

THOMAS, E. R. **Sociophonetics: an introduction**. Hampshire: Palgrave Macmillan, 2011. 356 p.

TRUBETZKOY, N. **Principles de phonologie**. Paris, Klincksieck, 1964 [1939].

TURVEY, M. T. **Preliminaries to a theory of action with reference to vision**. (R. Shaw, J. Bransford, Eds.) Perceiving, acting and knowing: Toward an ecological psychology. **Anais**. Hillsdale, NJ: Erlbaum, 1977. 492 p.

WARNER, N. Methods for studying spontaneous speech. Invited chapter in: A. Cohn, C. Fougeron, & M. Huffman (eds.). **The Oxford Handbook of Laboratory Phonology**. Oxford: Oxford University Press. 2012, p. 621-633.

WOLF, J. J. Efficient acoustic parameters for speaker recognition. **Journal of the Acoustical Society of America**, v. 51 n. 6, pp. 2044-2056, 1972.

YEHIA, H. C. **A study on the speech acoustic-to-articulatory mapping using morphological constraints**. Tese de doutorado. Universidade de Nagoya. 1997.

APÊNDICE A – Termo de consentimento livre e esclarecido

O Sr.(a) está sendo convidado(a) como voluntário(a) a participar da pesquisa **Variabilidade individual de parâmetros espectrais de curto termo em contóides nasais e vocóides orais como elemento para a identificação forense**, que tem como objetivo estudar parâmetros acústicos relacionados à perícia de Comparação de Locutor.

A sua participação no referido estudo será no sentido de fornecer amostras de fala com base em lista de cinquenta e três sílabas a serem pronunciadas dentro da seguinte frase: Ele dirá {sílabas} calmamente. Serão realizadas três gravações diárias, nos turnos da manhã, tarde e noite, em sete dias diferentes, totalizando, pelo menos, vinte e uma gravações. Além dessas gravações, será realizada uma entrevista de experiência pessoal, também gravada, e sem direcionamento de tópico, pois o alvo de interesse não é o conteúdo da fala e sim aspectos sonoros das sentenças proferidas. Ressaltamos que não são verificados riscos ou benefícios de qualquer natureza aos participantes.

Para participar deste estudo você não terá nenhum custo, nem receberá qualquer vantagem financeira. Você será esclarecido (a) sobre o estudo em qualquer aspecto que desejar e estará livre para participar ou recusar-se a participar. Poderá retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa em participar não acarretará qualquer penalidade ou modificação na forma em que é atendido pelo pesquisador. O pesquisador irá tratar a sua identidade com padrões profissionais de sigilo.

Os resultados da pesquisa estarão à sua disposição quando finalizada. Seu nome ou o material que indique sua participação não será liberado sem a sua permissão. O Sr não será identificado em nenhuma publicação que possa resultar deste estudo.

Este termo de consentimento encontra-se impresso em duas vias, sendo que uma cópia será arquivada pelo pesquisador responsável e a outra será fornecida a você.

Eu, «Nome», «idade» anos, «est_civil», «profissão», portador do documento de identidade «RG», residente à «endereço», fui informado dos objetivos do estudo denominado “Variabilidade individual de parâmetros espectrais de curto termo em contóides nasais e vocóides como elemento para a identificação forense” de forma clara e detalhada e esclareci minhas dúvidas. Sei que a qualquer momento poderei solicitar novas informações e modificar minha decisão de participar.

Declaro que concordo em participar desse estudo. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada à oportunidade de ler e esclarecer as minhas dúvidas.

Porto Alegre, _____ de _____ de 2014.

Nome Assinatura participante

Data

Nome Assinatura do pesquisador responsável

Data

Nome Assinatura do pesquisador responsável

Data

Em caso de dúvidas com respeito aos aspectos éticos deste estudo, você poderá consultar:

Comitê de Ética em Pesquisa
 Av. Ipiranga 6681, prédio 40 sala 505
 CEP 90619-900
 Porto Alegre, RS
 Fone: 55 51 3320.3345
 Horário de Funcionamento: das 8h às 12h e das 13h30min às 17h
 Site: www.pucrs.br/propesq
 Email: cep@pucrs.br

Profa. Dra. Cláudia Regina Brescancini (orientadora) (51) 3320-3676 ramal 8274
 Márcio Oppliger Pinto (pesquisador) 51) 9266-9943

APÊNDICE B – Termo de compromisso para utilização de dados



Pontifícia Universidade Católica do Rio Grande do Sul
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COMITÊ DE ÉTICA EM PESQUISA

TERMO DE COMPROMISSO PARA UTILIZAÇÃO DE DADOS

Título do Projeto

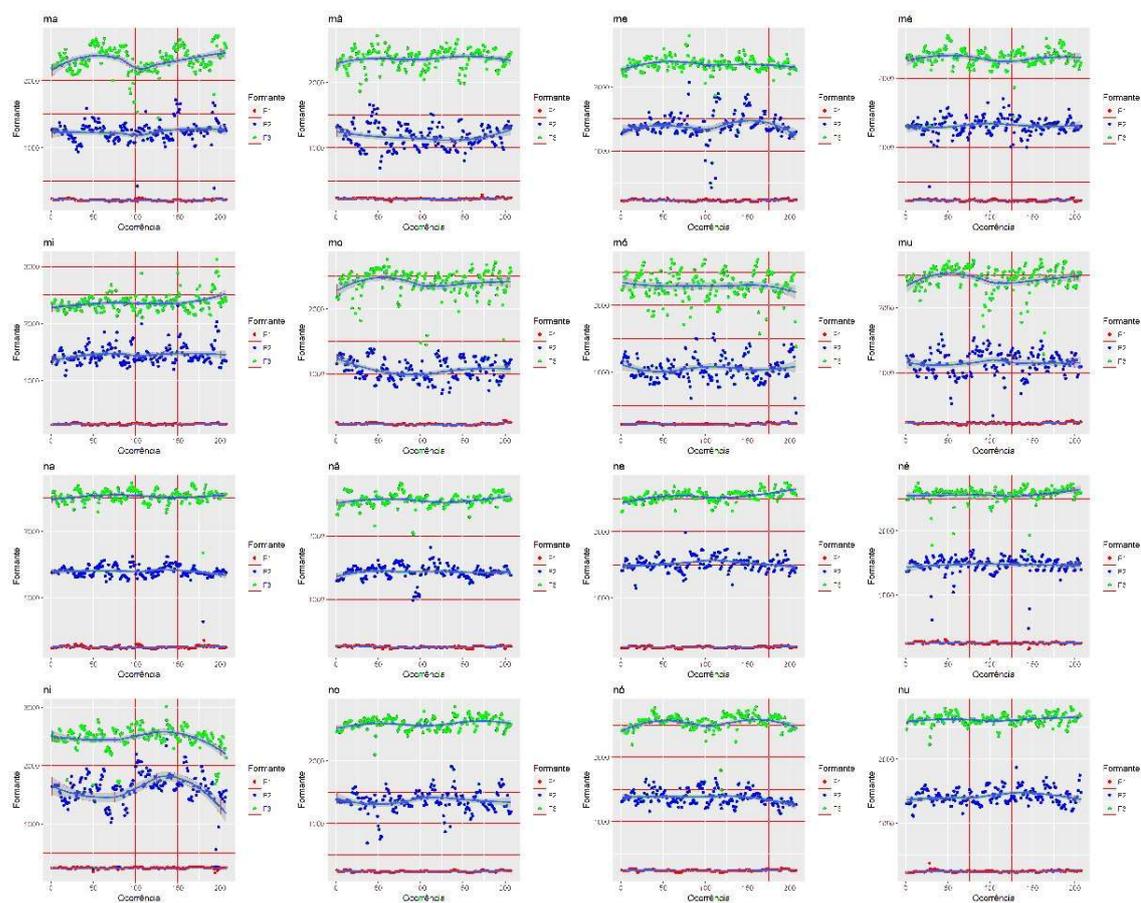
Variabilidade individual de parâmetros espectrais de curto termo em contóides nasais e vocóides orais como elemento para a identificação forense.

Os autores do projeto de pesquisa comprometem-se a manter o sigilo dos dados coletados em bases de dados referentes aos participantes da pesquisa intitulada *Variabilidade individual de parâmetros espectrais de curto termo em contóides nasais e vocóides orais como elemento para a identificação forense*. Os pesquisadores concordam, igualmente, que estas informações serão utilizadas única e exclusivamente com finalidade científica, preservando-se integralmente o anonimato dos participantes.

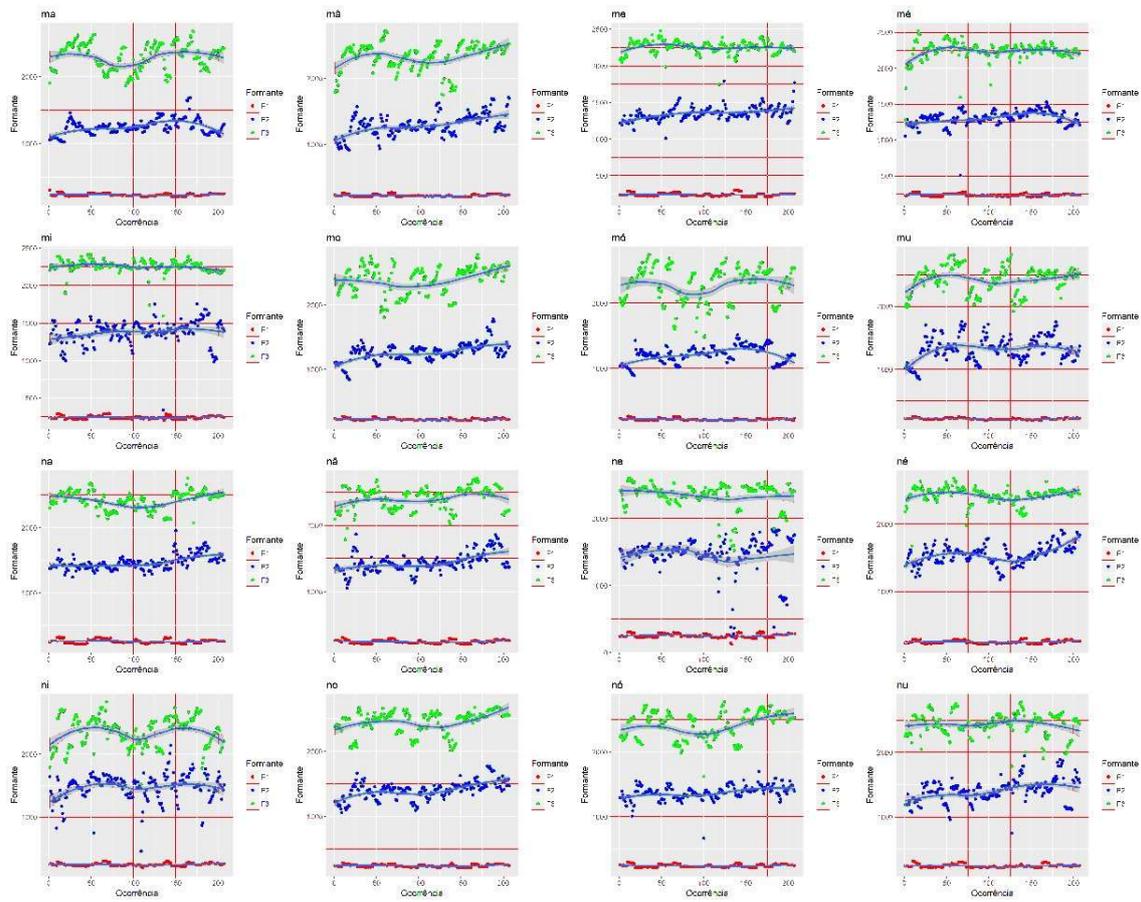
Porto Alegre, 19 de maio de 2014.

Autores do Projeto	
Nome	Assinatura
Cláudia Regina Bresciani	
MARCIO OPPLIGER PINTO	

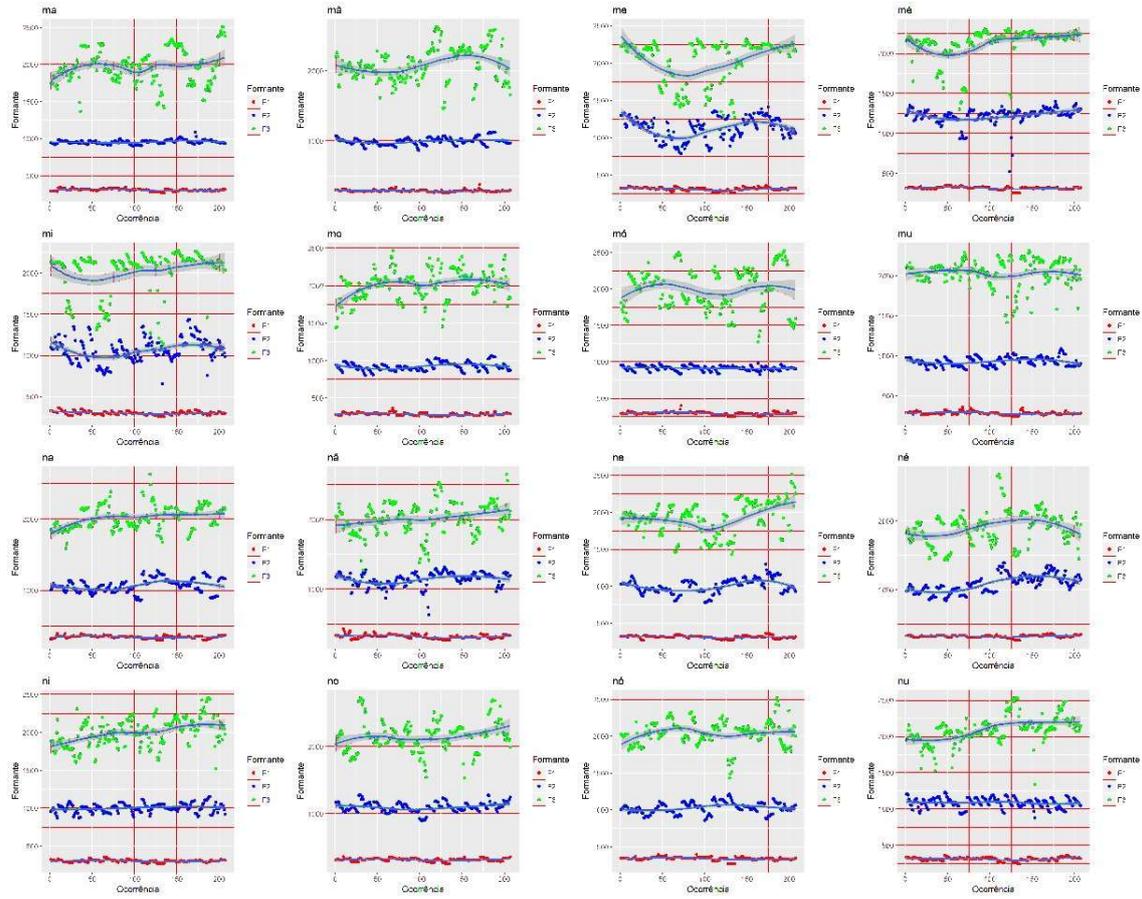
APÊNDICE D - Gráficos de distribuição dos formantes das consoantes nasais
Participante ACRR



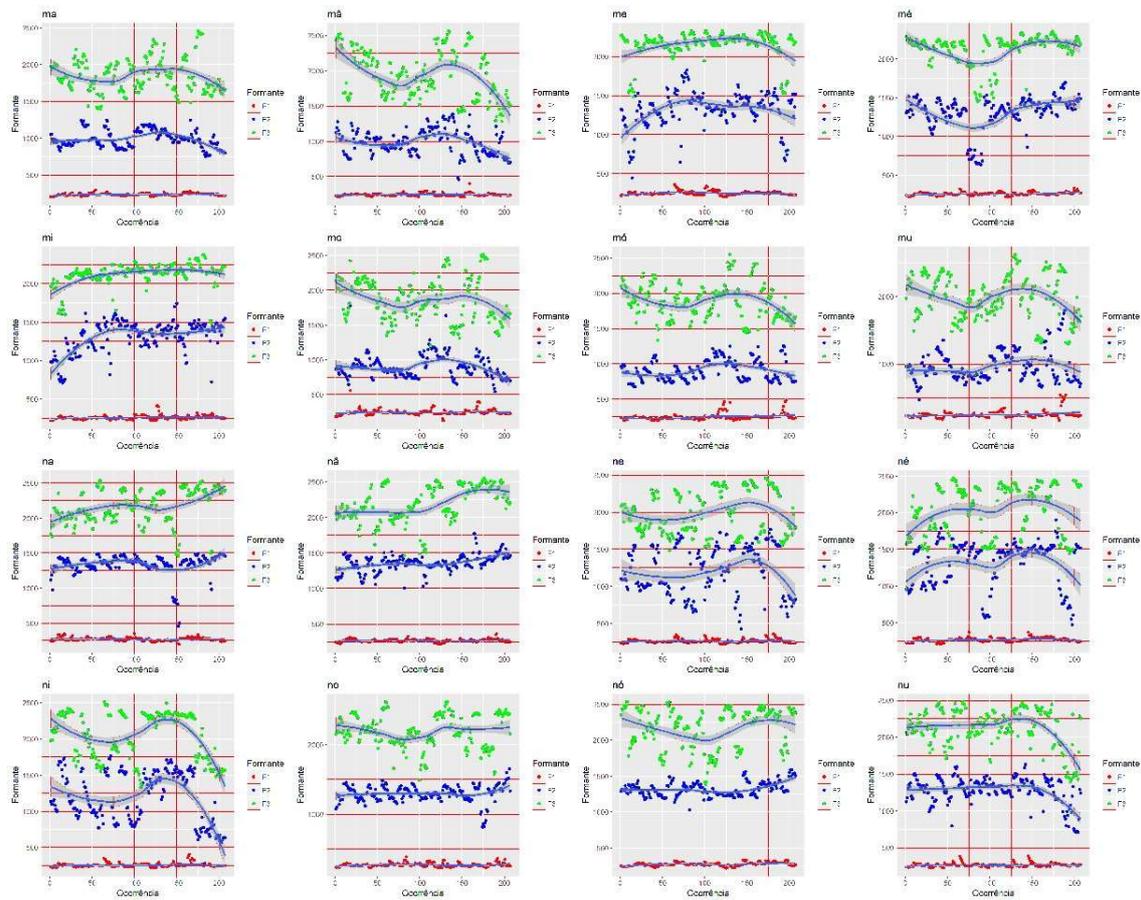
Participante AFM



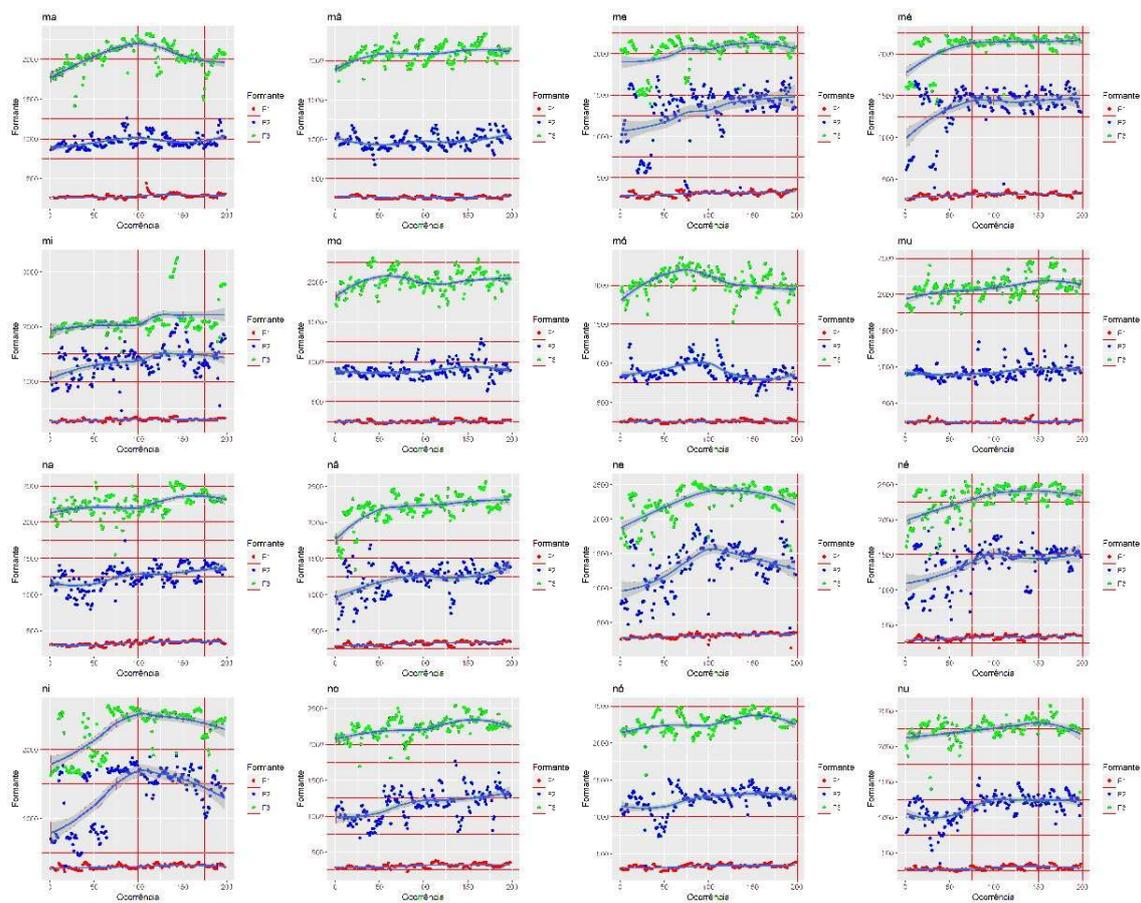
Participante APG



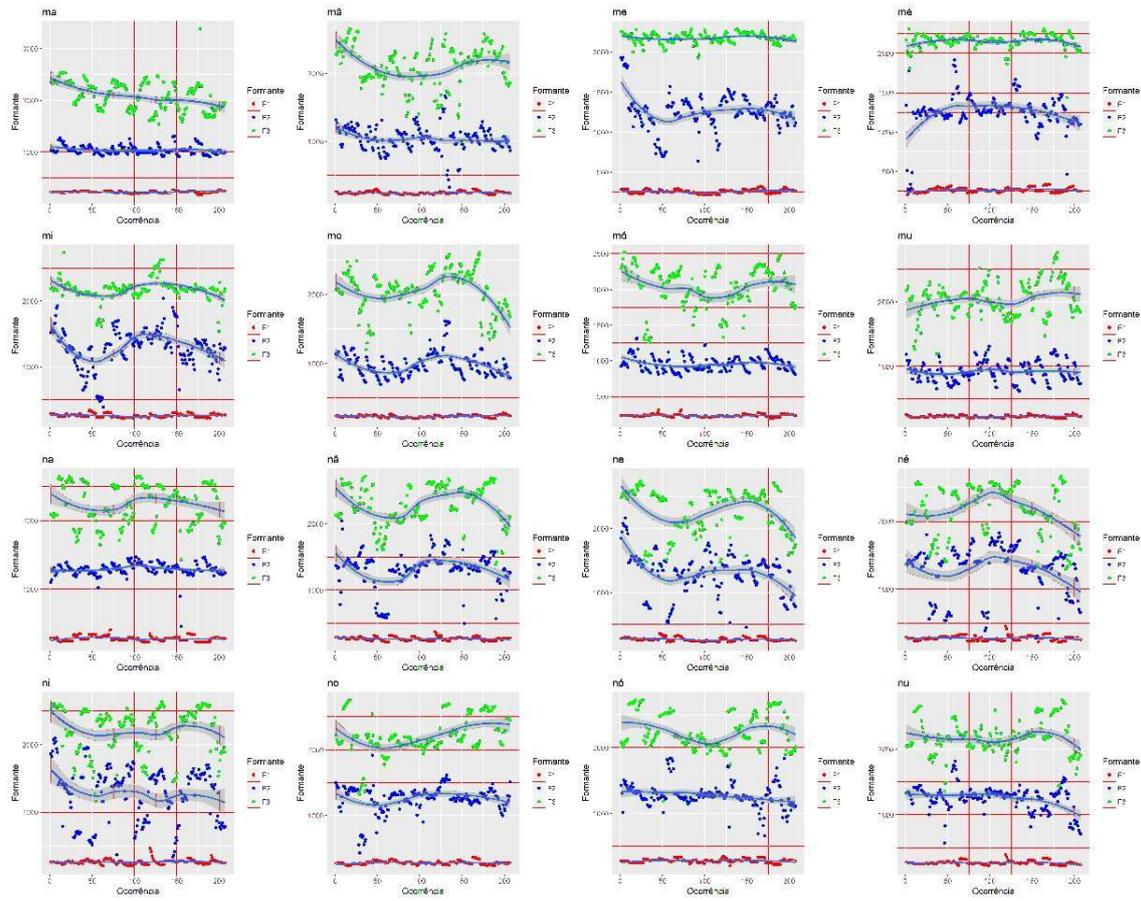
Participante CAPM



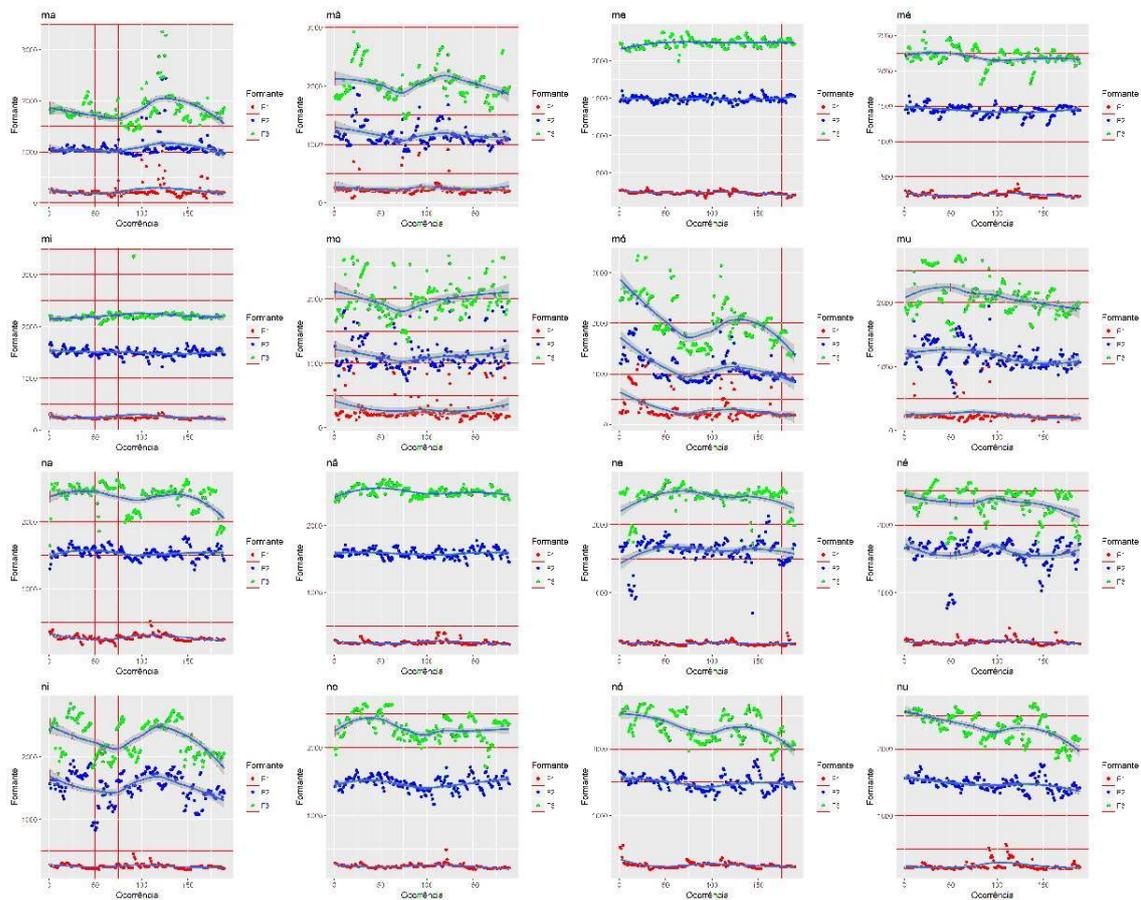
Participante CDK



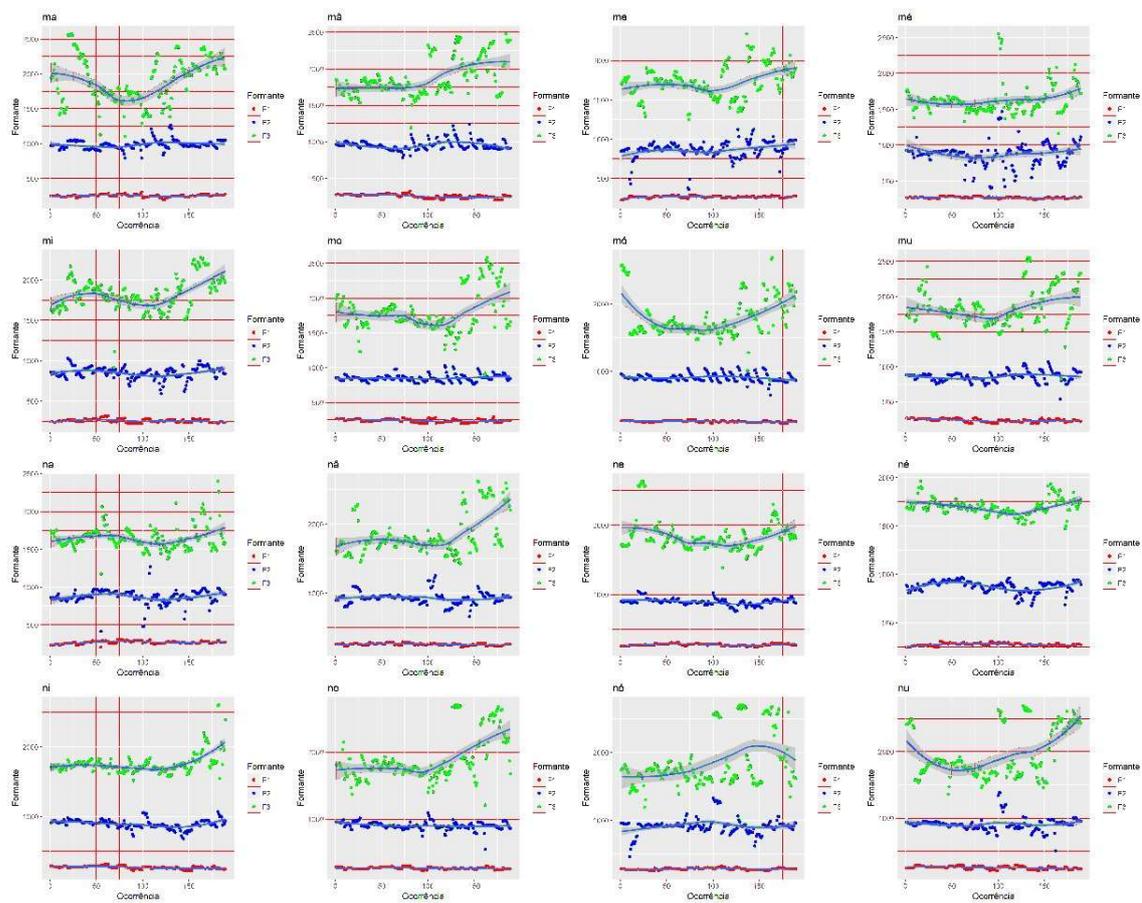
Participante CRD



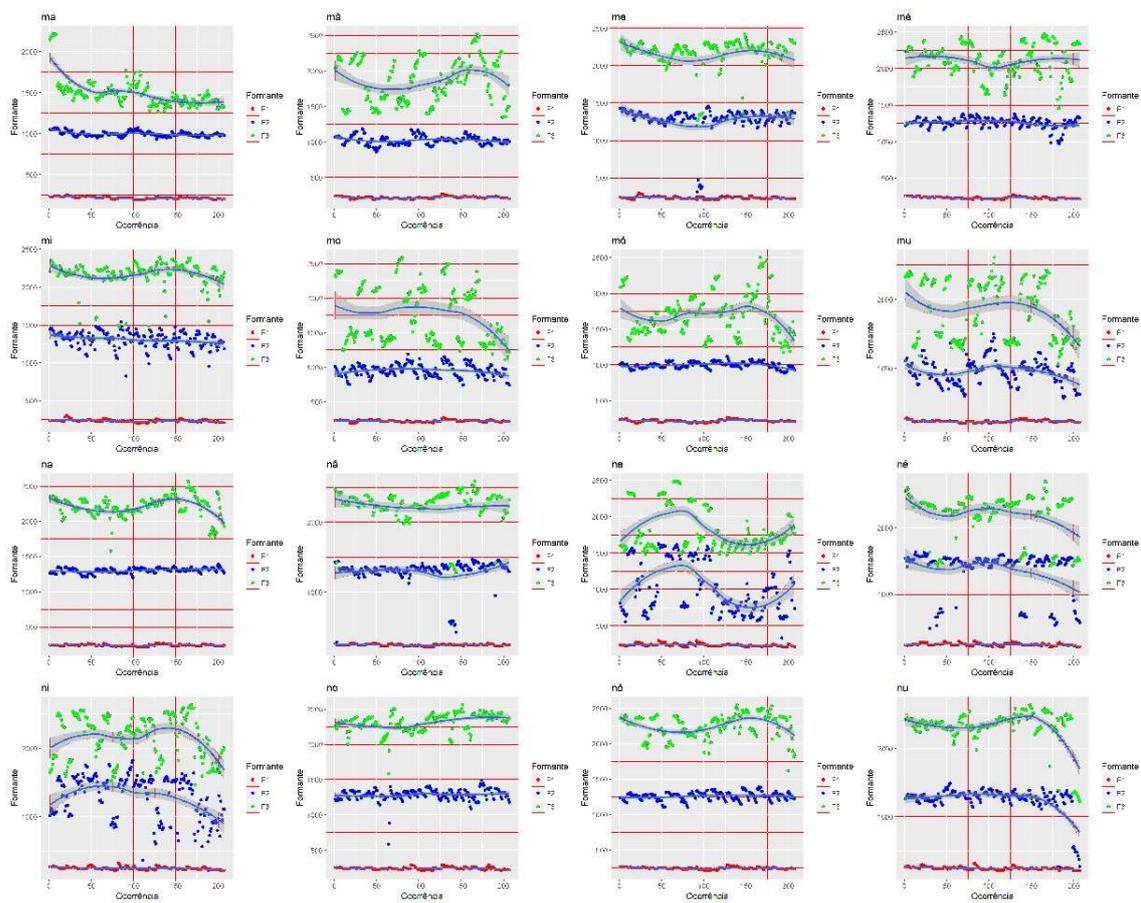
Participante EGS



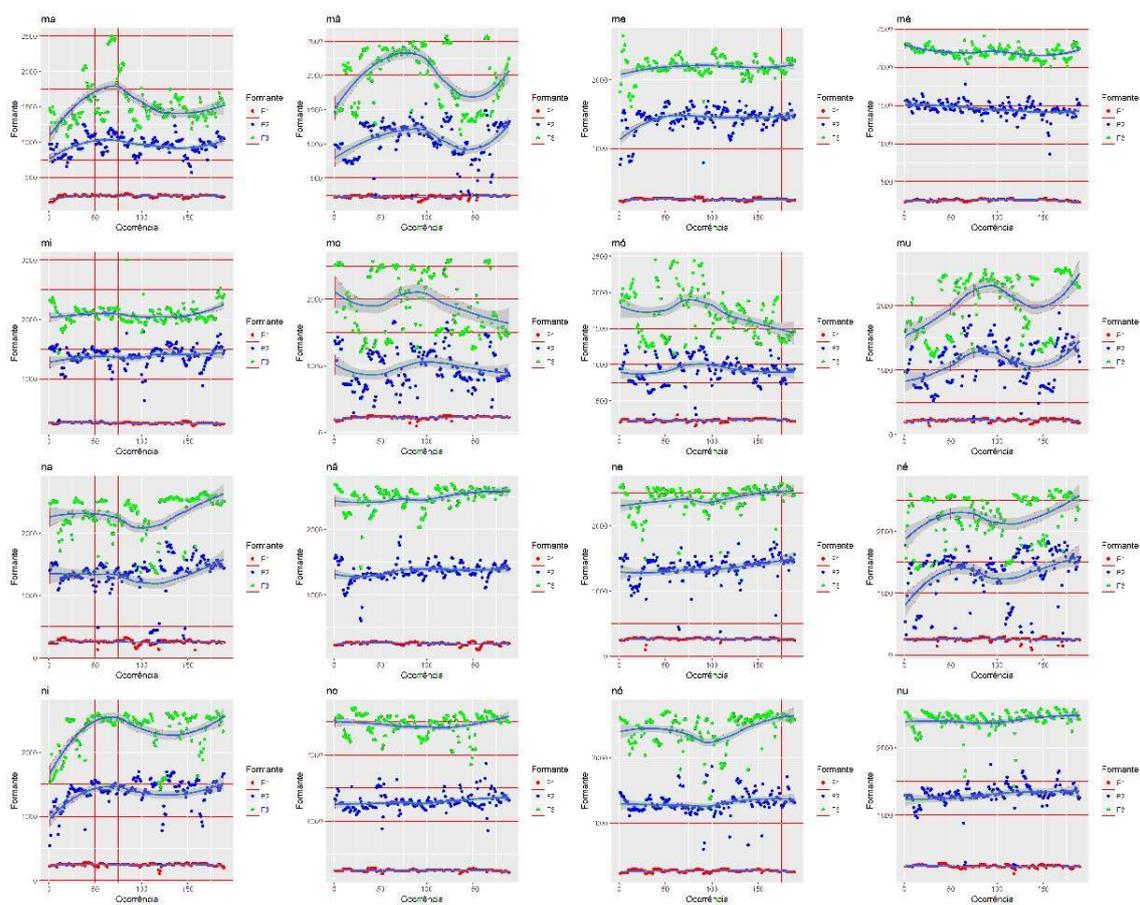
Participante FTC



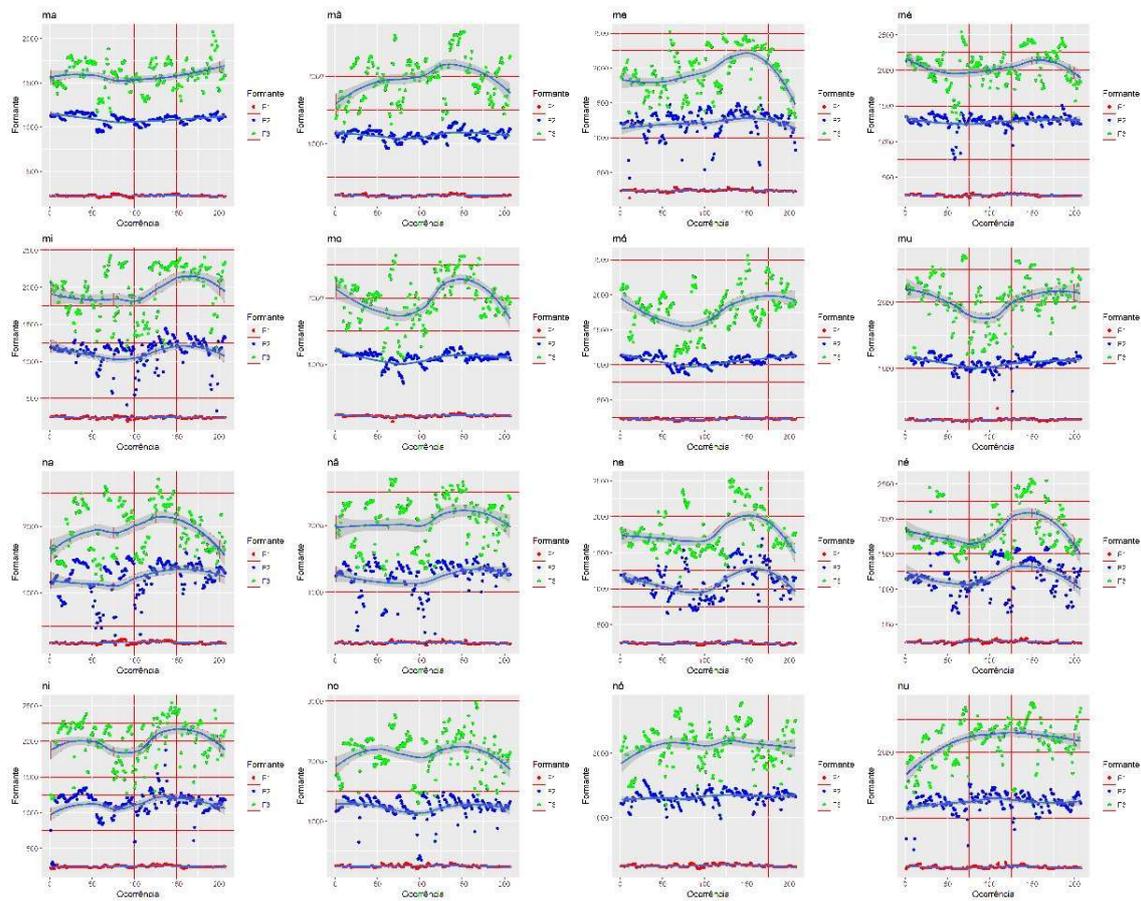
Participante IAB



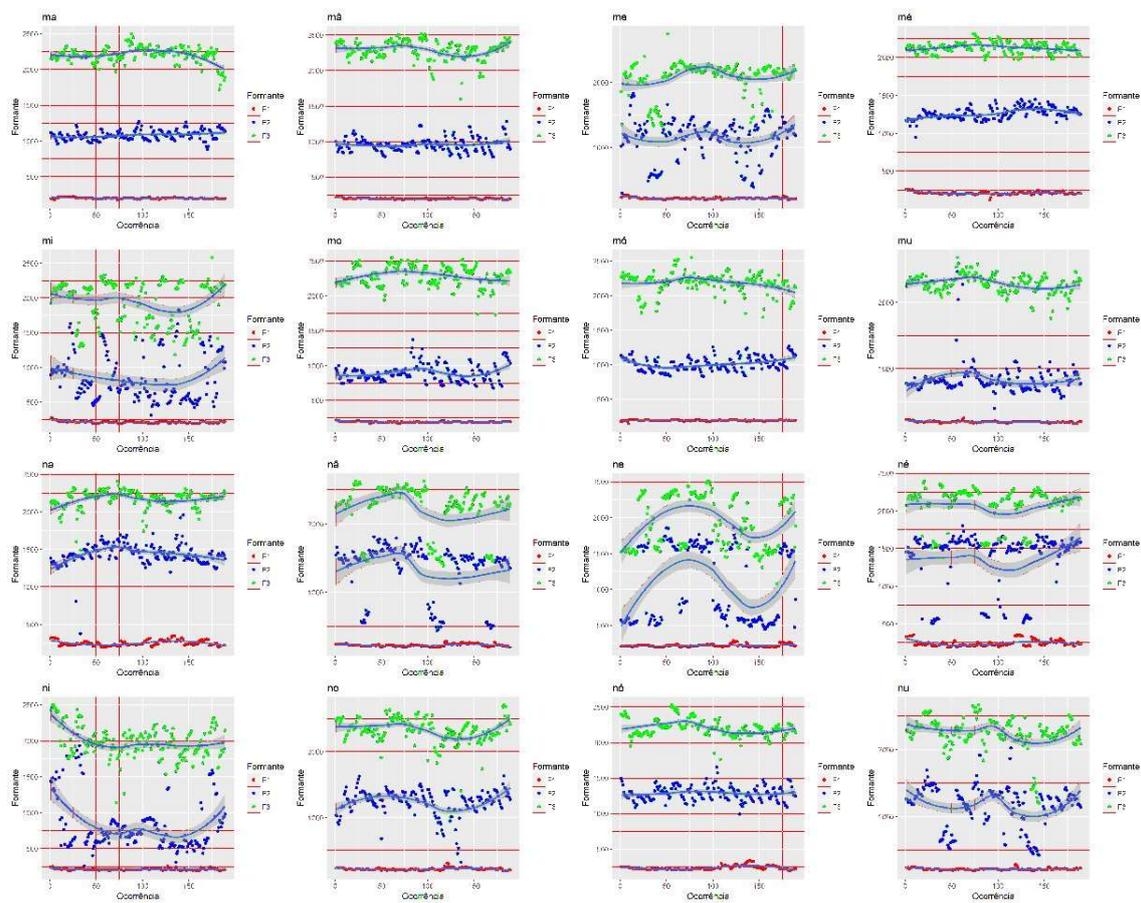
Participante MPG



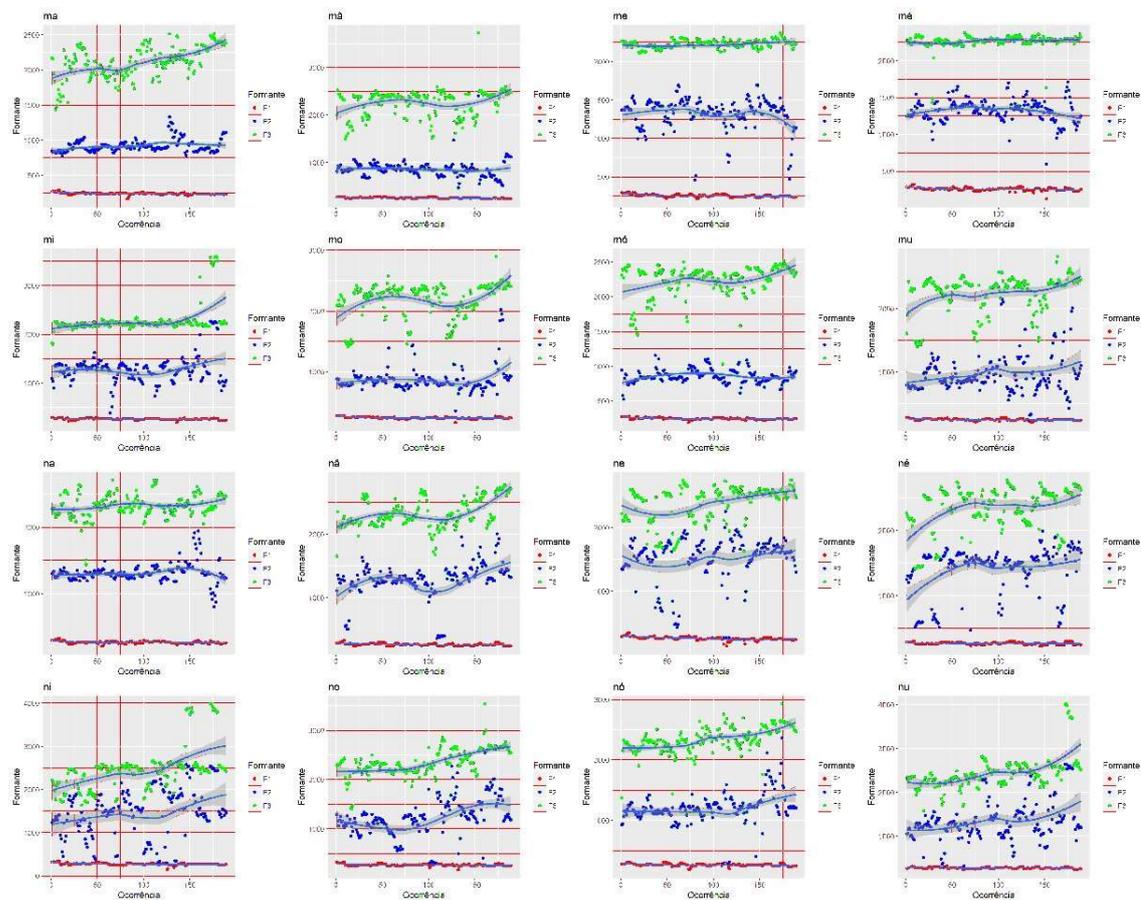
Participante MRR



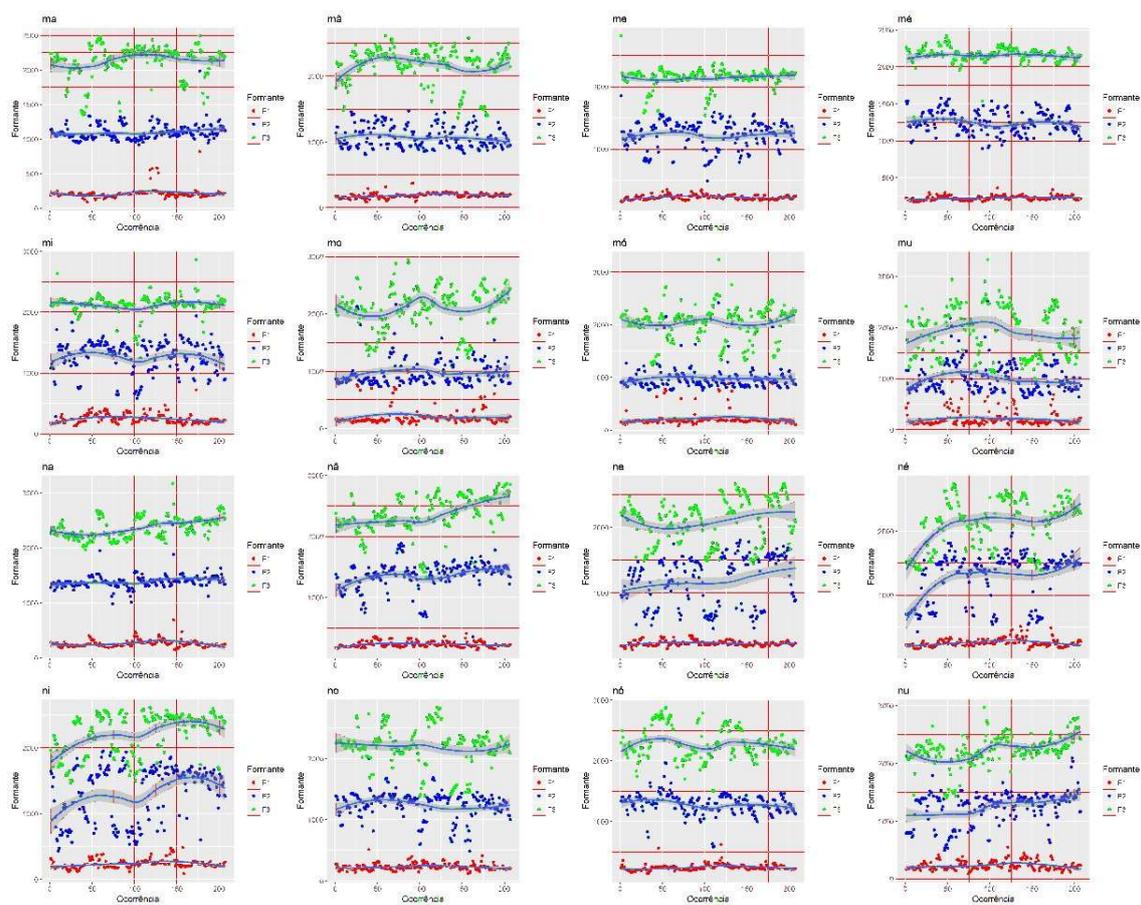
Participante MZ



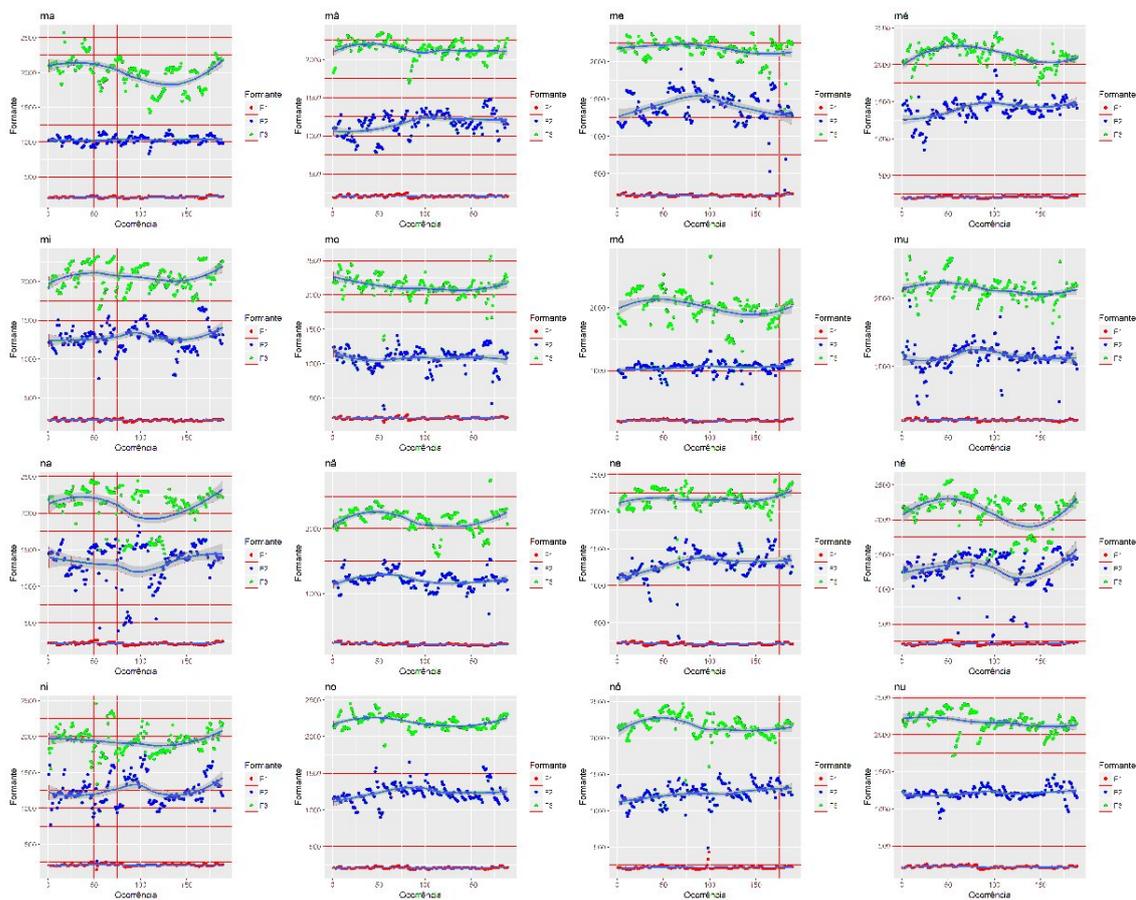
Participante PMP



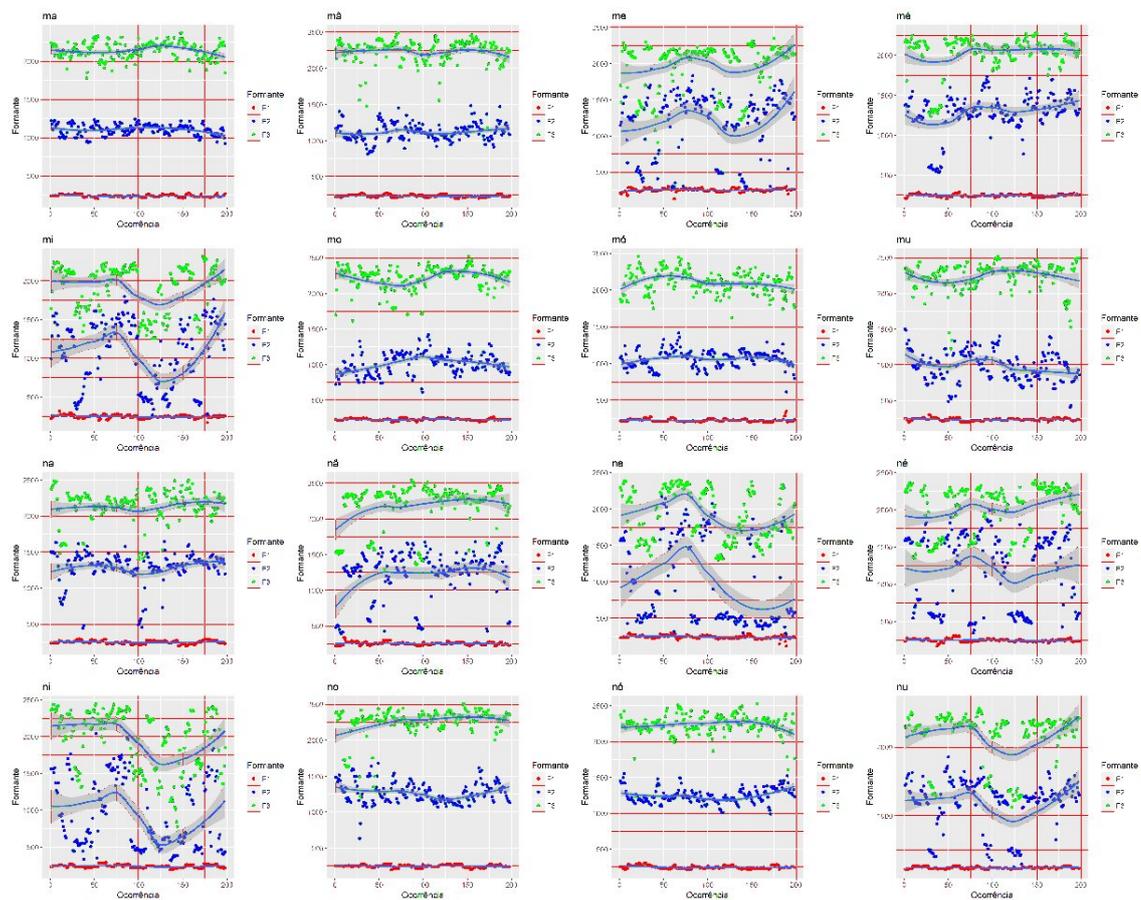
Participante PROF



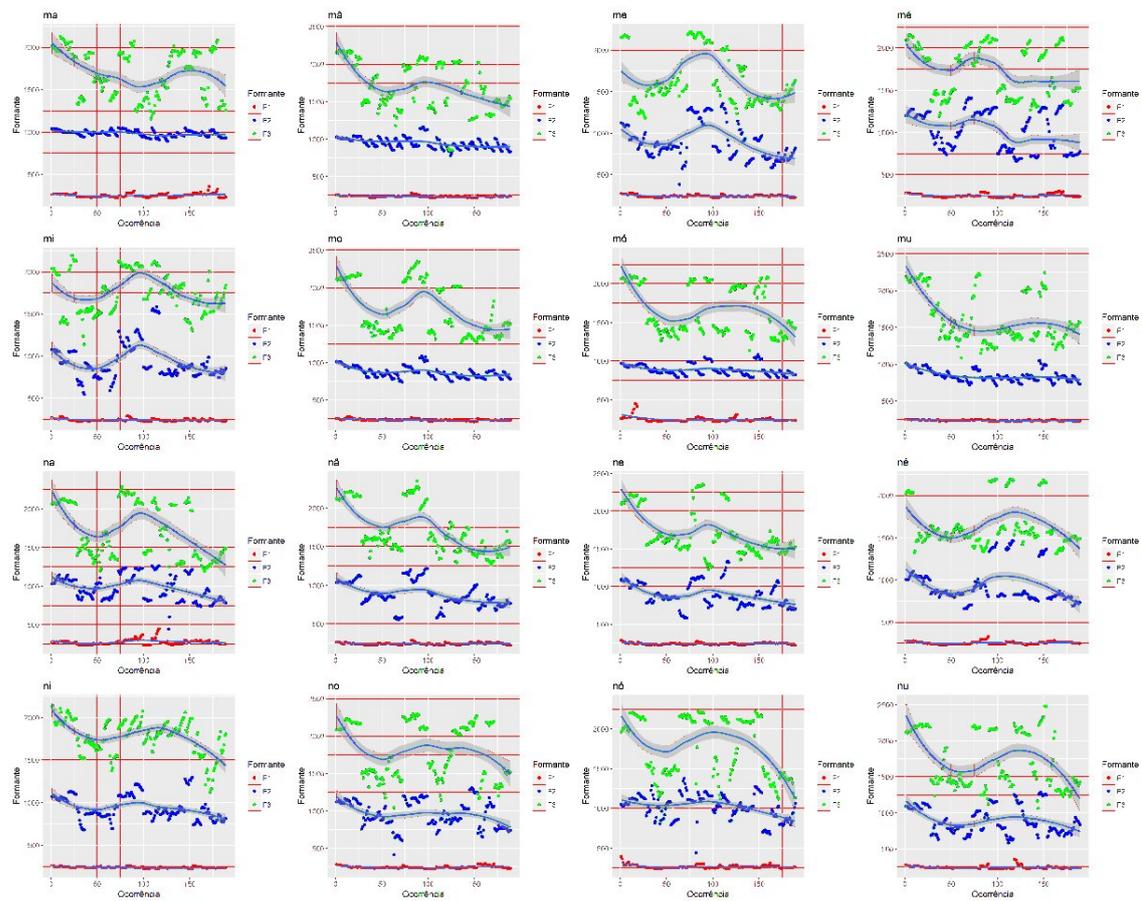
Participante RCS



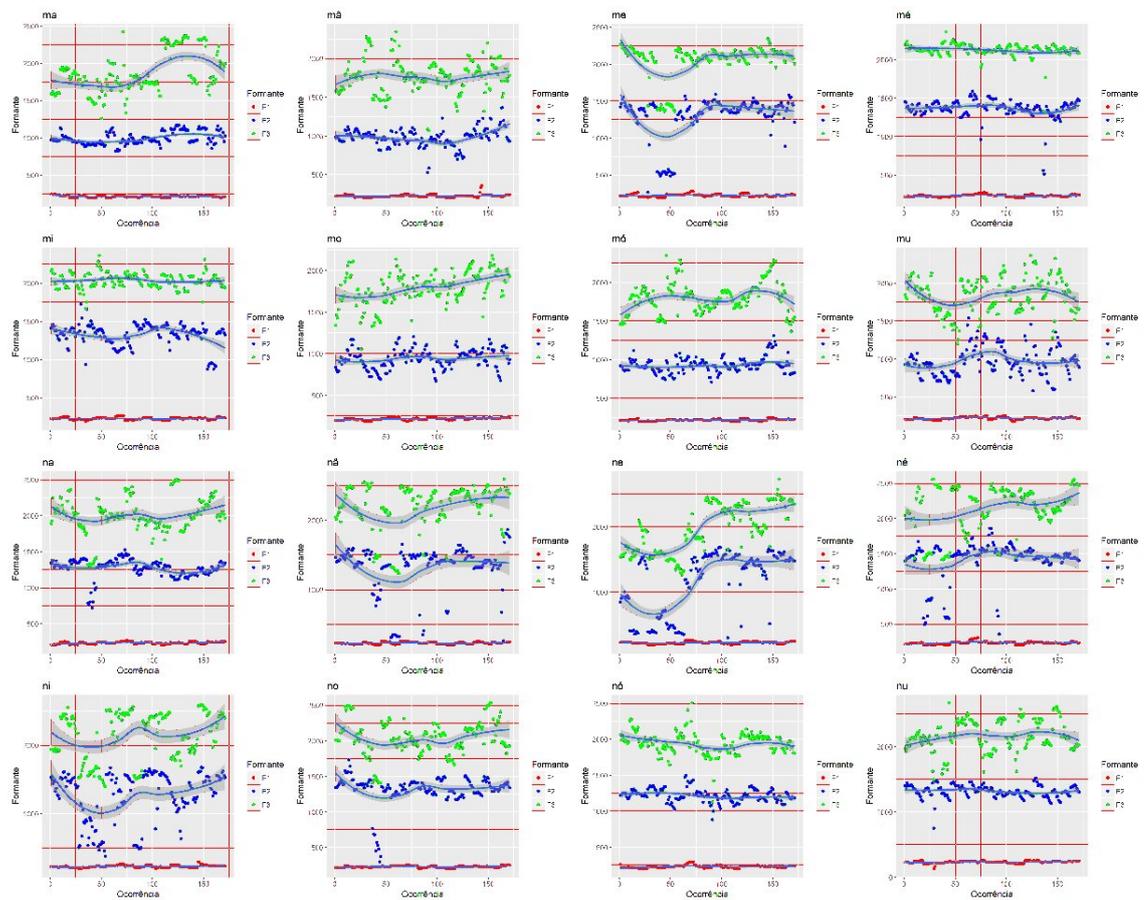
Participante RFS



Participante TBS



Participante VJP



ANEXO A – Alfabeto ARPABET retirado do manual do TADA*

Table I.1: ARPABET.

P	IY = /i/	heed
T	IH = /ɪ/	hid
K	EY = /eɪ/	hayed
B	EH = /ɛ/	head
D	AE = /æ/	had
G	AA = /ɑ/	hod
M	AO = /ɔ/	hawed
N	OW = /ou/	hoed
NX = /ŋ/	thing	UH = /u/	hood
F	UW = /u/	who'd
V	ER = /ɚ/	herd
TH = /θ/	thing	AH = /ʌ/	hud
DH = /ð/	this	AY = /aɪ/	hide
S	AW = /aʊ/	how
Z	OY = /ɔɪ/	boy
SH = /ʃ/	assure	AX = /ə/	about
ZH = /ʒ/	azure			
W			
R			
Y	= /j/			you
HH	= /h/			high
CH	= /tʃ/			chew
JH	= /dʒ/			jaw

* http://www.haskins.yale.edu/tada_download/doc/TADA_manual_v09.pdf

ANEXO B – Tabela I.2 retirada do manual do TADA

Stops

ARPA	Organ	Osc.	TV	Constr.	Target	Stiff.
B	Lips	clo.	LA	CLO	.	.
	Lips	rel.	LA	REL	.	.
	Velum	clo.	VEL	CLO	.	.
P	Lips	clo.	LA	CLO	.	.
	Lips	rel.	LA	REL	.	.
	Glottis	h	GLO	WIDE	.	.
	Velum	clo.	VEL	CLO	.	.
M	Lips	clo.	LA	CLO	.	.
	Lips	rel.	LA	REL	.	.
	Velum	n	VEL	WIDE	.	.
D	TT	clo.	TTCL	ALV	.	.
	TT	clo.	TTCD	CLO	.	.
	TT	rel.	TTCL	REL	.	.
	TT	rel.	TTCD	REL	.	.
	Velum	clo.	VEL	CLO	.	.
T	TT	clo.	TTCL	ALV	.	.
	TT	clo.	TTCD	CLO	.	.
	TT	rel.	TTCL	REL	.	.
	TT	rel.	TTCD	REL	.	.
	Glottis	h	GLO	WIDE	.	.
N	TT	clo.	TTCL	ALV	.	.
	TT	clo.	TTCD	CLO	.	.
	TT	rel.	TTCL	REL	.	.
	TT	rel.	TTCD	REL	.	.
	Velum	n	VEL	WIDE	.	.
G	TB	clo.	TBCL	VEL	.	.
	TB	clo.	TBCD	CLO	.	.
	TB	rel.	TBCD	REL	.	.
	Velum	clo.	VEL	CLO	.	.
K	TB	clo.	TBCL	VEL	.	.
	TB	clo.	TBCD	CLO	.	.
	TB	rel.	TBCD	REL	.	.
	Glottis	h	GLO	WIDE	.	.
	Velum	clo.	VEL	CLO	.	.
NX	TB	clo.	TBCL	VEL	.	.
	TB	clo.	TBCD	CLO	.	.
	TB	rel.	TBCD	REL	.	.
	Velum	n	VEL	WIDE	.	.

Fricatives

V	Lips.	crt	LA	CRIT.
	Lips.	rel	LA	REL
	Lips.	crt	LP.	DENT.
	Lips.	rel	LP.	REL
	Velum	crt	VEL.	CLO.
F	Lips.	crt	LA	CRIT.
	Lips.	rel	LA	REL
	Lips.	crt	LP.	DENT.
	Lips.	rel	LP.	REL
	Glottis	h	GLO.	WIDE
Velum	crt	VEL.	CLO.	
DH	TT	crt	TTCL	DENT.
	TT	crt	TTCD	CRIT.	-1.	..
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
	Velum	crt	VEL.	CLO.
TH	TT	crt	TTCL	DENT.
	TT	crt	TTCD	CRIT.	-1.	..
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
	Glottis	h	GLO.	WIDE
Velum	crt	VEL.	CLO.	
Z	TT	crt	TTCL	ALV.
	TT	crt	TTCD	CRIT.	0.16.	..
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
	TB	crt	TBCL	VEL.
TB	crt	TBCD	WIDE	
Velum	crt	VEL.	CLO.	
S	TT	crt	TTCL	ALV.
	TT	crt	TTCD	CRIT.
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
	TB	crt	TBCL	VEL.
TB	crt	TBCD	WIDE	
Glottis	h	GLO.	WIDE	
Velum	crt	VEL.	CLO.	
ZH	TT	crt	TTCL	ALVPAL.
	TT	crt	TTCD	CRIT.	0.5.	..
	TT	rel	TTCL	REL	40.	..
	TT	rel	TTCD	REL
	TB	crt	TBCL	PAL
TB	crt	TBCD	NAR	5.	..	
Velum	crt	VEL.	CLO.	
SH	TT	crt	TTCL	ALVPAL.
	TT	crt	TTCD	CRIT.
	TT	rel	TTCL	REL	40.	..
	TT	rel	TTCD	REL
	TB	crt	TBCL	PAL
TB	crt	TBCD	NAR	8.	..	
Glottis	h	GLO.	WIDE	
Velum	crt	VEL.	CLO.	

Affricates, Liquids, Glides

JH	TT	crt	TTCL	ALVPAL
	TT	crt	TTCD	CLO
	TT	rel	TTCL	REL	50	..
	TT	rel	TTCD	REL
	TB	crt	TBCL	PAL
	TB	crt	TBCD	NAR	5	..
	Velum	crt	VEL	CLO
CH	TT	crt	TTCL	ALVPAL
	TT	crt	TTCD	CLO
	TT	rel	TTCL	REL	50	..
	TT	rel	TTCD	REL
	TB	crt	TBCL	PAL
	TB	crt	TBCD	NAR	5	..
	Glottis	h	GLO	WIDE
Y	Velum	crt	VEL	CLO
	TB	nar	TBCL	PAL
	TB	nar	TBCD	NAR
W	Lips	nar	LA	V
	TB	nar	TBCL	UVU
	TB	nar	TBCD	NAR
	Lips	nar	LA	NAR
R	Lips	rel	LA	REL
	TT	nar	TTCL	PAL
	TT	nar	TTCD	NAR
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
L	Lips	voc	LA	NAR
	TT	nar	TTCL	ALV
	TT	nar	TTCD	NAR
	TT	rel	TTCL	REL
	TT	rel	TTCD	REL
	TB	voc	TBCL	UVU
HR	TB	voc	TBCD	NAR
	Glottis	h	GLO	WIDE
Q	Glottis	clo	GLO	CLO

Vowels

IY	Lips.	v.	LA	V.
	TB	v.	TBCL	PAL
	TB	v.	TBCD	V.	5.	..
IH	TB	v.	TBCL	PAL	..	3.
	TB	v.	TBCD	V.	8.	3.
EY	TB	v.	TBCL	PAL
	TB	v.	TBCD	V.	8.5.	..
EH	TB	v.	TBCL	PAL	..	3.
	TB	v.	TBCD	V.	11.5.	..
AE	TB	v.	TBCL	PHAR
	TB	v.	TBCD	V.	17.	..
AA	TB	v.	TBCL	PHAR
	TB	v.	TBCD	V.	11.	..
AO	TB	v.	TBCL	PHAR
	TB	v.	TBCD	V.	11.	..
	Lips.	v_rnd	LP.	PRO.	12.	..
	Lips.	v_rnd	LA	NAR	5.	..
UW	TB	v.	TBCL	UVU
	TB	v.	TBCD	V.	4.	..
	Lips.	v_rnd	LP.	PRO.	12.	..
	Lips.	v_rnd	LA	NAR	5.	..
UH	TB	v.	TBCL	UVU	..	3.
	TB	v.	TBCD	V.	6.	3.
	Lips.	v_rnd	LP.	PRO.	12.	..
	Lips.	v_rnd	LA	NAR	5.	..
OW	TB	v.	TBCL	UVUPHAR
	TB	v.	TBCD	V.	5.	..
	Lips.	v_rnd	LP.	PRO.	12.	..
	Lips.	v_rnd	LA	NAR	5.	..
AH	TB	v.	TBCL	UVUPHAR	..	3.
	TB	v.	TBCD	V.	6.5.	3.
AX	TB	v.	TBCL	UVU	..	3.
	TB	v.	TBCD	V.	8.5.	3.
ER	TT	v.	TTCL	PAL
	TT	v.	TTCD	NAR
	TB	v.	TBCL	PHAR
	TB	v.	TBCD	V.	11.	..

ANEXO C – *Script* utilizado para a medição dos dados

```
#####
##### #
# PRAAT SCRIPT "SEMI-AUTO FORMANT EXTRACTOR"
# This script semi-automates measuring formants from sound files with labeled TextGrids. It is
# based on and similar to the
# script "SemiAutoPitchAnalysis" by Daniel McCloy. This script cycles through a directory of
# TextGrids, finds associated
# sound files, opens them one at a time, displays a table of formant values for the interval at
# user-specified time points,
# and prompts the user to either (1) accept the formant measurements, (2) adjust the formant
# settings and recalculate, or
# (3) mark the interval as unmeasurable, before continuing on to the next interval or file.
#
# By default, the script will show a wideband spectrogram (from 0-5000Hz) and a formant
# track while running, but will
# not show pitch, pulses, intensity, etc. If you want to see additional analyses, change the
# arguments of the line that starts
# "Show analyses" (lines 165 & 168). If you want to change the underlying spectrogram
# settings, change lines 205-206.
# VERSION 0.3 (2012.05.05)
#
# CHANGELOG
# VERSION 0.3: Major reorganization of code (several complex control structures moved into
# procedures). Bug fixed where
# sometimes all time points through the interval were written out with values from the
# midpoint. Major
# improvements to efficiency by using formant values drawn from the editor window
# (since it's already open
# anyway) rather than extracting a sound slice and creating a formant object from it.
#
# VERSION 0.2: No significant code changes. Moderate improvements to documentation.
# License changed from CC to GPL.
#
```

```

# AUTHORS: DANIEL MCCLOY: (drmcclroy@uw.edu) & AUGUST MCGRATH

# LICENSED UNDER THE GNU GENERAL PUBLIC LICENSE v3.0 OR HIGHER:
http://www.gnu.org/licenses/gpl.html

# DEVELOPMENT OF THIS SCRIPT WAS FUNDED BY THE NATIONAL INSTITUTES OF HEALTH,
GRANT # R01DC006014 TO PAMELA SOUZA

#
#####
##### #

# COLLECT ALL THE USER INPUT

form Select directories for TextGrids and Sound files

    sentence Textgrid_directory
C:\Users\Márcio\Documents\Doutorado\Tese\Produção\COLETA\Marcados Tese\Silabas\

    sentence Sound_directory
C:\Users\Márcio\Documents\Doutorado\Tese\Produção\COLETA\Marcados Tese\Silabas\

    sentence Sound_extension .wav

    sentence Silaba

    comment Which TextGrid tier contains your segment labels?

    integer Label_tier 2

    comment You can pick up where you left off if you like:

    integer Starting_file_number 1

    comment How many seconds of the sound file do you want to
    comment see during analysis? (enter "0" to view the entire file)

    real Zoom_duration 1

    comment Full path of the output file:

    sentence Output_file
C:\Users\Márcio\Documents\Doutorado\Tese\Produção\COLETA\Marcados
Tese\FormantAnalysisResults.txt

    comment Default formant tracker settings (you can adjust "max formant" and
    comment "number of formants" later as you step through the intervals):

    positive Default_max_formant 5000

    positive Default_formant_number 5

    real Time_step 0.1

    real Preemphasis_from 50

```

```

positive Window_length 0.025
positive Dynamic_range 30
positive Dot_size 0.6
optionmenu Interval_measurement_option: 7
    option midpoint
    option onset, midpoint, offset
    option 20%, 50%, 80%
    option 25%, 50%, 75%
    option 10%, 30%, 50%, 70%, 90%
    option 5%, 10%, 20%, 50%, 80%, 90%, 95%
    option 5%, 10%, 20%, 35%, 50%, 65%, 80%, 90%, 95%
endform

# RUN SOME FUNCTIONS ON THE USER INPUT (TO BE USED LATER)
call pointsPerInterval

# BE FORGIVING IF THE USER FORGOT TRAILING PATH SLASHES OR LEADING FILE EXTENSION
DOTS
call cleanPath 'textgrid_directory$'
textgrid_dir$ = "cleanPath.out$"
call cleanPath 'sound_directory$'
sound_dir$ = "cleanPath.out$"
call cleanExtn 'sound_extension$'
sound_extn$ = "cleanExtn.out$"

# INITIATE THE OUTPUT FILE
if fileReadable (output_file$)
    beginPause ("The output file already exists!")
        comment ("The output file already exists!")
        comment ("You can overwrite the existing file, or append new data to the end
of it.")
    overwrite_setting = endPause ("Append", "Overwrite", 1)

```

```

if overwrite_setting = 2
    filedelete 'output_file$'
    call initializeOutfile
endif
else
    # THERE IS NOTHING TO OVERWRITE, SO CREATE THE HEADER ROW FOR THE NEW
    OUTPUT FILE
    call initializeOutfile
endif

# MAKE A LIST OF ALL TEXTGRIDS IN THE FOLDER
Create Strings as file list... list 'textgrid_dir$'*.*.TextGrid
file_list = selected("Strings")
file_count = Get number of strings

# LOOP THROUGH THE LIST OF FILES...
for current_file from starting_file_number to file_count

    # READ IN THE TEXTGRID & CORRESPONDING SOUND...
    select Strings list
    gridname$ = Get string... current_file
    Read from file... 'textgrid_dir$'gridname$'
    filename$ = selected$ ("TextGrid", 1)
    Open long sound file... 'sound_dir$'filename$"sound_extn$'
    total_duration = Get total duration

    # BOOLEAN TO PREVENT OPENING MULTIPLE EDITORS FOR THE SAME LONGSOUND
    new_file = 1

    # FIND THE LABELED INTERVAL...
    select TextGrid 'filename$'

```

```

num_intervals = Get number of intervals... label_tier
descarte$ = "fom nhó fim nhe fem nhi nho fam nhu nhã fum nha nhé "
# nasais$ = "ma mã me mé mi mo mó mu na ã ne né ni no nó nu "
intervalotiertres = 1
max_formant = default_max_formant
formant_number = default_formant_number

for interval to num_intervals

    select TextGrid 'filename$'

    label$ = Get label of interval... label_tier interval
    labelsil$ = Get label of interval... 3 intervalotiertres
    labelsil$ = labelsil$ + " "
    # IF THE LABEL IS NON-EMPTY, GET ITS ENDPOINTS
    #DEPURAÇÃO# pause 'labelsil$' 'interval' 'intervalotiertres'
    #ESTE IF SERVE PARA IR DIRETO À SÍLABA DESEJADA
    if labelsil$ = silaba$
    #ESTE IF SERVE PARA PERCORRER TODO O ARQUIVO
    #if label$ <> "" and index (descarte$, labelsil$) = 0
        start = Get starting point... label_tier interval
        end = Get end point... label_tier interval
        midpoint = (start+end)/2

        # PREVENT ZOOM DURATION FROM EXTENDING BEYOND THE ENDS
        OF THE FILE, BUT TRY TO MAINTAIN THE DESIRED WINDOW SIZE
        if not zoom_duration = 0
            left_edge = midpoint - zoom_duration/2
            right_edge = midpoint + zoom_duration/2
            right_excess = right_edge - total_duration

            if left_edge < 0
                zoom_start = 0

```

```

        if zoom_duration > total_duration
            zoom_end = total_duration
        else
            zoom_end = zoom_duration
        endif
    elif right_edge > total_duration
        zoom_end = total_duration
        if left_edge > right_excess
            zoom_start = zoom_end - zoom_duration
        else
            zoom_start = 0
        endif
    else
        zoom_start = left_edge
        zoom_end = right_edge
    endif
endif ; zoom_duration = 0
zoom_start = 0
zoom_end = total_duration
endif ; zoom_duration

if new_file = 1
    # IF THIS IS THE FIRST INTERVAL OF THE CURRENT FILE, SHOW
THE EDITOR WINDOW

    select LongSound 'filename$'
    plus TextGrid 'filename$'
    View & Edit

    # SINCE WE'RE IN THE FIRST LABELED INTERVAL, SET ALL THE
SETTINGS

    editor TextGrid 'filename$'

```

```

# FIRST, HIDE THE SPECTROGRAM ETC TO PREVENT
ANNOYING FLICKERING

Show analyses... no no no no no 10

Zoom... zoom_start zoom_end

# NOW SET ALL THE RELEVANT SETTINGS AND DISPLAY
WIDEBAND SPECTROGRAM

Spectrogram settings... 0 5000 0.005 50

Advanced spectrogram settings... 1000 250 Fourier

Gaussian yes 100 6 0

Formant settings... default_max_formant
default_formant_number window_length dynamic_range dot_size

Advanced formant settings... burg preemphasis_from

# SHOW THE FORMANT TRACKS

if not zoom_duration = 0

    # MAKE SURE THE "MAX ANALYSIS" SETTING IS
    LONG ENOUGH SO THE SPECTROGRAM ACTUALLY SHOWS UP

    Show analyses... yes no no yes no

zoom_duration*2

else

    # THE USER SPECIFIED "WHOLE FILE" SO WE
    ASSUME THE FILES ARE SHORT AND 10 SECONDS SHOULD BE ENOUGH

    Show analyses... yes yes no yes no 10

endif

endeditor

else

    # WE'RE NOT IN THE FIRST LABELED INTERVAL, SO EDITOR IS
    OPEN & SETTINGS ARE SET, SO JUST MOVE TO THE CURRENT INTERVAL

    editor TextGrid 'filename$'

    Zoom... zoom_start zoom_end

endeditor

endif

```

```

new_file = 0

# INITIALIZE SOME VARIABLES FOR THE PAUSE U.I.
clicked = 0
call getMeasureTimes
call getFormants
#pause 'labelsil$' 'label$'
call makeFormantTable
current_time_point = getMeasureTimes.time[1]
current_interval_measurement = 1

# PLACE CURSOR AT FIRST MEASUREMENT POINT
editor TextGrid 'filename$'
    Move cursor to... current_time_point
endeditor

# SHOW A U.I. WITH FORMANT TRACKER SETTINGS & MEASURED
FORMANT VALUES.

# KEEP SHOWING IT UNTIL THE USER ACCEPTS OR CANCELS THE
MEASUREMENT FOR THIS INTERVAL.
repeat
    beginPause ("Adjust formant tracker settings")
        comment ("File 'filename$' (file number 'current_file'
of 'file_count')")
        comment ("You can change these settings if the
formant track doesn't look right.")
        integer ("New_max_formant", max_formant)
        positive ("New_number_formants", formant_number)
        comment ("Clicking PLAY will play the sound in the
interval")
        comment ("Clicking REDRAW will redraw the formant
tracks with the settings above")
        comment ("Cicking SKIP will record all formants as zero
to mark for manual measurement")

```

```

comment (" ")
comment ("Formant measurements:")

# CREATE THE FORMANT TABLE
call getFormants
call makeFormantTable
comment ("makeFormantTable.header$")
comment ("makeFormantTable.f3$")
comment ("makeFormantTable.f2$")
comment ("makeFormantTable.f1$")
comment (" ")
sentence ("Notes_or_comments", "")
clicked = endPause ("Play", "Redraw", "Skip", "Accept", 4)

# IF THEY CLICKED "PLAY"
if clicked = 1
    editor TextGrid 'filename$'
        Play... start end
    endeditor

# IF THEY CLICKED "REDRAW"
elif clicked = 2
    max_formant = new_max_formant
    formant_number = new_number_formants
    editor TextGrid 'filename$'
        Formant settings... max_formant
    formant_number window_length dynamic_range dot_size
    endeditor
endif

until clicked >2

```

```
# END OF THE PAUSE U.I.
```

VALUES

```
# THE USER HAS EITHER ACCEPTED OR SKIPPED, SO WRITE OUT THE
```

```
for i from 1 to pointsPerInterval.pts
```

```
    time = getMeasureTimes.time[i]
```

```
    duracao = end-start
```

```
    percent = ((time-start)/duracao)*100
```

```
    if clicked = 3
```

```
        # MARK FOR HAND MEASUREMENT
```

```
        f1 = 0
```

```
        f2 = 0
```

```
        f3 = 0
```

```
        f4 = 0
```

```
        f5 = 0
```

```
    elif clicked = 4
```

```
        # GET MEASURED VALUES
```

```
        f1 = getFormants.f1[i]
```

```
        f2 = getFormants.f2[i]
```

```
        f3 = getFormants.f3[i]
```

```
        f4 = getFormants.f4[i]
```

```
        f5 = getFormants.f5[i]
```

```
        b1 = getFormants.b1[i]
```

```
        b2 = getFormants.b2[i]
```

```
        b3 = getFormants.b3[i]
```

```
        b4 = getFormants.b4[i]
```

```
        b5 = getFormants.b5[i]
```

```

        f0 = getFormants.f0[i]
    endif

    # WRITE OUT TO FILE

    resultline$ =
    "current_file"tab$filename$tab$labelsil$tab$label$tab$percent:0tab$duracao$tab$
    'f0"tab$f1"tab$b1"tab$f2"tab$b2"tab$f3"tab$b3"tab$f4"tab$b4"tab$f5"tab$b5"t
    ab$max_formant"tab$formant_number"tab$notes_or_comments$newline$

        fileappend "output_file$" 'resultline$'

    endfor ; EACH POINT IN THE INTERVAL

endif ; LABEL <> ""

# ISSO É PRA PAREAR O ANDAMENTO DOS DOIS CONTADORES DE INTERVALOS
if label$ = "n" or label$ = "m"
    intervalotiertres = intervalotiertres-2
endif

intervalotiertres = intervalotiertres + 1
endifor ; EACH INTERVAL IN THE FILE

# REMOVE ALL THE OBJECTS FOR THAT FILE AND GO ON TO THE NEXT ONE
select LongSound 'filename$'
plus TextGrid 'filename$'
Remove
select Strings list

endifor ; EACH FILE IN THE FOLDER

# REMOVE THE STRINGS LIST AND GIVE A SUCCESS MESSAGE
select Strings list
Remove
clearinfo

```

```
files_read = file_count - starting_file_number + 1
printline Done! 'files_read' files read.'newline$'
```

```
#####
```

```
# FUNCTIONS (A.K.A. PROCEDURES) THAT WERE CALLED EARLIER #
```

```
#####
```

```
procedure cleanPath .in$
```

```
    if not right$(.in$, 1) = "/"
```

```
        .out$ = ".in$" + "/"
```

```
    else
```

```
        .out$ = ".in$"
```

```
    endif
```

```
endproc
```

```
procedure cleanExtn .in$
```

```
    if not left$(.in$, 1) = "."
```

```
        .out$ = "." + ".in$"
```

```
    else
```

```
        .out$ = ".in$"
```

```
    endif
```

```
endproc
```

```
procedure initializeOutfile
```

```
    headerline$ =
```

```
"number'tab$'filename'tab$'sílabas'tab$'fone'tab$'posição'tab$'duração'tab$'f0'tab$'F1'tab$'B1'tab$'F2'tab$'B2'tab$'F3'tab$'B3'tab$'F4'tab$'B4'tab$'F5'tab$'B5'tab$'max
formant'tab$'number of formants'tab$'notes'newline$"
```

```
    fileappend "'output_file$" "headerline$"
```

```
endproc
```

```
procedure pointsPerInterval
# CALCULATE HOW MANY FORMANT MEASUREMENTS PER INTERVAL
  if interval_measurement_option = 1
    .pts = 1
  elif interval_measurement_option = 2 or interval_measurement_option = 3 or
interval_measurement_option = 4
    .pts = 3
  elif interval_measurement_option = 5
    .pts = 5
  elif interval_measurement_option = 6
    .pts = 7
  elif interval_measurement_option = 7
    .pts = 9
  endif
endproc
```

```
procedure getMeasureTimes
# MIDPOINT ONLY
  if interval_measurement_option = 1
    .time [1] = midpoint

# ONSET-MIDPOINT-OFFSET
  elif interval_measurement_option = 2
    .time [1] = start
    .time [2] = midpoint
    .time [3] = end

# 20-50-80
  elif interval_measurement_option = 3
    .time [1] = start + 0.2*(end-start)
    .time [2] = midpoint
```

```
.time [3] = start + 0.8*(end-start)
```

```
# 25-50-75
```

```
elif interval_measurement_option = 4
```

```
.time [1] = start + 0.25*(end-start)
```

```
.time [2] = midpoint
```

```
.time [3] = start + 0.75*(end-start)
```

```
# 10-30-50-70-90
```

```
elif interval_measurement_option = 5
```

```
.time [1] = start + 0.1*(end-start)
```

```
.time [2] = start + 0.3*(end-start)
```

```
.time [3] = midpoint
```

```
.time [4] = start + 0.7*(end-start)
```

```
.time [5] = start + 0.9*(end-start)
```

```
# 5-10-20-50-80-90-95
```

```
elif interval_measurement_option = 6
```

```
.time [1] = start + 0.05*(end-start)
```

```
.time [2] = start + 0.1*(end-start)
```

```
.time [3] = start + 0.2*(end-start)
```

```
.time [4] = midpoint
```

```
.time [5] = start + 0.8*(end-start)
```

```
.time [6] = start + 0.9*(end-start)
```

```
.time [7] = start + 0.95*(end-start)
```

```
# 5-10-20-35-50-65-80-90-95
```

```
elif interval_measurement_option = 7
```

```
.time [1] = start + 0.05*(end-start)
```

```
.time [2] = start + 0.1*(end-start)
```

```
.time [3] = start + 0.2*(end-start)
```

```

.time [4] = start + 0.35*(end-start)
.time [5] = midpoint
.time [6] = start + 0.65*(end-start)
.time [7] = start + 0.8*(end-start)
.time [8] = start + 0.9*(end-start)
.time [9] = start + 0.95*(end-start)
endif
endproc

procedure getFormants
  editor TextGrid 'filename$'

  for i from 1 to pointsPerInterval.pts
    Move cursor to... getMeasureTimes.time[i]
    .f5 [i] = Get formant: 5
    .b5 [i] = Get bandwidth: 5
    .f4 [i] = Get fourth formant
    .b4 [i] = Get fourth bandwidth
    .f3 [i] = Get third formant
    .b3 [i] = Get third bandwidth
    .f2 [i] = Get second formant
    .b2 [i] = Get second bandwidth
    .f1 [i] = Get first formant
    .b1 [i] = Get first bandwidth

    Show pitch
    .f0 [i] = Get pitch

    Show pitch
  endfor
  endeditor
endproc

```

procedure makeFormantTable

NOTE: THE EXTRA SPACES ARE INTENTIONAL, TO GET EVERYTHING TO LINE UP PROPERLY IN COLUMNS IN THE PAUSE WINDOW

MIDPOINT ONLY

if interval_measurement_option = 1

.header\$ = ""tab\$""tab\$'midpoint"

.f3\$ = ""tab\$'F3'tab\$' 'getFormants.f3[1]:0"

.f2\$ = ""tab\$'F2'tab\$' 'getFormants.f2[1]:0"

.f1\$ = ""tab\$'F1'tab\$' 'getFormants.f1[1]:0"

ONSET-MIDPOINT-OFFSET

elif interval_measurement_option = 2

.header\$ = ""tab\$""tab\$'onset'tab\$' mid'tab\$'offset"

.f3\$ = ""tab\$'F3'tab\$' 'getFormants.f3[1]:0'tab\$' 'getFormants.f3[2]:0'tab\$'
'getFormants.f3[3]:0"

.f2\$ = ""tab\$'F2'tab\$' 'getFormants.f2[1]:0'tab\$' 'getFormants.f2[2]:0'tab\$'
'getFormants.f2[3]:0"

.f1\$ = ""tab\$'F1'tab\$' 'getFormants.f1[1]:0'tab\$' 'getFormants.f1[2]:0'tab\$'
'getFormants.f1[3]:0"

20-50-80

elif interval_measurement_option = 3

.header\$ = ""tab\$""tab\$' 20%'tab\$' 50%'tab\$' 80%"

.f3\$ = ""tab\$'F3'tab\$' 'getFormants.f3[1]:0'tab\$' 'getFormants.f3[2]:0'tab\$'
'getFormants.f3[3]:0"

.f2\$ = ""tab\$'F2'tab\$' 'getFormants.f2[1]:0'tab\$' 'getFormants.f2[2]:0'tab\$'
'getFormants.f2[3]:0"

.f1\$ = ""tab\$'F1'tab\$' 'getFormants.f1[1]:0'tab\$' 'getFormants.f1[2]:0'tab\$'
'getFormants.f1[3]:0"

25-50-75

elif interval_measurement_option = 4

.header\$ = ""tab\$""tab\$' 25%'tab\$' 50%'tab\$' 75%"

```
.f3$ = ""tab$'F3'tab$' 'getFormants.f3[1]:0''tab$' 'getFormants.f3[2]:0''tab$'
'getFormants.f3[3]:0''
```

```
.f2$ = ""tab$'F2'tab$' 'getFormants.f2[1]:0''tab$' 'getFormants.f2[2]:0''tab$'
'getFormants.f2[3]:0''
```

```
.f1$ = ""tab$'F1'tab$' 'getFormants.f1[1]:0''tab$' 'getFormants.f1[2]:0''tab$'
'getFormants.f1[3]:0''
```

```
# 10-30-50-70-90
```

```
elif interval_measurement_option = 5
```

```
.header$ = ""tab$''tab$' 10%'tab$' 30%'tab$' 50%'tab$' 70%'tab$' 90%''
```

```
.f3$ = ""tab$'F3'tab$' 'getFormants.f3[1]:0''tab$' 'getFormants.f3[2]:0''tab$'
'getFormants.f3[3]:0''tab$' 'getFormants.f3[4]:0''tab$' 'getFormants.f3[5]:0''
```

```
.f2$ = ""tab$'F2'tab$' 'getFormants.f2[1]:0''tab$' 'getFormants.f2[2]:0''tab$'
'getFormants.f2[3]:0''tab$' 'getFormants.f2[4]:0''tab$' 'getFormants.f2[5]:0''
```

```
.f1$ = ""tab$'F1'tab$' 'getFormants.f1[1]:0''tab$' 'getFormants.f1[2]:0''tab$'
'getFormants.f1[3]:0''tab$' 'getFormants.f1[4]:0''tab$' 'getFormants.f1[5]:0''
```

```
# 5-10-20-50-80-90-95
```

```
elif interval_measurement_option = 6
```

```
.header$ = ""tab$' 5%'tab$''tab$' 10%'tab$' 20%'tab$' 50%'tab$' 80%'tab$'
90%'tab$' 95%''
```

```
.f3$ = ""tab$'F3'tab$' 'getFormants.f3[1]:0''tab$' 'getFormants.f3[2]:0''tab$'
'getFormants.f3[3]:0''tab$' 'getFormants.f3[4]:0''tab$' 'getFormants.f3[5]:0''tab$'
'getFormants.f3[6]:0''tab$' 'getFormants.f3[7]:0''
```

```
.f2$ = ""tab$'F2'tab$' 'getFormants.f2[1]:0''tab$' 'getFormants.f2[2]:0''tab$'
'getFormants.f2[3]:0''tab$' 'getFormants.f2[4]:0''tab$' 'getFormants.f2[5]:0''tab$'
'getFormants.f2[6]:0''tab$' 'getFormants.f2[7]:0''
```

```
.f1$ = ""tab$'F1'tab$' 'getFormants.f1[1]:0''tab$' 'getFormants.f1[2]:0''tab$'
'getFormants.f1[3]:0''tab$' 'getFormants.f1[4]:0''tab$' 'getFormants.f1[5]:0''tab$'
'getFormants.f1[6]:0''tab$' 'getFormants.f1[7]:0''
```

```
# 5-10-20-35-50-65-80-90-95
```

```
elif interval_measurement_option = 7
```

```
.header$ = ""tab$''tab$' 5%'tab$' 10%'tab$' 20%'tab$' 35%'tab$' 50%'tab$'
65%'tab$' 80%'tab$' 90%'tab$' 95%''
```

```
.f3$ = ""tab$'F3'tab$' 'getFormants.f3[1]:0''tab$' 'getFormants.f3[2]:0''tab$'  
'getFormants.f3[3]:0''tab$' 'getFormants.f3[4]:0''tab$' 'getFormants.f3[5]:0''tab$'  
'getFormants.f3[6]:0''tab$' 'getFormants.f3[7]:0''tab$' 'getFormants.f3[8]:0''tab$'  
'getFormants.f3[9]:0''
```

```
.f2$ = ""tab$'F2'tab$' 'getFormants.f2[1]:0''tab$' 'getFormants.f2[2]:0''tab$'  
'getFormants.f2[3]:0''tab$' 'getFormants.f2[4]:0''tab$' 'getFormants.f2[5]:0''tab$'  
'getFormants.f2[6]:0''tab$' 'getFormants.f2[7]:0''tab$' 'getFormants.f2[8]:0''tab$'  
'getFormants.f2[9]:0''
```

```
.f1$ = ""tab$'F1'tab$' 'getFormants.f1[1]:0''tab$' 'getFormants.f1[2]:0''tab$'  
'getFormants.f1[3]:0''tab$' 'getFormants.f1[4]:0''tab$' 'getFormants.f1[5]:0''tab$'  
'getFormants.f1[6]:0''tab$' 'getFormants.f1[7]:0''tab$' 'getFormants.f1[8]:0''tab$'  
'getFormants.f1[9]:0''
```

```
endif
```

```
endproc
```