

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO  
PUC-SP

Maria Augusta Forte Svicero

Caracterização acústica e de imagens de ultrassonografia  
das vogais orais do Português Brasileiro

Mestrado em Linguística Aplicada e Estudos da Linguagem

SÃO PAULO

2012

Maria Augusta Forte Svicero

Caracterização acústica e de imagens de ultrassonografia  
das vogais orais do Português Brasileiro

MESTRADO EM LINGUÍSTICA APLICADA E ESTUDOS DA LINGUAGEM

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de MESTRE em Linguística Aplicada e Estudos da Linguagem sob a orientação da Profa. Dra. Zuleica Camargo.

SÃO PAULO

2012

Banca Examinadora

-----

-----

-----

Data: \_\_\_\_/\_\_\_\_/\_\_\_\_

Svicero, Maria Augusta Forte

Caracterização acústica e de imagens de ultrassonografia das vogais orais do Português Brasileiro / Svicero, Maria Augusta Forte - São Paulo, 2012

Dissertação (mestrado) Pontifícia Universidade Católica de São Paulo. Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem.

Título em inglês: Acoustic and ultrasound images characterization of oral vowels in Brazilian Portuguese

### **AUTORIZAÇÃO**

Autorizo, exclusivamente para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação por processos de fotocopiadoras ou eletrônicos.

Maria Augusta Forte Svicero \_\_\_\_\_

São Paulo,        de                de 2012.

## DEDICATÓRIA

AOS MEUS PAIS

Por dedicarem suas vidas aos filhos  
Por transmitirem valores íntegros  
Por ensinarem que a felicidade é simples  
Por acreditarem que na fé somos mais fortes  
Por serem minha fortaleza quando fraquejo  
Por serem sempre meu porto seguro  
Por guiarem meus caminhos  
Por serem minha inspiração  
Por serem como são!

Agradeço por tê-los como meus pais e  
por tudo o que por mim sempre fizeram!

Amo vocês!

## AGRADECIMENTOS

A Deus por tantas benções.

À Profa. Dra. Zuleica Camargo agradeço pelos conhecimentos transmitidos com tanta sabedoria, propriedade e paixão ao longo destes anos. Obrigada pelas orientações, pelas oportunidades e pela amizade!

À Profa. Dra. Sandra Madureira, pela serenidade e generosidade com que transmite tudo o que sabe. Obrigada pelas sugestões ao longo do trabalho!

À Profa. Dra. Luisa Barsaghi-Ficker pela leitura detalhada e contribuições no exame de qualificação.

Ao Prof. Dr. Chi Nan Pai, pela participação nas coletas experimentais e no exame de qualificação. Obrigada pelas ideias e orientações.

Ao Prof. Dr. Luiz Carlos Rusilo pela análise estatística.

Ao Dr. Renê Keizo Arikawa pela gentileza e disponibilidade constantes ao permitir o desenvolvimento de parte importante dessa pesquisa! Obrigada!

À Dra. Sílvia Pinho por transmitir seus valiosos conhecimentos e pelas oportunidades. Ao Dr. Ricardo Pinho, com saudades, pela paciência e ensinamentos.

Ao grupo do LIAAC por tantos momentos que dividimos e vencemos! Especialmente, à Cristina Canhetti e Andrea Sacco, pela alegria, trocas de experiências e amizade; à Aline Pessoa pelo ajuda com o material da coleta e amizade; à Fabiana Gregio, pela cumplicidade, generosidade e amizade; à Luciana Oliveira pela presença e apoio no exame de qualificação, por esses anos de amizade e aprendizados; à Lilian Kuhn, pelos dados que contribuíram na construção desse trabalho; à Roberta Cury por tantas ajudas nas coletas experimentais e nas discussões do projeto; ao Marcelo Barbosa pelas boas risadas e revisão do inglês; à Perpétua e Thamires, dupla muito prestativa; ao

Sergio, Vanessa, Alice, Janaina e Paulo pelas trocas nas atividades que realizamos! Foi muito bom estar com todos vocês!

À Maria Lucia do LAEL e à Fátima Albuquerque do LIAAC, sempre muito prestativas e atenciosas!

Ao CEFAC pela disponibilidade do espaço e material para parte da coleta de dados. Agradeço aos atenciosos e queridos funcionários e também à Dra. Irene Marchesan, sonhadora e realizadora de conquistas para a nossa área.

À CAPES pela auxílio financeiro fornecido para a pesquisa.

À Ser Clínica, especialmente à Birgit por compreender e incentivar meu percurso nestes anos, à Inês pela gentileza e carinho, e à Priscilla que organizou meus horários nesse período!

Aos meus pais – Elói e Carolina – meu eterno agradecimento por tudo! Obrigada também pela ideia, desenvolvimento e confecção de parte importante deste trabalho!

Ao Tato, meu irmão querido. Agradeço todos os dias por todas as fases que vivemos juntos e peço para que continuemos sempre unidos. Sou sua maior torcedora e estarei sempre ao seu lado!

À Tia Cris, super madrinha! Presença constante em todas as fases da minha vida! Obrigada pelo amor, carinho, apoio, torcida, desenhos, projetos, surpresas... por tudo o que já fez por mim!

À amada D. Lucia, minha avó do coração! Por cada abraço, cada roupa cuidada com carinho, cada oração por toda a família, pela dedicação de tantos anos! Muito obrigada! Saudades eternas!

Ao Lipe e ao amor que nasceu entre nós. Sua alegria me conquistou, seu cuidado me guiou, seu amor me transformou. Que bom que encontrei você! Que a felicidade e o amor nos acompanhem sempre!



À minha `grande` família – avô Carlito (avós Norma, Nair e Argemiro, e Tio Elcio – no coração com saudade!), tios, primos e afilhados. É muito bom tê-los por perto!

À querida Família Façanha – Edu, Tonia, Carlos, Tati, Gui, D. Maria Augusta – por me acolherem com tanto carinho e compreenderam minhas ausências nesse último ano de trabalho! Ganhei mais uma família!

Aos meus amigos que acompanharam, torceram e entenderam os encontros desmarcados...obrigada por estarem sempre por perto!

Aos sujeitos participantes dessa pesquisa que se disponibilizaram voluntariamente e contribuíram para o seu desenvolvimento. Obrigada pela participação, disponibilidade e amizade!

## RESUMO

Estudos de natureza fonético-acústica e articulatória revelam a possibilidade de compreender a produção da fala por meio de instrumentais de decomposição sonora (Fonética Acústica) e de descrição de imagens da configuração do trato vocal (Fonética Articulatória). A teoria acústica da produção da fala propicia o entendimento dos ajustes realizados na produção dos sons vocálicos de uma língua. A análise acústica e a ultrassonografia (USG) propiciam, respectivamente, pistas acústicas e articulatórias do comportamento das estruturas envolvidas na fala, de forma não invasiva. O objetivo deste trabalho foi caracterizar acusticamente (frequências formânticas) as vogais orais do português brasileiro (PB), variante paulistana, e apresentar dados preliminares da investigação de vogais por imagens de USG, apontando possíveis correlações com a esfera acústica. A primeira parte da coleta consistiu na gravação em áudio de 16 falantes do PB de ambos os gêneros, sem queixa e/ou alterações de fala, auditivas e cranio-faciais. Para a segunda parte da coleta um sujeito do gênero masculino foi submetido à gravação em áudio simultânea à gravação com o aparelho de USG. Foi desenvolvido para a presente pesquisa um suporte que permitiu a acoplagem do transdutor de USG e a sua fixação durante a produção de fala. Ambas as coletas consistiram na produção das sete vogais orais do PB inseridas em sentenças-veículo com estrutura “Digo [pVpv] baixinho”, em que V é a vogal oral produzida em posição tônica e v vogal similar àquela indicada em V, porém em posição pós-tônica. As gravações acústicas foram digitalizadas em extensão .wav e analisadas pelo *software* de livre acesso *Praat* para extração dos valores de frequência (Hz) dos três primeiros formantes (F1, F2 e F3). As amostras de vídeo de USG foram submetidas à digitalização, edição e sincronização aos dados de áudio pelo *software VideoPad Video Editor*. As imagens de USG foram submetidas a duas abordagens de análise de dados: 1- tratamento das imagens e geração de traçados manuais do contorno da língua e 2- utilização do *software* de livre acesso *Osiris* para extração de medidas de deslocamento da língua nos eixos vertical e horizontal. Os dados acústicos foram submetidos a análise estatística (análise discriminante) e revelaram que: (1) os valores das frequências formânticas contribuíram para a diferenciação das emissões entre os gêneros; (2) no gênero masculino os valores de F1 foram os mais influentes na segregação das sete vogais orais; (3) no gênero feminino, os valores de F2 foram os mais influentes na segregação das sete vogais orais. Os dados de imagem de USG revelaram o contorno do dorso de língua nos eixos vertical e horizontal durante a produção das vogais orais do PB, sugerindo melhor diferenciação para deslocamentos no eixo vertical. Os valores de referência para as frequências formânticas das sete vogais orais de falantes dos gêneros masculino e feminino foram apresentados, bem como imagens de USG revelaram potencial aplicação no estudo das correlações acústico-articulatórias das vogais.

**Palavras-chave ou descritores:** Fala; Acústica da Fala; Medida de Produção da Fala; Ultrassonografia; Fonética

## ABSTRACT

Studies on acoustic and articulatory phonetics reveal a possibility of understanding speech production through sound decomposing instrumentals (Acoustic Phonetics) and vocal tract imaging (Articulatory Phonetics). The acoustic theory of speech production provides an understanding of the adjustments made in the production of the vowel sounds of a language. Acoustic analysis and ultrasonography (US) enable, respectively, the description of acoustic and articulatory clues of the speech, by means of a non-invasive approach. The aim of this study was to present the acoustic description (formant frequencies) of oral vowels of Brazilian Portuguese (BP), São Paulo's variant, and present preliminary research data of vowels by images of US, pointing out possible acoustic-articulatory correlations. The first part of the data collection consisted on audio recording from 16 BP speakers, both genders, without speech, hearing or cranium-facial impairments or alteration. For the second part of the data collection, a male subject was submitted to simultaneous audio and US recordings. For this research a supporting device was developed to enable the transducer to be coupled to the US equipment. The recordings comprised the production of seven oral vowels of BP within the structure of the carrier sentences "Diga [pVpV] baixinho" (Say pVpv softly), where V is the oral vowel produced in stressed position and v post stressed position to the similar vowel. The acoustic recordings were digitized in .wav extensions and the F1, F2 and F3 measures were extraceted by means of software PRAAT. The US video samples were edited in order synchronize them to the audio data (software VideoPad Video Editor). The US images underwent two analysis approaches: 1<sup>st</sup> –images were manually treated in order to retrieve the tongue contours 2<sup>nd</sup> - extraction of tongue displacement measures in vertical and horizontal axes. The acoustic data were statistically treated by means of discriminant analysis. The results pointed out that: (1) the formant frequency values contributed to the differentiation of the speech production's gender; (2) male F1 values were the most influential acoustic measures for the segregation of the seven oral vowels; (3) female F2 values were the most influential in the segregation of the seven oral vowels. US image data revealed the tongue body contour in the vertical and horizontal axis during the production of oral vowels of BP, suggesting better differentiation for displacements on the vertical axis. The reference values for the formant frequency of seven oral vowels of the male and female genders were presented, as well as images of US revealed potential application in study of acoustic-articulatory for correlations vowels sounds.

**Keywords:** Speech; Acoustic; Speech Production Measures; Ultrasonography; Phonetics

## **Lista de Quadros**

Quadro 1-	Descrição dos sujeitos participantes: gênero, idade e
escolaridade.....	31

## Lista de Figuras

Figura 1- Exemplo de sentença-veículo apresentada ao sujeito durante a gravação de <i>corpus</i> de pesquisa .....	33
Figura 2- Instrução para a leitura do <i>corpus</i> de pesquisa .....	33
Figura 3 - Traçado de forma da onda (janela superior) e espectrograma de banda larga (janela inferior): seleção de ponto estacionário (indicado por flecha) da emissão da vogal [a] em posição tônica, com medidas geradas pelo extrator automático (janela <i>Praat Info</i> ) .....	34
Figura 4 - Traçado de forma da onda (janela superior) e espectrograma de banda estreita (janela inferior): seleção de ponto estacionário (indicado por flecha) da emissão da vogal [a] em posição tônica e geração do espectro FFT (janela <i>Spectrum</i> ) para conferência dos picos espectrais .....	34
Figura 5- Espectro FFT (linha contínua) e espectro LPC (linha tracejada) .....	35
Figura 6 - Esquema ilustrativo do suporte desenvolvido para a fixação do transdutor de ultrassonografia para coletas de fala .....	38
Figura 7 - Suporte para acoplagem do transdutor de ultrassonografia desenvolvido para a presente pesquisa (a) visão lateral do suporte montado e (b) visão frontal do suporte montado; (c) pedestal de teclado; (d) cabo de madeira de 16cm; (e) suporte de microfone YH4; (f) suporte de microfone do tipo cachimbo de pedestal; (g) visão lateral do suporte com o transdutor de USG; (h) visão frontal do suporte com o transdutor de ultrassonografia .....	39

Figura 8 -	Posicionamento do sujeito participante da pesquisa (a) e figura alvo para manutenção da posição (b) .....	40
Figura 9 -	Edição e sincronização dos vídeos gerados pelo aparelho de ultrassonografia com os registros de áudio pelo <i>software VideoPad Video Editor</i> .....	41
Figura 10 -	Imagem de ultrassonografia em negativo .....	42
Figura 11 -	Esquema de análise das imagens de ultrassonografia: padrão quadriculado para estimar o deslocamento do dorso de língua para as distintas emissões vocálicas (a) e a linha contínua para delinear o contorno do dorso de língua (b) na abordagem 1 .....	42
Figura 12 -	Pontos de referência da extração das medidas (em milímetros) de deslocamento do dorso de língua a partir das imagens de ultrassonografia pelo <i>software Osiris</i> : (a) ponto de partida no centro inferior da imagem; (b) deslocamento vertical máximo até o ponto do contorno de língua visualizado; (c) e (d) deslocamento horizontal máximos, respectivamente, até o ponto anterior visualizado e o ponto posterior visualizado do contorno de língua para a abordagem 2 .....	43
Figura 13 -	Distribuição dos valores de frequências formânticas (F1, F2 e F3) das vogais orais do PB de falantes dos gêneros masculino e feminino .....	45
Figura 14 -	Trapézio das vogais orais traçado por meio do <i>script CartaFormantes.Log.praat</i> para os gêneros masculino e feminino .....	46
Figura 15 -	Matriz de confusão para os resultados de validação cruzada da análise discriminante de medidas acústicas (F1, F2 e F3) e gêneros .....	46

Figura 16 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 para identificação do gênero do falante) .....	47
Figura 17 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (valor segregatório das medidas de F1, F2 e F3 de vogais orais para o gênero masculino) .....	47
Figura 18 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (valor segregatório das medidas de F1, F2 e F3 de vogais orais para o gênero feminino) .....	48
Figura 19 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 por vogais orais para falantes do gênero masculino) .....	48
Figura 20 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 por vogais orais para falantes do gênero feminino) .....	49
Figura 21 - Imagem de ultrassonografia (abordagem 1), sendo: 1. a imagem em corte sagital permitida pelo aparelho de ultrassonografia e pelo transdutor endocavitário; 2. o traçado branco que indica a posição do dorso de língua; 3. a estrutura do corpo da língua; 4 e 5. indicam a direção da imagem, respectivamente anterior e posterior .....	50
Figura 22 - Imagem de ultrassonografia (abordagem 1) extraída no momento de postura habitual (repouso) de língua. O traçado branco indica a posição do dorso de língua .....	50
Figura 23 - Imagens de ultrassonografia (abordagem 1) extraídas (a) no momento de preparação da língua para o início da deglutição; (b) no momento	

de início da deglutição; (c) no momento da deglutição propriamente dita; (d) ao final da deglutição ..... 51

Figura 24 - Imagens de ultrassonografia (abordagem 1) das vogais orais: vogal [a]; vogal [ɛ]; vogal [e]; vogal [i]; vogal [ɔ]; vogal [o]; vogal [u]. O traçado da linha branca acompanha o contorno da língua, exceto na vogal [u], em que não foi possível a sua visualização dado o movimento posterior ..... 51

Figura 25 - Imagens de ultrassonografia (abordagem 1) geradas em padrão negativo das vogais orais: vogal [a]; vogal [ɛ]; vogal [e]; vogal [i]; vogal [ɔ]; vogal [o]; vogal [u] ..... 53

Figura 26 - Imagem gerada em sobreposição do traçado que acompanha o contorno da língua: vogal [a] – linha vermelha; vogal [ɛ] – linha amarela; vogal [e] – linha azul; vogal [i] – linha verde; vogal [ɔ] linha rosa; vogal [o] linha laranja (abordagem 1) ..... 54

Figura 27 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (das informações de ultrassonografia para as medidas acústicas F1, F2 e F3) das sete vogais orais do português brasileiro ..... 54

Figura 28 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3) para identificação das vogais ..... 55

Figura 29 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (medidas de deslocamento da língua e medidas acústicas (F1, F2 e F3) para as vogais orais ..... 56

Figura 30 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de deslocamento do dorso de língua para as medidas de F1, F2 e F3) das vogais orais ..... 56



## **Lista de Tabelas**

Tabela 1 - Valores de média, desvio-padrão (DP) e intervalos de confiança (IC) das medidas de frequências (Hz) dos três primeiros formantes (F1, F2 e F3) das vogais orais dos falantes dos gêneros masculino e feminino .....	44
--	----

## **Lista de Abreviaturas**

DA - deficiência auditiva

DP - desvio padrão

EGG - eletroglotografia

EMA - articulografia magnética

EMG - eletromiografia

EPG - eletropalatografia

F1 - primeiro formante

F2 - segundo formante

F3 - terceiro formante

Hz – *Hertz*

IC – intervalo de confiança

IRM - imagens de ressonância magnética

LAEL - Linguística Aplicada e Estudos da Linguagem

mm - milímetros

PEPG – Programa de Estudos Pós Graduated

PB - português brasileiro

TCLE – Termo de Consentimento Livre e Esclarecido

USG - ultrassonografia

## SUMÁRIO

Resumo .....	x
Listas .....	xii
1. INTRODUÇÃO .....	20
2. REVISÃO DE LITERATURA .....	24
3. METODOLOGIA .....	31
4. RESULTADOS .....	44
5. DISCUSSÃO .....	57
6. CONCLUSÃO .....	62
7. ANEXOS .....	63
8. REFERÊNCIAS BIBLIOGRÁFICAS .....	66

## 1. INTRODUÇÃO

Considerar a produção da fala humana como um processo dinâmico conduz à demanda de investigação da interface entre os processos de sua produção e percepção e impulsiona o desenvolvimento de tecnologias. Todo o comportamento humano é dinâmico e resulta da coordenação de movimentos que se ajustam para realizar uma tarefa específica, que pode incluir a produção de fala (KELSO, 1995).

O trato vocal humano (região que se estende desde as pregas vocais até os lábios ou cavidade nasal) é um tubo flexível que pode ser moldável pela ação dos órgãos fonoarticulatórios (BECK, 2006). As vogais, do grupo de sons ressoantes da fala, são compreendidas como o resultado do som produzido nas pregas vocais modificado pela ação dos órgãos fonoarticulatórios (lábios, língua, dentes, mandíbula, palato duro e mole, faringe, cavidade nasal).

Abordaremos as relações entre acústica e articulação dos sons vocálicos, com respaldo na teoria acústica da produção da fala (FANT, 1960), especialmente do modelo fonte-filtro.

A teoria acústica de Fant (1960) enfoca a produção de fala com a construção do modelo do trato vocal, baseada em modelo matemático linear. Destaca-se que a validade dos resultados encontrados na produção da fala são dependentes do modelamento do trato vocal, especialmente no que diz respeito ao seu comprimento e largura (LADEFOGED, HARSHMAN, 1979).

O modelo fonte-filtro aplicado à produção das vogais define a produção do sinal de fala como o produto da fonte de energia (produto acústico da vibração de pregas vocais) somado à ação de filtros (ressoadores - cavidades supraglóticas) (FANT, 1960).

Tal modelo impulsionou os estudos em análise acústica da fala, com a vantagem de permitir a investigação do sinal de fala de forma não invasiva. No caso das vogais, a caracterização acústica reside na descrição das ressonâncias do trato vocal supraglótico, conhecidas como frequências formânticas ou formantes. Tais elementos caracterizam-se como as ressonâncias do trato vocal, traduzidos na função de filtro, a partir da fonte sonora propiciada pela vibração

das pregas vocais. O padrão de formantes é definido por sua frequência, intensidade e largura de banda (STEVENS, 2000).

As vogais são elementos importantes na estruturação do sinal de fala, na medida em que veiculam as informações de acento lexical e frasal, ritmo, melodia e entoação, demarcando variações das propriedades acústicas básicas (frequência, intensidade e duração), as quais são determinantes dos aspectos prosódicos, que colaboram para a veiculação de sentido (MADUREIRA, 2004).

Dessa forma, a caracterização acústico-articulatória das vogais é relevante para a compreensão do sistema de uma determinada língua, sendo, no caso do português brasileiro (PB), uma demanda ainda a ser contemplada em contexto mais amplo e de forma a respeitar a variedade fonética no território. O estudo dos sons da fala, portanto, pode ser abordado nas vertentes da Fonética Acústica, Articulatória e Perceptiva, as quais têm dialogado com as áreas de conhecimento congregadas nos estudos da Fala.

As correlações acústico-articulatórias tornam-se, nesse sentido, uma opção de abordagem do inventário fonético das línguas, e permitem uma inferência sobre a configuração do trato vocal supraglótico, ou seja, uma estimativa da movimentação dos articuladores, tais como a postura de língua, lábios, mandíbula e palato mole.

A caracterização dos ajustes supraglóticos do trato vocal nas vogais pode correlacionar as medidas acústicas do padrão de formantes (frequência, intensidade e banda de formantes) com a descrição articulatória, por meio de instrumentais desenvolvidos pelas Ciências da Fala para imagens do trato vocal. As técnicas de ultrassonografia (USG), de imagens de ressonância magnética (IRM), de eletroglotografia (EGG) podem ampliar os procedimentos de investigação e avaliação de produções de fala sem alterações e a futura exploração das manifestações de alterações das mais diversas naturezas.

Neste contexto, MENDES (2003) analisou a relação entre produção e percepção das sete vogais orais do PB em um sujeito deficiente auditivo e outro sujeito referência (ouvinte). Neste estudo, a análise acústica foi o instrumento

utilizado para a compreensão dos dados articulatórios e perceptivos da fala, com base na Fonética Acústica.

Os estudos reunidos nesta dissertação, que relacionam a Fonética Acústica e a Fonética Articulatória, incorporam diferentes instrumentais que contemplam a análise acústica, a EGG, a USG, a IRM, a eletromiografia (EMG), a eletropalatografia (EPG) e a articulografia magnética (EMA).

Mais recentemente, voltando-se aos avanços da incorporação das Tecnologias de Fala nas Ciências Fonéticas, o método de IRM tem sido usado nas pesquisas de fala por ser um procedimento não invasivo, inócuo, no sentido de que não há exposição à radiação, e provedor de boa qualidade de imagem.

Entretanto, a grande desvantagem, reconhecida para o estudo dos processos dinâmicos da fala, relaciona-se ao posicionamento exigido ao longo da captação de imagens (posição deitada, com cabeça fixa). Tal situação interfere no posicionamento dos articuladores no ato da fala, além da geração de ruído que dificulta a captação simultânea dos sons produzidos para estudos de correlações acústico-articulatórias.

Apesar de tais limitações, GREGIO (2006) caracterizou a configuração do trato vocal supraglótico na produção das vogais orais do PB, por IRM, tratando-se de estudo recente do PB com relação à descrição articulatória das vogais produzidas em padrão sustentado por um falante da cidade de São Paulo.

Neste cenário, a técnica de USG para captação de imagens do trato vocal supraglótico tem se destacado nas últimas décadas, especialmente para a descrição das vogais (KELSEY *et al.*, 1969) e por propiciar dados da configuração do trato vocal supraglótico em repouso e em movimento (FINGËR *et al.*, 2004). Dessa forma, procura-se investigar o comportamento das estruturas supraglóticas envolvidas na produção da fala, mais especificamente dos sons das vogais, sem exposição à radiação e em postura mais natural de fala.

A utilização do instrumental de USG é descrita como um procedimento simples e que propicia a análise do trato vocal supraglótico de forma não invasiva (FINGËR *et al.*, 2004), com possibilidade de gravação simultânea de áudio e de emissões de fala encadeada.

Diante destes avanços, torna-se possível resgatar o enfoque das questões dinâmicas da fala, não propiciadas por estudo das vogais sustentadas.

Busca-se, portanto, ampliar a investigação das características fonéticas de configuração do trato vocal supraglótico na produção das vogais orais do PB, em sua variante paulistana, com aprofundamento da descrição de seus correlatos acústicos (frequências formânticas) e articulatórios (imagens de USG da língua), para possibilitar respostas às seguintes demandas:

- descrição acústica (frequências formânticas) das vogais orais do PB, variante paulistana;
- investigação da possibilidade de descrição das vogais orais por meio da técnica de USG;
- descrição preliminar de correspondências entre os dados acústicos (frequências formânticas) e imagens de USG do trato vocal supraglótico (mais especificamente da postura de língua).

As demandas expostas suscitam a possibilidade de compreensão da atividade do trato vocal supraglótico na produção das vogais, as quais, futuramente, poderão influenciar o ensino de fonética do PB, além de procedimentos clínicos em avaliação e terapia de distúrbios da fala de diversas naturezas.

Concernente à linha de Pesquisa Linguagem e Patologias da Linguagem do Programa de Estudos Pós Graduated (PEPG) em Linguística Aplicada e Estudos da Linguagem (LAEL), que investiga os modelos de natureza dinâmica de produção de fala e a relação entre produção e percepção, este estudo teve por **objetivo** caracterizar acusticamente (frequências formânticas) as vogais orais do PB, variante paulistana, e apresentar dados preliminares da investigação de vogais por imagens de USG, apontando possíveis correlações com a esfera acústica.

## 2. REVISÃO DE LITERATURA

A produção da fala resulta da associação de três componentes fisiológicos básicos, descritos usualmente para melhor compreensão funcional, tendo-se como referência a glote (espaço anatomicamente definido pelas pregas vocais): os sistemas respiratório (componente subglótico), fonatório (componente glótico) e articulatório/ressonantal (componente supraglótico). A estreita relação entre as estruturas ao longo do trato vocal humano, e sua inerente possibilidade de mobilização constante, confere dinamismo à produção dos sons da fala (GREGIO *et al*, 2006).

A teoria acústica da produção de fala (FANT, 1960) contempla a proposição do modelo fonte-filtro para o entendimento dos ajustes realizados na produção dos sons vocálicos, por meio da descrição de correspondências acústico-articulatórias. A fonte define-se pelo produto acústico da vibração das pregas vocais, enquanto o filtro é resultante das modificações promovidas pelos ajustes realizados ao longo do trato vocal.

No caso das vogais, especificamente, a fonte é caracterizada pela fonação, que por sua vez é resultante da vibração das pregas vocais, localizadas na laringe. A corrente de ar egressiva, vinda dos pulmões, é a fonte de energia que estimula a vibração das pregas vocais e promove sua consequente ação de fonação (LIEBERMAN, BLUMSTEIN, 1988).

O filtro, por sua vez, compreende a influência da região supraglótica do trato vocal e define, pelo seu comprimento e largura, as condições de ressonância das cavidades supraglóticas (LIEBERMAN, BLUMSTEIN, 1988). Ao longo do trato vocal, os articuladores ajustam-se de diferentes formas para produzir os sons das vogais. Há articuladores ativos, como a língua, lábios, mandíbula, palato mole e pregas vocais, e ainda, os articuladores passivos, como os dentes, o palato duro e a parede posterior da faringe (KENT, READ, 1992). Tais ajustes definem cavidades com seletividade distinta para diferentes faixas de frequência do sinal acústico (as ressonâncias) e revelam a correspondência entre fenômenos de ordem acústica e articulatória (fisiológica).



KENT, READ (1992) salientam que o modelo fonte-filtro é relevante para a interpretação de diversos procedimentos de análise acústica da fala e de suas relações com os aspectos articulatórios e perceptivos dos sons.

A produção das vogais, portanto, é entendida como o resultado de uma fonte vibratória associada a um tubo ressonador (FANT, 1960). Considera-se que as vogais são produzidas pela vibração das pregas vocais, seguidas de relativa abertura do trato vocal supraglótico, que funciona como um filtro de transmissão de energia, no qual determinadas frequências são enfatizadas em detrimento de outras (KENT, READ, 1992). É característica das vogais apresentar o trato vocal livre de obstruções à passagem da corrente de ar (FANT, 1960; CUKIER, 2006).

A classificação articulatória das vogais é definida tradicionalmente pelo posicionamento vertical e horizontal da língua na cavidade oral, graus de arredondamento dos lábios e de abertura da cavidade oral, além da abertura ou fechamento do segmento velofaríngeo (GREGIO, 2006; CAMARGO *et al*, 2008).

A produção dos sons pode ser analisada por meio da análise acústica, uma vez que essa análise permite inferir os ajustes realizados pelo trato vocal supraglótico. A decomposição do sinal acústico nas três propriedades acústicas do som – frequência, intensidade e duração – permite inferir as configurações realizadas pelos articuladores durante a produção de fala (MENDES, 2003).

As vogais são caracterizadas acusticamente pelos formantes, tidos como as frequências naturais de vibração do trato vocal. Em conjunto, constituem a função de transferência do trato vocal (KENT, READ, 1992), sendo descrito o padrão de formantes em função dos parâmetros de frequência, intensidade e largura de banda.

Dessa forma, o modelo fonte-filtro possibilita estimar as ressonâncias do trato vocal (formantes) na produção das vogais. A configuração do trato vocal supraglótico definida pelos articuladores irá delimitar as propriedades acústicas de cada vogal (JOHNSON, 2000).

Segundo FANT (1960) é infinita a geração de formantes pelo trato vocal. A frequência do primeiro formante (F1) relaciona-se especialmente ao

movimento da língua no eixo vertical e da mandíbula. A frequência do segundo formante (F2) diz respeito à movimentação da língua no eixo horizontal (ântero-posterior). O terceiro formante (F3), por sua vez, é relacionado ao grau de constrição entre a língua e a faringe (KENT, READ, 1992). Entretanto, na produção das vogais, os dois primeiros formantes são apontados como os principais para a atribuição de sua identidade fonética na maioria das línguas.

Portanto, as medidas acústicas de frequências formânticas, especialmente F1 e F2, extraídas por meio de procedimentos de análise acústica, podem se relacionar com os ajustes articulatórios do posicionamento de língua na cavidade oral, identificados por técnicas de imagem. Os parâmetros de intensidade e de banda de formantes fornecem informações adicionais de resposta das cavidades de ressonância, especialmente das condições de amortecimento e ainda são pouco enfocados em caracterizações fonético-acústicas de vogais em nosso meio (MAGRI e col, 2009; OLIVEIRA, 2011).

A caracterização dos aspectos acústicos e articulatórios das vogais em fala sem alterações é fundamental para propiciar a descrição da variação de fala, e também, para fomentar a discussão sobre os eventos presentes nos falantes com alteração. MENDES (2003) referiu que as novas tecnologias têm permitido melhor compreensão dos mecanismos de produção dos sons de fala, pois possibilitam a análise simultânea do sinal de fala na perspectiva acústica e articulatória. Desta forma, abordagem desta natureza pode colaborar com avanços nas descrições fonéticas do PB e, futuramente, nortear estudos com variados distúrbios de fala.

A descrição acústica das vogais do PB apresenta estudos de variantes de alguns grupos de falantes. MEIRELLES (2008) descreve acusticamente as vogais tônicas (em palavras paroxítonas) de falantes do PB falado no Rio Grande do Sul. Há também a descrição das vogais tônicas posteriores do PB falado no Amazonas (CAMPOS, 2010).

O estudo de OLIVEIRA (2011) analisou, do ponto de vista fonético-acústico, a produção das vogais orais do PB em crianças respiradoras orais e respiradoras nasais. A descrição acústica das vogais permitiu a inferência do posicionamento de língua e a mobilização dos demais articuladores durante a

fala de respiradores orais. Os dados apresentados pela autora apresentaram mudança na estrutura formântica das vogais orais das crianças que tinham alteração na respiração (respiração oral). Além disso, nas crianças com respiração oral, os valores de intensidade de formantes também se mostraram elevados em relação às crianças com respiração nasal, o que pode ser justificado pelo aumento global da intensidade da vogal relacionado aos ajustes de hiperfunção laríngea encontrados neste grupo.

Em relação à descrição dos aspectos acústicos, usualmente os dados são coletados em Laboratórios de Fala. Xu (2010) descreve o que chamou de mitos referentes ao uso dos laboratórios de fala atualmente, e defende que esse tipo de experimento controlado pode representar com naturalidade a “fala de laboratório”. O autor apresenta argumentos que desfazem os mitos que consideram a fala de laboratório lenta, não natural, monótona, com prosódia inadequada e sem comunicação funcional. Segundo ele, é possível orientar os falantes nas coletas, de forma a não limitar o laboratório de fala com considerações infundadas.

A produção articulatória das vogais orais do PB foi estudada por PINHO *et al.* (1988), por meio da xerorradiografia (tipo de raio-X) de cabeça e pescoço. As vogais nasais, por sua vez, foram descritas por meio da mesma técnica por MASTER (1991). Ambos os estudos foram realizados com base em um falante da cidade de São Paulo. Dadas as limitações da técnica e a exposição à grande quantidade de radiação, não houve continuidade de uso, apesar de promover imagens estáticas do trato vocal de excelente qualidade para emissões de vogais sustentadas.

As caracterizações fonético-acústicas e articulatórias das línguas datam de algumas décadas. Entretanto, o desenvolvimento tecnológico tem permitido novas incursões e detalhamentos, não possíveis há pouco tempo. Tal evolução permitiu o avanço de descrições de sons isolados para abordagens dinâmicas, que levam em conta a fala em situação mais definida em termos de contextos.

Estudos que tomam como pressupostos teóricos a relação acústico-articulatória em perspectiva dinâmica reconhecem a caracterização de variabilidade das manifestações e a imprecisão na produção de fala de sujeitos

com alterações. Sujeitos com deficiência auditiva (DA) tiveram amostras de fala coletadas e avaliadas no grupo das vogais (MENDES, 2003) e das consoantes (FICKER, 2003).

No caso das vogais, foi detectada semelhança em relação à distribuição de valores de frequências formânticas nas vogais produzidas por um sujeito ouvinte e por um sujeito com DA. Entretanto, houve maior dispersão dos valores para a produção do sujeito com DA, com tendência à sobreposição de espaços acústico-articulatórios das suas produções, mais acentuada em vogais posteriores (MENDES, 2003).

O emprego de ferramentas desenvolvidas pelas Tecnologias de Fala permite o uso de técnicas de imagem para investigar o sinal de fala, especialmente, a visualização dos ajustes dos articuladores supraglóticos.

DEMOLIN *et al* (2006) descreveram o uso da IRM em tempo real para estudar a coordenação dos movimentos da articulação durante a produção das vogais. As imagens em tempo real foram comparadas às imagens estáticas similares de IRM. De acordo com os autores, a tecnologia da IRM em tempo real permite estudar os movimentos dos principais articuladores envolvidos na produção da fala: lábios, língua, laringe, mandíbula e véu palatino.

DEMOLIN e MEDEIROS (2006) apresentaram um estudo exploratório de IRM para verificar o abaixamento do véu palatino no momento da produção da vogal nasal do PB. O experimento envolveu um sujeito do sexo masculino, falante do PB. Foram analisados cinco estímulos de pares mínimos contendo as vogais orais *versus* vogais nasais. As imagens de IRM identificaram a posição do véu palatino elevado nas vogais orais e abaixado nas vogais nasais. Além disso, observou-se diferença na posição da língua entre vogais orais e nasais.

Os autores referem ainda que não foram gravados dados acústicos simultâneos à coleta dos dados articulatórios, pois o estudo pretendia descrever a posição dos articuladores. Dessa forma, salientaram a necessidade de outros estudos com maior número de sujeitos e a captura de dados acústicos para extração de medidas acústicas que possam confirmar os achados articulatórios.

Ainda sobre as vogas nasais, o estudo de GREGIO (2006) apontou que, em relação às vogais orais, as vogais nasais apresentam três diferentes fases

marcadas pelo posicionamento dos articuladores (dorso de língua e palato mole) são elas: fase oral, fase nasal e fase nasal com movimentação de ponta de língua.

De acordo com o estudo de BERTI (2010), o equipamento que apresenta melhor custo-benefício para a investigação da produção de fala é o aparelho de USG. A autora fez o levantamento de outros instrumentos, tais como, raio-X, IRM, EMA, micro raio-X, e considerou os seguintes requisitos de análise: segurança, portabilidade, resolução temporal, custo do equipamento, imagem do articulador passivo, do trato vocal (parcial) e, da superfície da língua. Dentre estes requisitos, e suas contribuições de análise da produção de fala, a técnica de USG se destaca.

As imagens de USG são capturadas por meio de transdutores que têm como elemento ativo um material piezoelétrico. Por material piezoelétrico entende-se um elemento, como a cerâmica, que sofre deformação mecânica sob a ação de um campo elétrico e que, ao mesmo tempo, gera um campo elétrico quando sofre a deformação mecânica. A descoberta do efeito piezoelétrico ocorreu em 1880 pelos irmãos Curie. A primeira aplicação ocorreu durante a Primeira Guerra Mundial para o desenvolvimento de sonares ultrassônicos (PEREIRA, 2010; SANTOS, 2010).

No corpo humano, o método é usado para diagnóstico de imagem de órgãos do organismo. A USG capta sons com frequências superiores a 20 kHz (Hz). Os pulsos elétricos emitidos pelo ultrassom provocam uma vibração do transdutor piezoelétrico que emite uma onda sonora. O som é refletido, produzindo o eco, quando encontra diferentes densidades das estruturas do corpo. As imagens produzidas são efeito do inverso da sequência anterior percorrida: o retorno da onda sonora pelo transdutor transforma a vibração em pulsos elétricos que o aparelho de USG transforma em imagem pelo processo de *scanner* (BAUERMANN, 2009).

Segundo KOCJANCIC (2008), a USG configura-se como um método eficiente e não invasivo para a descrição dos movimentos da língua durante a fala sem nenhuma obstrução, relacionada ao posicionamento dos articuladores. ZHARKOVA *et al* (2008) realizaram estudo sobre a coarticulação da língua – natureza e variabilidade - em adultos e crianças por meio da USG.

Ainda são restritos os estudos que relacionam os aspectos acústicos e articulatórios da produção das vogais do PB. A possibilidade de descrever e relacionar os ajustes articulatórios (por meio de imagens da configuração supraglótica com USG) com as medidas acústicas referentes aos padrões formânticos pode trazer nova perspectiva de descrição e detalhamento dos processos dinâmicos da fala. Os fundamentos da teoria acústica de produção da fala (FANT, 1960) propiciam esta discussão.

### 3. METODOLOGIA

O capítulo de metodologia apresenta as duas coletas de dados realizadas separadamente – acústicos e de imagem de USG (estudo de caso). O projeto de pesquisa foi submetido ao Comitê de Ética e Pesquisa da PUC-SP e aprovado com o Protocolo de Pesquisa nº224/2011 (anexo 1).

#### 1. Dados Acústicos

##### 1.1. Sujeitos

Foram registrados 16 sujeitos, falantes do PB, variante paulistana, na faixa etária de 20 a 35 anos (média de 26 anos e 7 meses), sendo 8 falantes do gênero masculino (faixa etária média de 26 anos e 8 meses) e 8 falantes do gênero feminino (faixa etária média de 26 anos e 6 meses).

Como critério de inclusão, os sujeitos apresentaram ausência de dificuldades e/ou queixas relacionadas à produção da fala, de alterações crânio-faciais e auditivas. Os indivíduos responderam a um questionário elaborado especialmente para a pesquisa (anexo 2), com o objetivo de levantar as informações referentes aos critérios acima mencionados (Quadro 1).

Quadro 1 – Descrição dos sujeitos participantes: gênero, idade e escolaridade

<b>Sujeito</b>	<b>Gênero</b>	<b>Idade</b>	<b>Escolaridade</b>
S1	masculino	24	ensino superior
S2	masculino	23	ensino superior
S3	masculino	26	ensino superior
S4	masculino	29	ensino superior
S5	masculino	24	ensino superior
S6	masculino	30	ensino superior
S7	feminino	27	ensino superior
S8	feminino	26	ensino superior
S9	feminino	29	ensino superior
S10	feminino	27	ensino superior
S11	masculino	34	ensino superior

S12	feminino	26	ensino superior
S13	masculino	25	ensino superior
S14	feminino	28	ensino superior
S15	feminino	20	ensino superior
S16	feminino	30	ensino superior

Os participantes foram informados da pesquisa, orientados sobre os procedimentos a serem realizados, além de solicitados para a assinatura do termo de consentimento livre e esclarecido (TCLE) (anexo 3).

## 1.2. Local de coleta e instrumental

A coleta dos dados acústicos foi realizada no Laboratório de Fala e Voz do CEFAC – Saúde e Educação. O instrumental utilizado foi: cabina acústica; computador com placa de som; mesa de som *Behringer*, modelo BCA 2000; microfone unidirecional AKG C4100; *Software Sound Forge Edit* (versão 10).

## 1.3. Procedimentos de coleta de dados

Em cada gravação, os sujeitos foram acomodados na cabina acústica, em posição sentada e ereta. O microfone foi colocado a uma distância fixa de 5 cm da comissura labial esquerda do falante. O ambiente da gravação foi controlado para impedir a interferência de ruídos externos.

No início da gravação, foi solicitado aos sujeitos que falassem o nome, a idade, a data da gravação e um trecho de fala semi-espontânea, a partir da solicitação: “Fale sobre a cidade onde você nasceu” para ajustar os controles de ganho da gravação.

Após a primeira etapa, cada sujeito realizou a emissão das sete vogais orais do PB inseridas em sentenças-veículo com estrutura “Digo [pVpv] baixinho”, em que **V** é a vogal oral produzida em posição tônica ([a], [ε], [e], [i], [ɔ], [o], [u]) e v vogal similar àquela indicada em **V**, porém em posição pós-tônica.



As frases foram impressas em folhas (A4) separadas e apresentadas aos sujeitos uma a uma do lado externo da cabine (Figura 1).

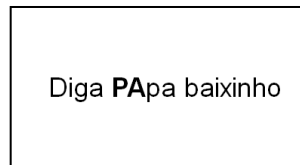


Figura 1 – Exemplo de sentença-veículo apresentada ao sujeito durante a gravação de *corpus* de pesquisa

A instrução dada ao sujeito pelo pesquisador quanto à leitura das frases é apresentada na Figura 2.

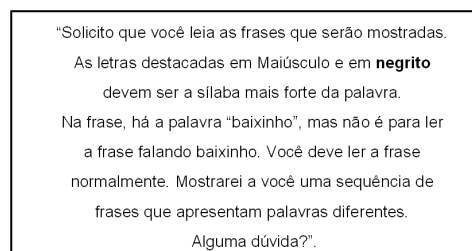


Figura 2 – Instrução para a leitura do *corpus* de pesquisa

A apresentação das frases ocorreu em ordem aleatorizada, com cinco repetições de cada estímulo.

#### 1.4 Procedimentos de análise de dados – Frequências Formânticas

As gravações em áudio foram digitalizadas no formato *.wav*, frequência de amostragem de 22.050Hz, 16 bits e analisadas pelo *software* de livre acesso *Praat* – versão 5.1.34 (BOERSMA, WEENINK, 2002) - disponível em: <http://www.fon.hum.uva.nl/praat> - (*Institute of Phonetic Sciences - University of Amsterdam*) para caracterização acústica das frequências formânticas (Hz) das sete vogais orais do PB (F1, F2 e F3).

Para a extração das medidas acústicas das vogais orais do PB foram reproduzidos os procedimentos abaixo:

- 1º- Segmentação das sentenças-veículo em todas as repetições com demarcação da **vogal em posição tônica** no vocábulo chave pVpv;
- 2º- Extração das medidas das frequências formânticas (F1, F2 e F3), conforme Figuras 3, 4 e 5.

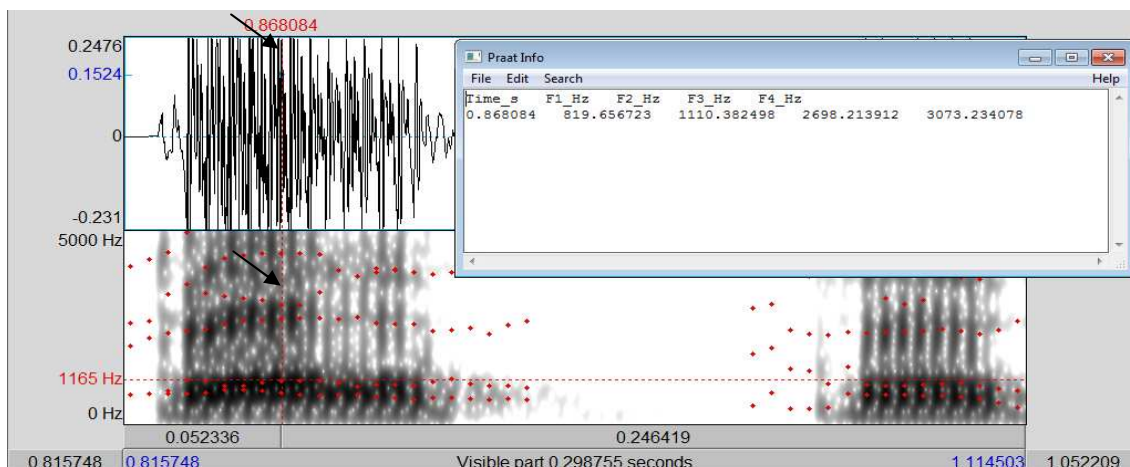


Figura 3 – Traçado de forma da onda (janela superior) e espectrograma de banda larga (janela inferior): seleção de ponto estacionário (indicado por flecha) da emissão da vogal [a] em posição tônica, com medidas geradas pelo extrator automático (janela *Praat Info*)

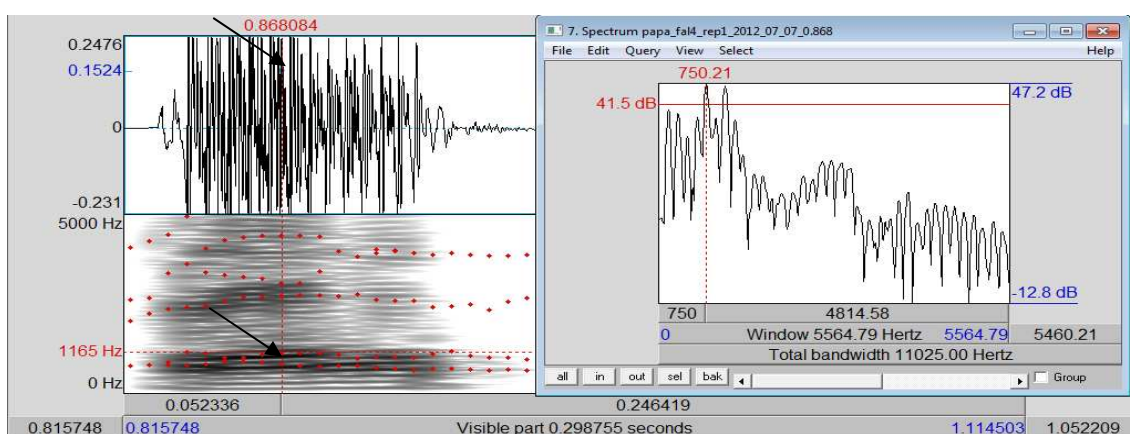


Figura 4 – Traçado de forma da onda (janela superior) e espectrograma de banda estreita (janela inferior): seleção de ponto estacionário (indicado por flecha) da emissão da vogal [a] em posição tônica e geração do espectro FFT (janela *Spectrum*) para conferência dos picos espectrais

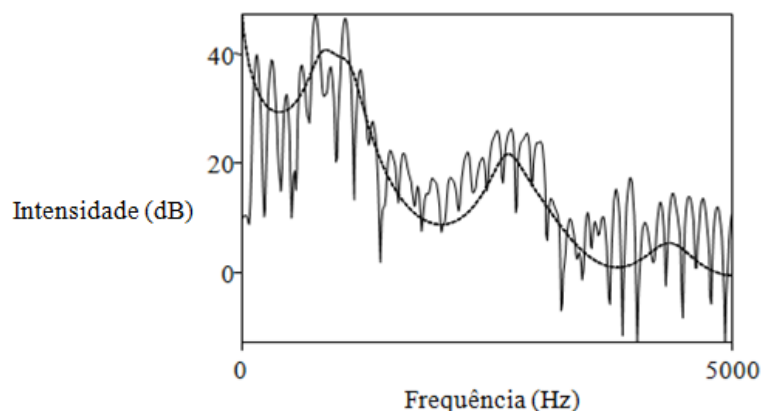


Figura 5 – Espectro FFT (linha contínua) e espectro LPC (linha tracejada)

As medidas acústicas de frequências formânticas foram submetidas ao *script* CartaFormantesLog.praat (aplicado ao *software Praat*) cedido e desenvolvido por Eduardo Velásquez (adaptado de Mietta Lennes, 2004<sup>1</sup>) para a confecção do trapézio das vogais orais do PB (F1 x F2).

As medidas acústicas foram submetidas à análise estatística. Os dados de frequências formânticas (F1, F2 e F3) foram explorados em termos de apresentação de médias, desvio-padrão (DP) e intervalos de confiança (IC), a fim de traçar estimativas de referência para futuros estudos. Além disso, os dados foram submetidos à análise discriminante com referência ao poder discriminatório das medidas de frequências formânticas (F1, F2 e F3) para detecção do gênero dos falantes e das diferentes vogais orais estudadas.

---

<sup>1</sup> *Script*: CartaFormantesLog.praat. Desenvolvido por Eduardo Velásquez (adaptado de Mietta Lennes, 2004).

## **2. Dados de imagem de USG – Estudo de caso**

### **2.1. Sujeito**

Um sujeito falante do PB, variante paulistana, gênero masculino, 29 anos de idade, foi submetido à captura simultânea das imagens de USG e registros em áudio.

O sujeito cujas imagens foram coletadas nessa etapa participou também da etapa de coleta dos dados acústicos (item 1). Portanto, já contemplava os critérios de inclusão referidos e havia sido informado sobre os termos da pesquisa, tendo assinado o TCLE.

### **2.2. Local de coleta e instrumental**

A coleta das imagens ultrassonográficas, acompanhada de gravação em áudio, foi realizada no setor de Radiologia do Hospital Santa Paula em São Paulo. O setor de Radiologia realiza diagnósticos por imagens: USG, IRM, tomografia computadorizada, raio-X, etc.

A captação das imagens ultrassonográficas foi realizada com o aparelho de USG *Sonix SP – Ultrasonix*. O instrumental utilizado para coleta simultânea de áudio foi: MD-minidisc MZ-R70 *Sony*, microfone unidirecional *Áudio 400 DSP Plantronics*.

### **2.3. Das demandas e adaptações para a coleta de imagens de USG de fala**

A USG não é usualmente utilizada para captura e diagnóstico de fala em nosso meio. Dessa forma, as coletas experimentais demonstraram a escassez de adaptação do instrumental para uma coleta controlada e possível de ser replicada, tanto no que diz respeito às configurações de gravação (imagens estáticas ou sequência de vídeo) do aparelho de USG, quanto aos transdutores de USG já existentes para a análise de outros órgãos do corpo humano.

Portanto, a cada coleta experimental realizada, uma nova demanda surgia no sentido de adequação do instrumental para coletas de fala.

A primeira escolha foi quanto ao aparelho de USG. O local da coleta dispunha de dois equipamentos de USG de qualidade similar, porém com funções diferentes. Após algumas tentativas, foi selecionado para a coleta, o aparelho de USG que permitia a captura de vídeos (vários quadros por segundo – *frames*), uma vez que era preciso o registro de natureza dinâmica da fala.

A segunda escolha referiu-se à definição do transdutor a ser utilizado, dentre aqueles comumente utilizados: convexo, endocavitário, linear, cardíaco, 4D, *doppler* sem imagens, etc. Os dois critérios para a definição pelo transdutor endocavitário foram: a qualidade das imagens e estruturas que o transdutor permitia visualizar, além de seu encaixe na região submandibular (região que possibilita analisar o posicionamento e o deslocamento da língua) (SCOBIE *et al*, 2008).

Entretanto, apesar do melhor encaixe do transdutor escolhido na região submandibular, a cada tentativa de começar a falar uma sequência de frases, o transdutor deslocava-se do lugar inicial e, por consequência, a imagem também era alterada.

Dessa forma, várias sessões de coleta foram conduzidas para a adequação dos procedimentos, especialmente no sentido de desenvolver algum mecanismo de fixação do transdutor na região submandibular durante a fala, ao mesmo tempo em que permitisse o movimento natural das estruturas envolvidas na fala, especialmente da mandíbula.

Foi desenvolvido, portanto, um suporte que permitiu a acoplagem do transdutor e a sua fixação durante a produção de fala. O suporte em questão é apresentado em forma de esquema na Figura 6.

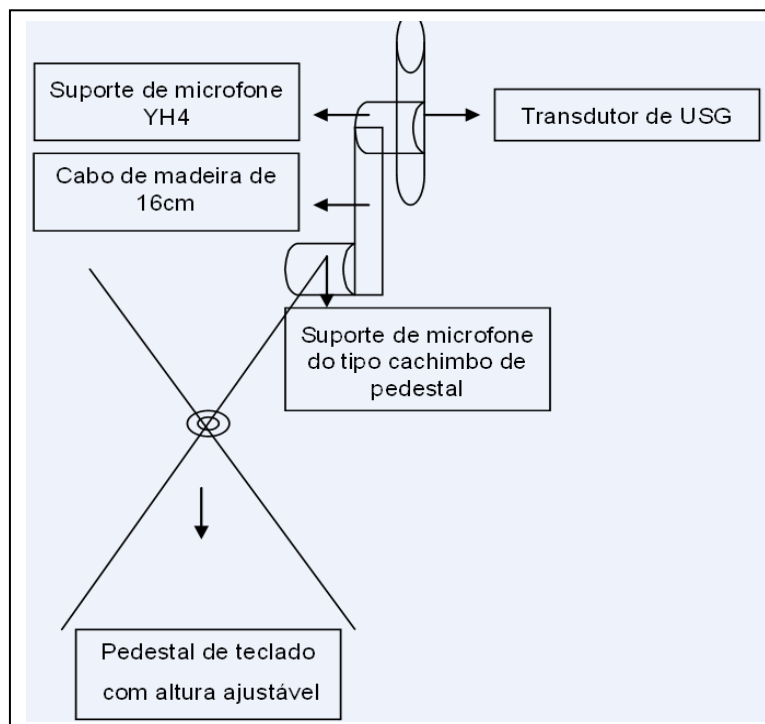


Figura 6 – Esquema ilustrativo do suporte desenvolvido para a fixação do transdutor de ultrassonografia para coletas de fala

O desenvolvimento do suporte (Figura 7 – (a) e (b)) foi fundamental para a fixação do transdutor de USG na posição necessária para a coleta de dados de fala. O suporte desenvolvido consiste em: pedestal de teclado ajustável à altura do sujeito (Figura 7 – (c)); cabo de madeira de 16cm para ligação entre os suportes (Figura 7 – (d)); suporte de microfone YH4 (Figura 7 – (e)); suporte de microfone do tipo cachimbo de pedestal (Figura 7 – (f)); que se prende ao transdutor de USG (Figura 7 – (g) e (h)).

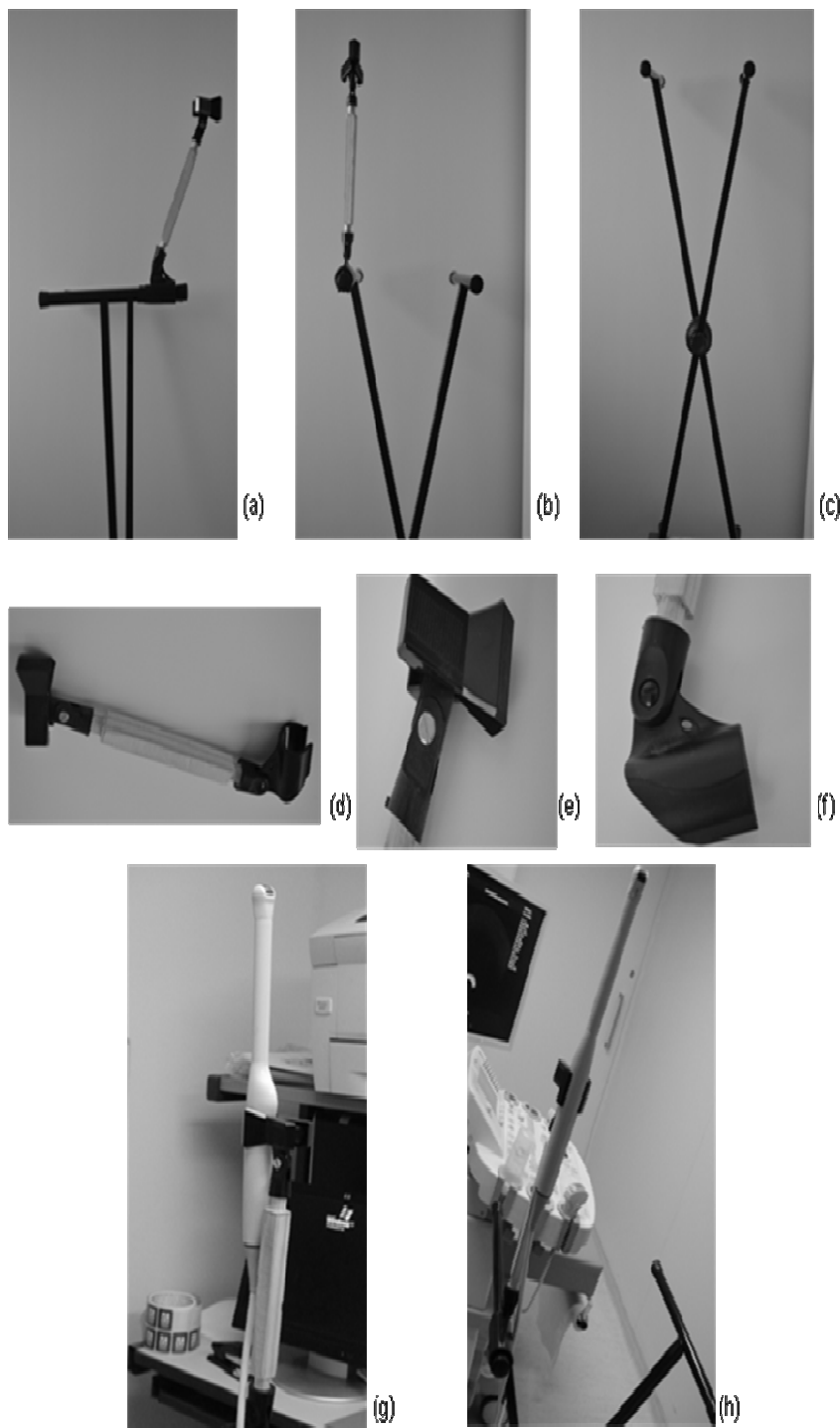


Figura 7 – Suporte para acoplagem do transdutor de ultrassonografia desenvolvido para a presente pesquisa (a) visão lateral do suporte montado e (b) visão frontal do suporte montado; (c) pedestal de teclado; (d) cabo de madeira de 16cm; (e) suporte de microfone YH4; (f) suporte de microfone do tipo cachimbo de pedestal; (g) visão lateral do suporte com o transdutor de USG; (h) visão frontal do suporte com o transdutor de ultrassonografia

## 2.4. Procedimentos de coleta de dados

O sujeito selecionado foi orientado quanto aos procedimentos de gravação: permanecer em pé com postura ereta, olhar em um ponto fixo (Figura 8 – (a)) indicado por uma figura na parede (Figura 8 – (b)) e seguir as instruções da sequência de gravação explicadas pelo pesquisador.

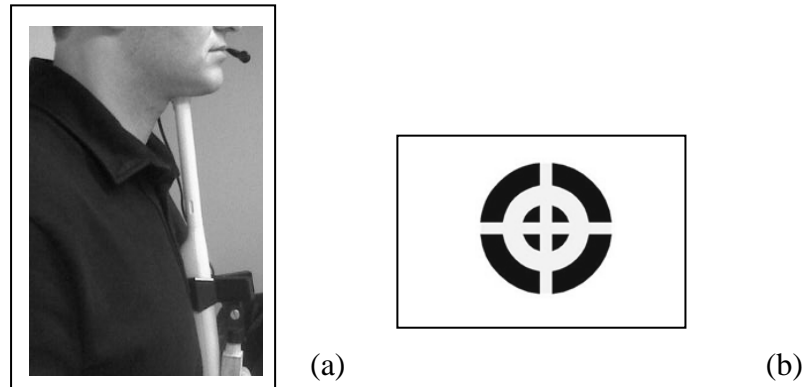


Figura 8 – Posicionamento do sujeito participante da pesquisa (a) e figura alvo para manutenção da posição (b)

Para o início da coleta, após o posicionamento adequado do sujeito, o transdutor de USG foi posicionado na região submandibular do sujeito (imagens em plano sagital), devidamente preparado com o gel Ultra-Gel (MultiGel Ind e Com.) e encaixado no suporte desenvolvido (Figura 8 – (a)).

No primeiro momento, foram registradas as imagens do trato vocal em posição habitual (repouso) e com a execução de uma deglutição de saliva para a visualização do contorno da língua e referência de seu deslocamento mínimo e máximo.

Num segundo momento, o sujeito foi orientado quanto à realização de emissão das sete vogais orais do PB, descritas quanto ao *corpus* de gravação acústica no item 1.3 anteriormente descrito.

Os dados de imagens de USG foram capturados de forma concomitante aos dados acústicos. As medidas acústicas extraídas dessa gravação foram utilizadas apenas para a análise integrada dos dados acústicos, conforme item 3.



## 2.5. Procedimentos de análise de dados – Imagens de USG

Os vídeos gravados de USG da produção das repetições das vogais orais (inseridas em sentença-veículo em posição tônica de acentuação) foram digitalizados, editados e sincronizados às gravações de áudio pelo *software VideoPad Video Editor*, conforme Figura 9.



Figura 9 – Edição e sincronização dos vídeos gerados pelo aparelho de ultrassonografia com os registros de áudio pelo *software VideoPad Video Editor*

Foram selecionadas imagens do posicionamento da língua a partir dos vídeos dos seguintes momentos: posição habitual (repouso), deglutição e período estacionário das sete vogais orais do PB ([a], [ɛ], [e], [i], [ɔ], [o], [u]) em posição tônica. As imagens selecionadas de USG também passaram por tratamento pelo *software PhotoPad Image Editor* que propiciou a geração de suas imagens em negativo (Figura 10), de forma a permitir a melhor definição do contorno de língua.

A partir do tratamento das imagens, duas abordagens de interpretação de dados foram realizadas. Na abordagem 1, foram gerados esquemas de deslocamento do dorso de língua – nos eixos vertical e horizontal – a partir da postura habitual de língua (final da deglutição), tomado como ponto de referência. Na Figura 11 estão indicados os esquemas traçados para identificação e geração do contorno do dorso de língua: a divisão da tela de visualização em padrão quadriculado permitiu estimar o

deslocamento do dorso de língua para as distintas emissões vocálicas (a) e a linha contínua sobre a imagem da língua para delinear o contorno do dorso de língua (b).

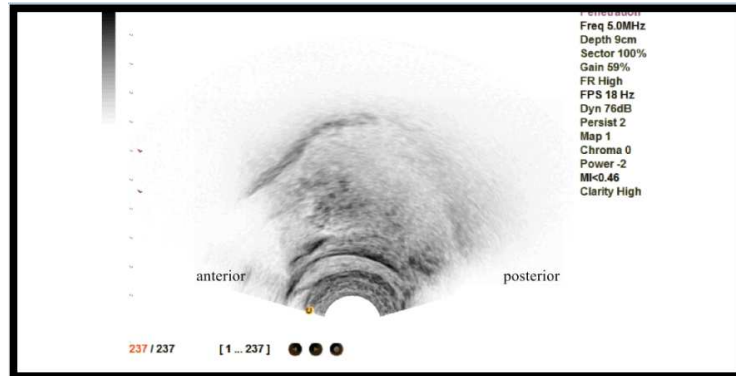


Figura 10 – Imagem de ultrassonografia em negativo

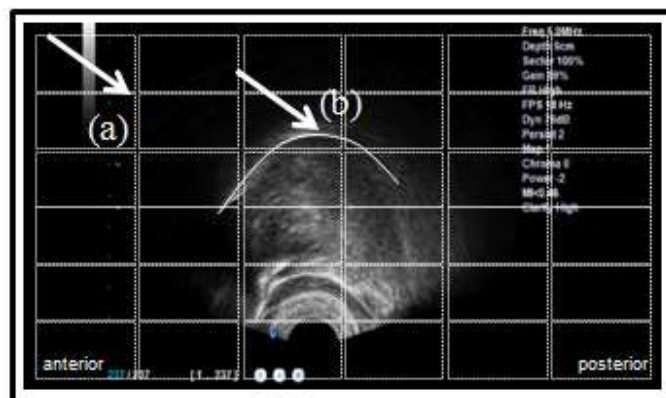


Figura 11 – Esquema de análise das imagens de ultrassonografia: padrão quadriculado para estimar o deslocamento do dorso de língua para as distintas emissões vocálicas (a) e a linha contínua para delinear o contorno do dorso de língua (b) na abordagem 1

Na abordagem 2, as imagens de USG foram submetidas ao *software* de livre acesso *Osiris* ([http://www.sim.hcuge.ch/osiris/01\\_Osiris\\_Presentation\\_EN.htm](http://www.sim.hcuge.ch/osiris/01_Osiris_Presentation_EN.htm)) para a mensuração do deslocamento de língua nos eixos vertical e horizontal em milímetros (mm). Foram tomados como referência para a extração em mm os seguintes pontos: (a) ponto de partida no centro inferior da imagem; (b) deslocamento vertical máximo até o ponto do contorno de língua visualizado; (c) e (d) deslocamento horizontal

máximos, respectivamente, até os pontos anterior e posterior visualizados do contorno de língua (Figura 12).

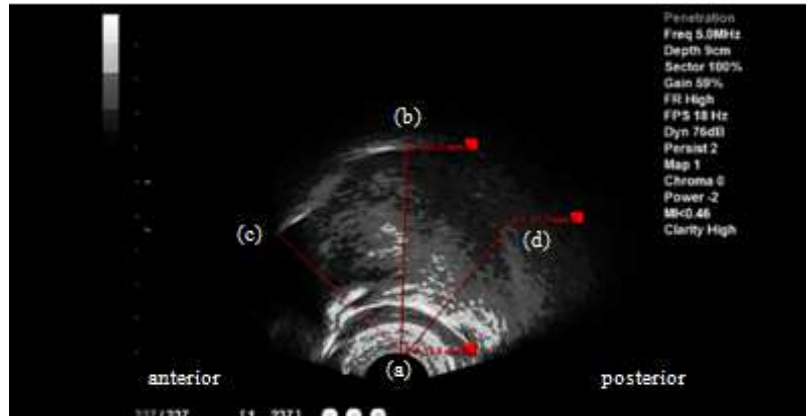


Figura 12 – Pontos de referência da extração das medidas (em milímetros) de deslocamento do dorso de língua a partir das imagens de ultrassonografia pelo *software Osiris*: (a) ponto de partida no centro inferior da imagem; (b) deslocamento vertical máximo até o ponto do contorno de língua visualizado; (c) e (d) deslocamento horizontal máximos, respectivamente, até o ponto anterior visualizado e o ponto posterior visualizado do contorno de língua para a abordagem 2

### 3. Análise integrada de dados acústicos e de imagens de USG

As informações de dados acústicos e de imagens de USG foram tratadas por meio da análise discriminante e análise de correlação canônica por meio do *software XLStat (Addinsoft)*.

Para a análise discriminante foi avaliado o poder segregatório das análises de imagens geradas pelas abordagens 1 e 2 (descritas no item 2.5) em prever as vogais orais do PB.

A análise de correlação canônica foi aplicada com o fim de avaliar as correlações entre as informações de deslocamento de dorso de língua (eixos vertical e horizontal) extraídas das imagens de USG (abordagens 1 e 2) as medidas de frequências formânticas (F1, F2 e F3) das amostras de áudio extraídas no momento da coleta das imagens do sujeito avaliado.

## 4. RESULTADOS

### 1. Dados acústicos - Medidas de Frequências Formânticas

As medidas acústicas das frequências formânticas (F1, F2 e F3) das sete vogais orais do PB são apresentadas na Tabela 1.

**Tabela 1** – Valores de média, desvio-padrão (DP) e intervalos de confiança (IC) das medidas de frequências (Hz) dos três primeiros formantes (F1, F2 e F3) das vogais orais dos falantes dos gêneros masculino e feminino

	F1 Masculino Média (Hz) (DP) IC	F1 Feminino Média (Hz) (DP) IC	F2 Masculino Média (Hz) (DP) IC	F2 Feminino Média (Hz) (DP) IC	F3 Masculino Média (Hz) (DP) IC	F3 Feminino Média (Hz) (DP) IC
[a]	803 (79) de 779 a	996 (93) de 967 a	1204 (129) de 1164 a	1489 (107) de 1456 a	2638 (120) de 2601 a	2772 (255) de 2693 a
[ε]	828 576 (36) de 565 a	1025 711 (61) de 691 a	1244 1822 (97) de 1792 a	1522 2208 (101) de 2176 a	2675 2669 (147) de 2623 a	2851 3025 (120) de 2987 a
[e]	587 381 (28) de 372 a	730 473 (58) de 455 a	1853 1920 (163) de 1869 a	2241 2432 (135) de 2389 a	2715 2766 (152) de 2719 a	3063 3161 (135) de 3119 a
[i]	390 273 (38) de 261 a	492 325 (52) de 309 a	1970 2111 (140) de 2068 a	2475 2695 (115) de 2660 a	2813 3072 (174) de 3018 a	3204 3515 (111) de 3481 a
[ɔ]	285 657 (65) de 636 a	341 726 (106) de 693 a	2155 1590 (514) de 1431 a	2731 1066 (117) de 1029 a	3126 2939 (451) de 2799 a	3550 2777 (214) de 2710 a
[o]	677 465 (67) de 444 a	759 531 (100) de 498 a	1750 1338 (595) de 1151 a	1103 938 (129) de 896 a	3079 2703 (304) de 2607 a	2844 2832 (261) de 2747 a
[u]	486 400 (98) de 369 a	564 416 (98) de 385 a	1525 1242 (733) de 1015 a	980 895 (356) de 785 a	2798 2650 (314) de 2553 a	2918 3002 (222) de 2934 a
	430	446	1470	1005	2748	3071

A distribuição dos valores de frequências formânticas (F1, F2 e F3) para os grupos de falantes dos gêneros masculino e feminino é apresentada na Figura 13.

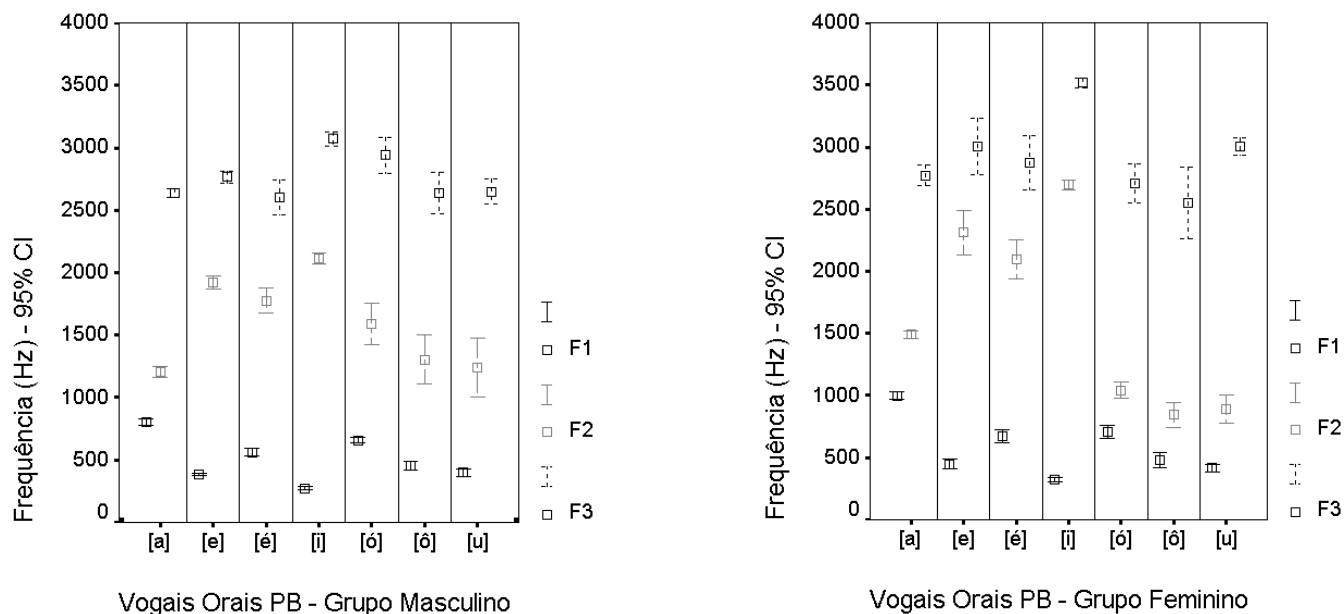


Figura 13 – Distribuição dos valores de frequências formânticas (F1, F2 e F3) das vogais orais do PB de falantes dos gêneros masculino e feminino

Os valores da distribuição das frequências de F1 e F2 são apresentadas na Figura 14 referente ao trapézio vocálico das vogais orais para os grupos de falantes dos gêneros masculino e feminino.

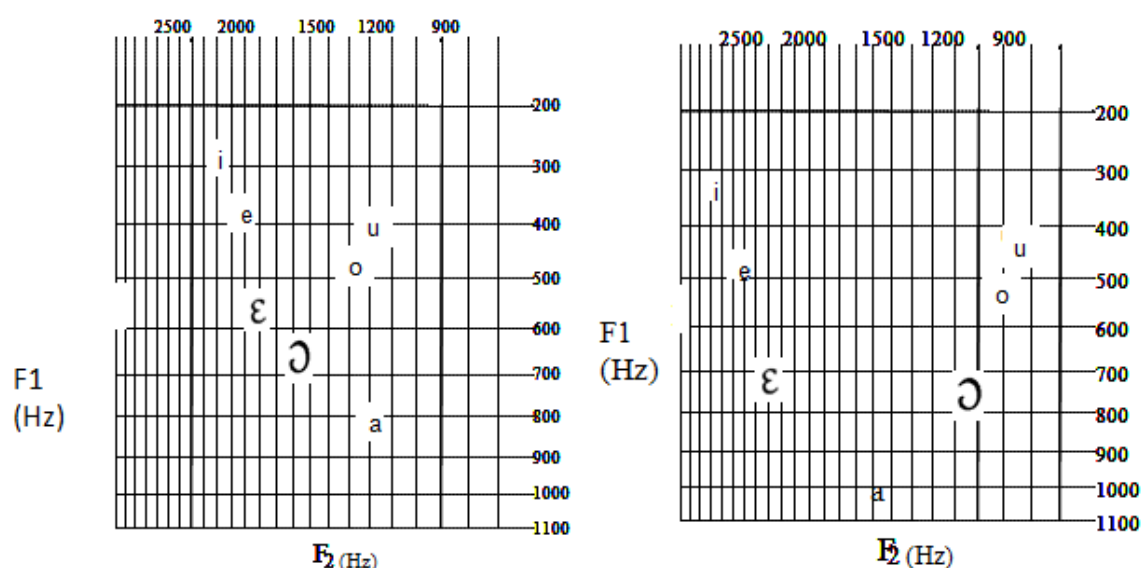


Figura 14 – Trapézio das vogais orais traçado por meio do *script* CartaFormantes.Log.praat para os gêneros masculino e feminino

A análise discriminante aplicada às medidas acústicas (F1, F2 e F3) por gênero dos falantes revelou poder segregatório parcial (entre 73, 74 e 77,49%), conforme Figura 15. O conjunto de variáveis que influenciou a segregação dos gêneros masculino e feminino (Figura 16) foi, em ordem decrescente, F3 (fator de correlação de 70,5%), e F1 (fator de correlação de 41,7%).

Gêneros de \ a	Masculino	Feminino	Total	% correto
Masculino	205	73	278	73,74%
Feminino	61	210	271	77,49%
Total	266	283	549	75,59%

Figura 15 - Matriz de confusão para os resultados de validação cruzada da análise discriminante de medidas acústicas (F1, F2 e F3) e gêneros

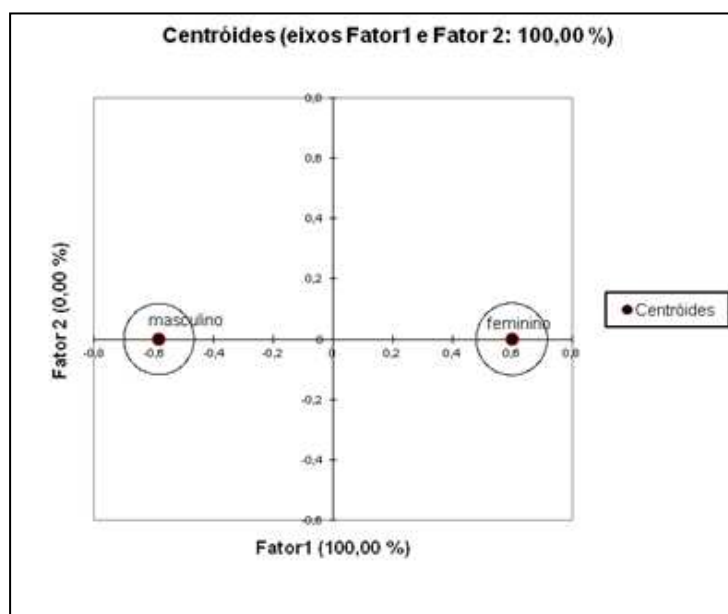


Figura 16 – Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 para identificação do gênero do falante)

A análise discriminante aplicada às medidas acústicas (F1, F2 e F3) por vogais (diferenciada por grupos de falantes dos gêneros masculino e feminino) revelou resultados similares. No caso do gênero masculino, o poder segregatório das medidas acústicas (F1, F2 e F3) para o conjunto de vogais foi de 78,06%. Os dados referentes ao poder segregatório (%) de cada vogal podem ser observados na Figura 17.

Vogais Orais de \ a	[a]	[ε]	[e]	[i]	[ɔ]	[o]	[u]	Total	% correto
[a]	38	0	0	0	2	0	0	40	95,00%
[ε]	0	37	0	0	2	0	0	39	94,87%
[e]	0	0	37	0	0	2	1	40	92,50%
[i]	0	0	1	39	0	0	0	40	97,50%
[ɔ]	8	3	0	0	28	1	0	40	70,00%
[o]	0	6	4	0	2	13	14	39	33,33%
[u]	0	4	5	1	1	4	25	40	62,50%
Total	46	50	47	40	35	20	40	278	78,06%

Figura 17 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (valor segregatório das medidas de F1, F2 e F3 de vogais orais para o gênero masculino)

No caso do gênero feminino, o poder segregatório das medidas acústicas (F1, F2 e F3) para o conjunto de vogais foi de 82,66%. Os dados referentes ao poder segregatório (%) de cada vogal podem ser observados na Figura 18.

Vogais Orais de \ a	[a]	[ε]	[e]	[i]	[ɔ]	[o]	[u]	Total	% correto
[a]	37	0	0	0	3	0	0	40	92,50%
[ε]	0	38	0	0	0	0	0	38	100,00%
[e]	0	2	32	4	0	0	0	38	84,21%
[i]	0	0	0	40	0	0	0	40	100,00%
[ɔ]	3	0	0	0	29	7	0	39	74,36%
[o]	0	0	0	0	8	17	11	36	47,22%
[u]	0	4	0	0	0	5	31	40	77,50%
Total	40	44	32	44	40	29	42	271	82,66%

Figura18 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (valor segregatório das medidas de F1, F2 e F3 de vogais orais para o gênero feminino)

O conjunto de variáveis que influenciou a segregação de vogais (Figura 17 e 19) para falantes do gênero masculino concentrou-se em medidas de F1 (93,59%) e F2 (3,98%). Os dados de distribuição das centróides dos valores de F1, F2 e F3 são apresentados na Figura 19.

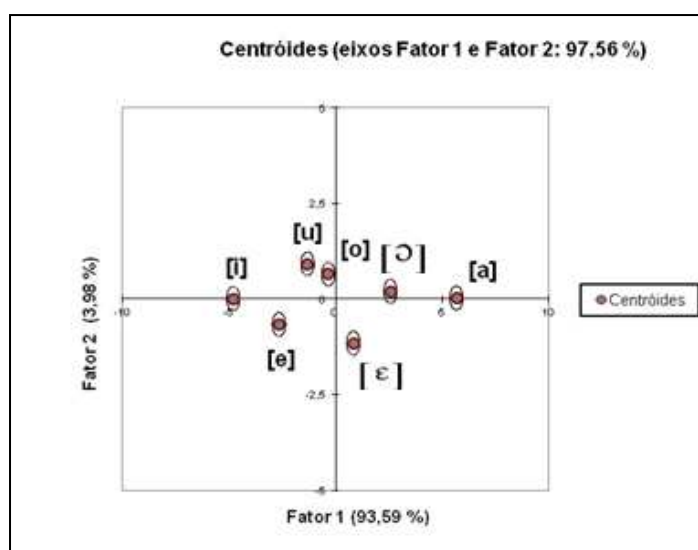


Figura 19 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 por vogais orais para falantes do gênero masculino)



O conjunto de variáveis que influenciou a segregação de vogais (Figura 18 e 20) para falantes do gênero feminino concentrou-se em medidas de F2 e F3 (84,33%) e F1 (15,33%). Os dados de distribuição das centróides dos valores de F1, F2 e F3 são apresentados na Figura 20.

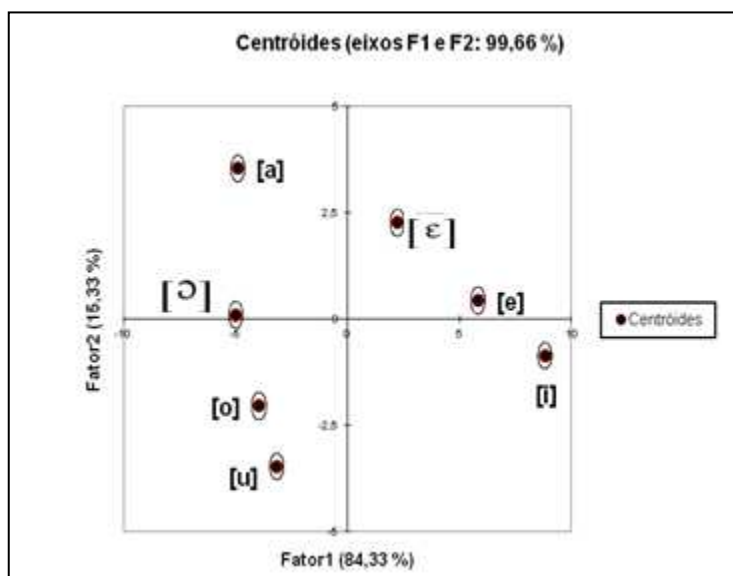


Figura 20 - Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3 por vogais orais para falantes do gênero feminino)

## 2. Dados de imagens de USG

As imagens de USG extraídas dos vídeos coletados são apresentadas nas Figuras 21, 22, 23, 24, 25 e 26.

Na Figura 21 está indicada uma imagem referência para acompanhamento das imagens subsequentes (abordagem 1), sendo 1. a imagem em corte sagital gerada pelo aparelho de USG e pelo transdutor endocavitário; 2. o traçado branco que indica a posição do dorso de língua; 3. a estrutura do corpo da língua; 4 e 5. indicam a direção da imagem, respectivamente anterior e posterior.

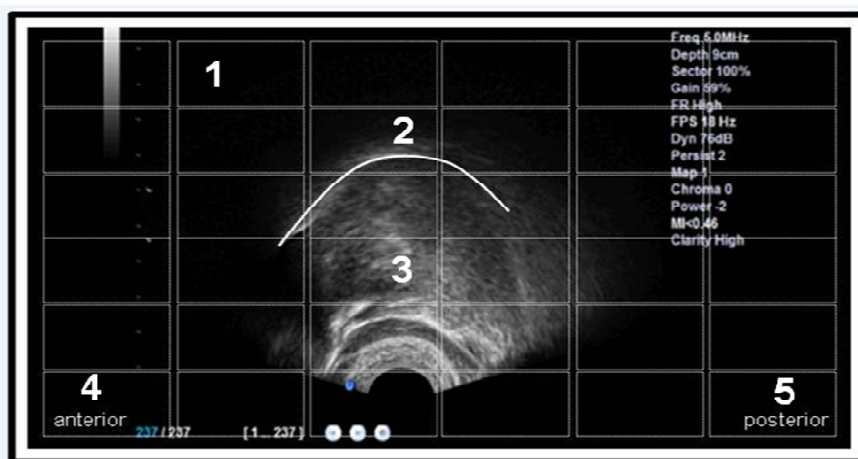


Figura 21 – Imagem de ultrassonografia (abordagem 1), sendo: 1. a imagem em corte sagital permitida pelo aparelho de ultrassonografia e pelo transdutor endocavitário; 2. o traçado branco que indica a posição do dorso de língua; 3. a estrutura do corpo da língua; 4 e 5. indicam a direção da imagem, respectivamente anterior e posterior

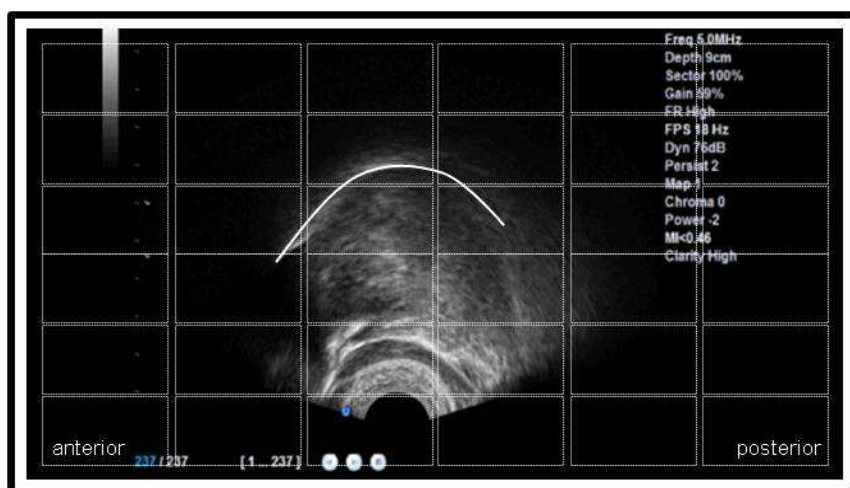


Figura 22 – Imagem de ultrassonografia (abordagem 1) extraída no momento de postura habitual (repouso) de língua. O traçado branco indica a posição do dorso de língua

A sequência de uma deglutição realizada pelo sujeito para estimativa de referência de extensão dos movimentos do dorso de língua na fala é apresentada na Figura 23.

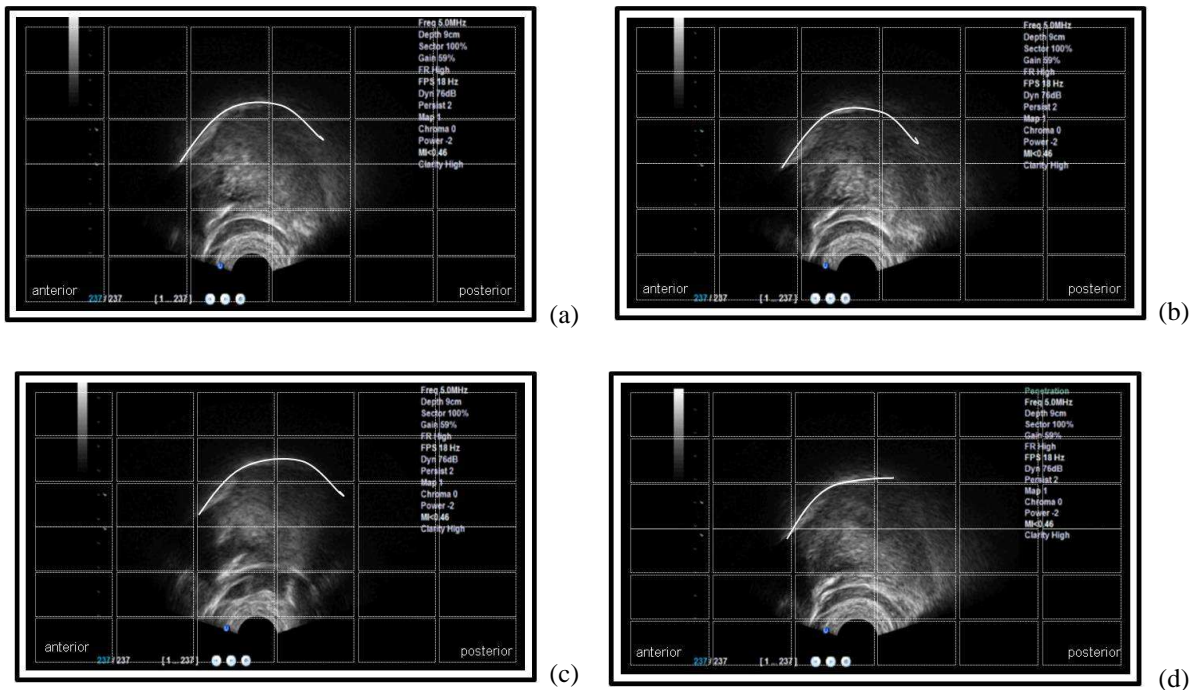
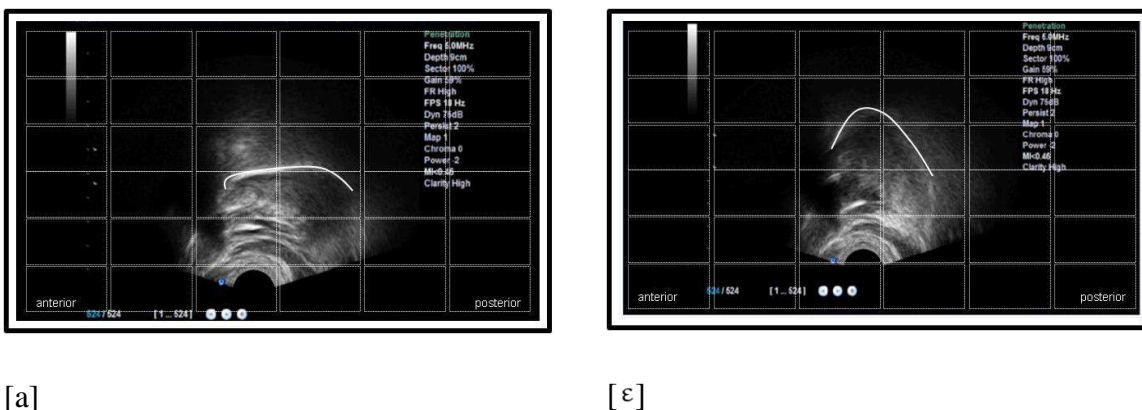
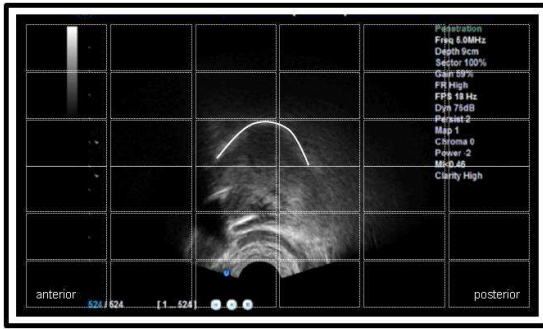


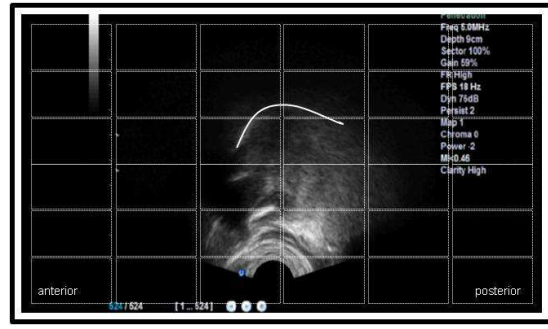
Figura 23 –Imagens de ultrassonografia (abordagem 1) extraídas (a) no momento de preparação da língua para o início da deglutição; (b) no momento de início da deglutição; (c) no momento da deglutição propriamente dita; (d) ao final da deglutição

Os traçados da postura do dorso de língua para as sete vogais orais do PB são apresentadas na Figura 24 em imagens selecionadas no período estacionário de cada uma delas e de suas respectivas imagens em negativo (Figura 25). Na Figura 26 é apresentado a sobreposição dos traçados que acompanham o contorno da língua em cada vogal. Vale ressaltar que não foi possível delinear o contorno do dorso de língua para a vogal [u].

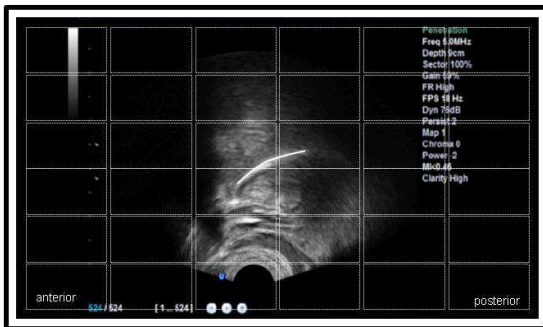




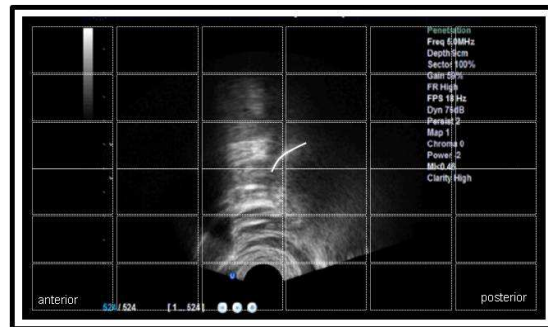
[e]



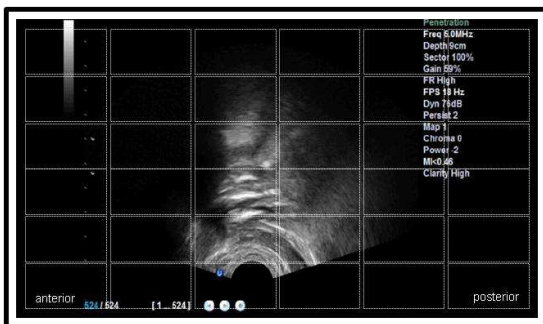
[i]



[ɔ]



[o]

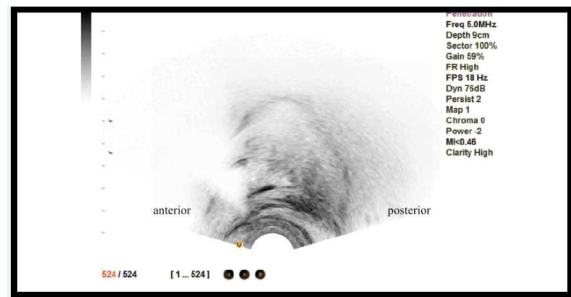


[u]

Figura 24 – Imagens de ultrassonografia (abordagem 1) das vogais orais: vogal [a]; vogal [ɛ]; vogal [e]; vogal [i]; vogal [ɔ]; vogal [o]; vogal [u]. O traçado da linha branca acompanha o contorno da língua, exceto na vogal [u], em que não foi possível a sua visualização dado o movimento posterior



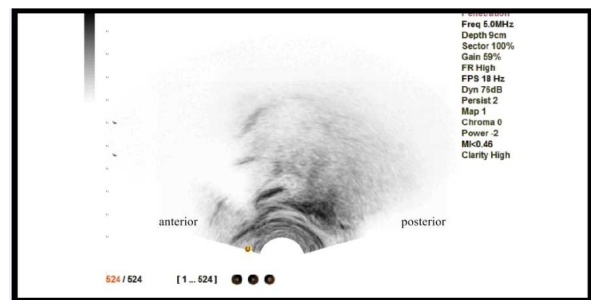
[a]



[ε]



[e]



[i]



[ɔ]



[o]



[u]

Figura 25 – Imagens de ultrassonografia (abordagem 1) geradas em padrão negativo das vogais orais: vogal [a]; vogal [ε]; vogal [e]; vogal [i]; vogal [ɔ]; vogal [o]; vogal [u]

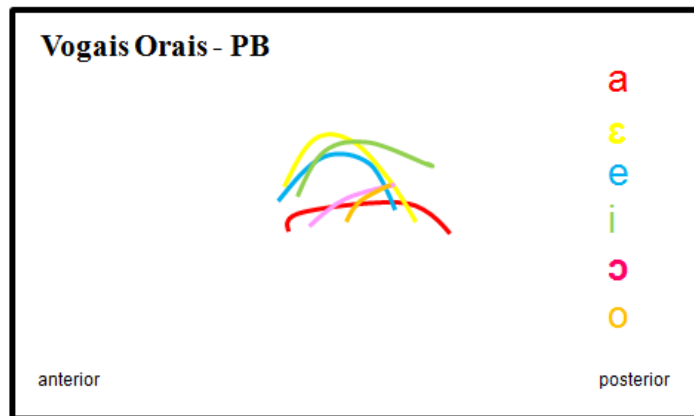


Figura 26 – Imagem gerada em sobreposição do traçado que acompanha o contorno da língua: vogal [a] – linha vermelha; vogal [ε] – linha amarela; vogal [e] – linha azul; vogal [i] – linha verde; vogal [ʊ] linha rosa; vogal [o] linha laranja (abordagem 1)

### 3. Análise integrada de dados acústicos e de imagens de USG

A análise discriminante do poder segregatório das informações de deslocamento do dorso de língua (na abordagem 1) em predizer as vogais orais avaliadas revelou poder segregatório parcial do total das amostras (em 52,94%). Para cada vogal, o poder segregatório variou de 0% (vogais [e] e [ε]) a 100% (vogais [i] e [o]) conforme Figura 27. A avaliação da vogal [u] não foi realizada por não ter sido possível gerar o traçado do contorno de língua.

O fator que influenciou tal nível de segregação das vogais (Figura 28) foi a posição vertical da língua com 79,1% de influência.

de \ a	[a]	[e]	[i]	[o]	[ε]	[ʊ]	Total	% correto
[a]	2	0	0	0	0	1	3	66,67%
[e]	1	0	2	0	0	0	3	0,00%
[i]	0	0	2	0	0	0	2	100,00%
[o]	0	0	0	3	0	0	3	100,00%
[ε]	0	0	2	0	0	1	3	0,00%
[ʊ]	1	0	0	0	0	2	3	66,67%
Total	4	0	6	3	0	4	17	52,94%

Figura 27 - Matriz de confusão para os resultados de validação cruzada

da análise discriminante (das informações de ultrassonografia para as medidas acústicas F1, F2 e F3)  
das sete vogais orais do português brasileiro

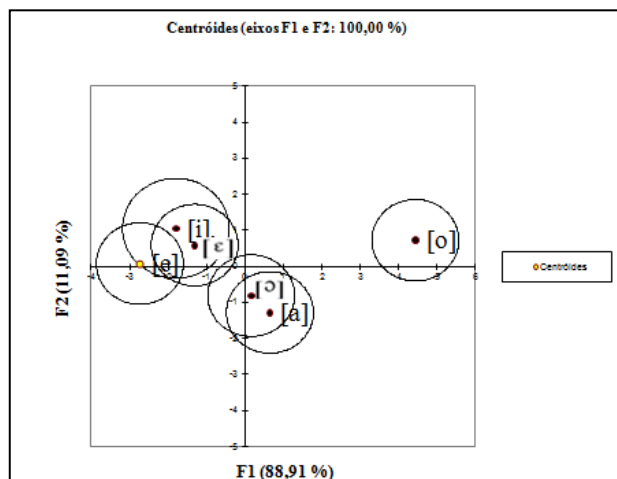


Figura 28 – Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de F1, F2 e F3) para identificação das vogais

Os dados de análise de correlação canônica para informações de deslocamento do dorso de língua (na abordagem 1) e medidas acústicas de formantes F1, F2 e F3 revelaram as seguintes correlações: medidas de F2 e deslocamentos verticais (75,4%) e horizontal (59,9%) da língua.

A análise discriminante do poder segregatório das informações de deslocamento do dorso de língua (na abordagem 2) em prever as vogais orais avaliadas revelou poder segregatório parcial do total das amostras em 29,41%. Para cada vogal, o poder segregatório variou de 0% (vogais [a] e [o]) a 100% (vogal [i]), conforme Figura 29. A avaliação da vogal [u] não foi realizada por não ter sido possível gerar o traçado do contorno de língua.

O fator que influenciou tal nível de segregação das vogais (Figura 30) foi a posição vertical da língua com 66,5% de influência, aliada à posição horizontal (posterior) em 61,2%.

de \ a	[a]	[e]	[i]	[o]	[ε]	[ɔ]	Total	% correto
[a]	0	0	0	2	0	1	3	0,00%
[e]	0	2	0	0	1	0	3	66,67%
[i]	0	0	2	0	0	0	2	100,00%
[o]	2	0	0	0	1	0	3	0,00%
[ε]	0	1	0	1	1	0	3	33,33%
[ɔ]	1	0	1	1	0	0	3	0,00%
Total	3	3	3	4	3	1	17	29,41%

Figura 29 - Matriz de confusão para os resultados de validação cruzada da análise discriminante (medidas de deslocamento da língua e medidas acústicas (F1, F2 e F3) para as vogais orais

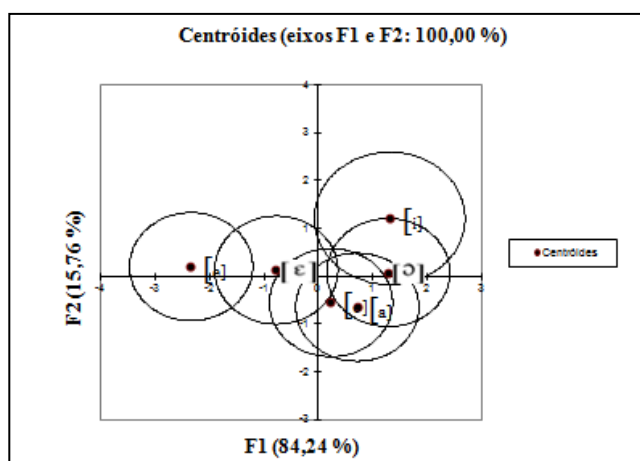


Figura 30 – Gráfico das centróides gerado pela análise discriminante (valor segregatório das medidas de deslocamento do dorso de língua para as medidas de F1, F2 e F3) das vogais orais

Os dados de análise de correlação canônica para informações de deslocamento do dorso de língua (na abordagem 1) e medidas acústicas de formantes F1, F2 e F3 revelaram as seguintes correlações: medidas de F2 e deslocamento vertical (67,6%) e horizontal – anterior (63,4%) da língua, além de medidas de F3 e deslocamento horizontal – posterior (39,8%).



## 5. DISCUSSÃO

Os recursos disponibilizados nos campos da Fonética Acústica e da Fonética Articulatória possibilitam a descrição e a compreensão integradas dos eventos de fala.

Discutimos aqui as medidas acústicas das frequências formânticas de falantes sem alteração de fala, além das informações de imagens de USG, especialmente do traçado do dorso de língua a partir de um estudo de caso. Procuramos, portanto, descrever, de forma preliminar, as possíveis correlações entre achados acústicos e articulatórios por imagens de USG.

A análise acústica fornece a possibilidade de traçar inferências a respeito do movimento realizado pelos articuladores na produção do sinal de fala, o qual influencia a configuração dos espaços das cavidades supraglóticas no caso das vogais. Por sua vez, os estudos com imagens do trato vocal supraglótico têm revelado detalhes importantes da atividade dos articuladores durante a produção dos sons, complementando os dados articulatórios inferidos da análise acústica.

A descrição acústica das vogais do PB contempla estudos de diversas naturezas. As vogais tônicas dos estados brasileiros do Rio Grande do Sul (MEIRELLES, 2008) e do Amazonas (CAMPOS, 2010), a caracterização acústica da produção e percepção das vogais e consoantes por DA (MENDES, 2003; FICKER, 2003), a descrição acústica de crianças respiradoras orais e respiradoras nasais (OLIVEIRA, 2011).

No âmbito dos estudos articulatórios destacam-se as técnicas que são desenvolvidas no campo das Tecnologias de Fala. A xerorradiografia (tipo de raio-X) permite excelente qualidade de imagem. No entanto, o sujeito submetido à técnica é exposto à grande quantidade de radiação e existe uma limitação quanto ao *corpus*, concentrado em uma imagem de um momento de emissões sustentadas (PINHO *et al.*, 1988; MASTER, 1991).

A IRM, por sua vez, possibilita a descrição da configuração do trato vocal supraglótico, especialmente na produção de vogais orais e nasais do PB (GREGIO, 2006). Apesar do registro dinâmico da fala e da imagem captada incluir estruturas e cavidades, a limitação da técnica reside no posicionamento do falante (deitado ao

longo da captação das imagens), o que interfere na postura habitual dos articuladores durante a produção de fala.

A USG para captura de imagens do trato vocal revelou-se como uma técnica não invasiva e propícia para a avaliação de fala articulada nas últimas décadas (KELSEY *et al*, 1969; FINGER *et al*, 2004; BERTI, 2010).

Pela escassez de estudos voltados à extração de dados acústicos que descrevam a produção de vogais do PB em população sem alterações de fala, analisamos acusticamente um grupo de 16 falantes dos gêneros masculino e feminino.

O parâmetro acústico analisado – frequências formânticas – imprime a identidade fonética das vogais na maioria das línguas (FANT, 1960). De acordo com o mesmo autor, as frequências dos dois primeiros formantes (F1 e F2) são responsáveis, respectivamente, pelo movimento da língua no eixo vertical (e de mandíbula) e no eixo horizontal.

A análise discriminante aplicada aos dados acústicos revelou informações que merecem destaque em nossa discussão. Os dados acústicos encontrados apontam que os valores de F1 apresentam a maior influência na segregação das vogais orais do gênero masculino (correlação de 97,3%) e o F2 para o gênero feminino (correlação de 91,9%).

Os valores das frequências formânticas contribuíram para discriminar os gêneros, ou seja, do total de emissões, 73,74% foram identificadas como sendo do gênero masculino, e 77,59% do gênero feminino. Nesse sentido, estes dados colaboram para a distinção de falantes de diferentes gêneros. Os valores de F3 definem-o como o principal formante para a diferenciação dos gêneros, o que podemos relacionar às diferentes configurações estabelecidas pelo trato vocal, especialmente ao espaço da cavidade oral durante a produção dos sons de fala (KENT, READ, 1992).

Para os falantes do gênero masculino, as medidas de frequências formânticas (F1, F2 e F3) permitiram diferenciar as vogais orais em 78,06% das emissões registradas. Quando considerado o poder segregatório das medidas acústicas para detecção de cada vogal oral (Figuras 17 e 19), as vogais posteriores [ɔ], [o] e [u] revelaram os menores índices, ao contrário das vogais [a], [ɛ], [e] e [i] com melhores índices de discriminação.

No caso do gênero feminino, as medidas de frequências formânticas (F1, F2 e F3) permitiram diferenciar as vogais orais em 82,66% das emissões registradas. Quando considerado o poder segregatório das medidas acústicas para detecção de cada vogal oral (Figuras 18 e 20), novamente as vogais posteriores [ɔ], [o] e [u] revelaram os menores índices, ao contrário das vogais [a], [ɛ], [e] e [i].

Os dados acústicos de frequências formânticas em uma população sem alterações e/ou queixas de fala apontam para a demanda de ampliação de pesquisas que buscam padrões de referência na produção de vogais orais do PB (nesse caso, variante paulistana com a utilização de *software* de ampla aplicação na comunidade científica - *Praat*). Além disso, facilitam a prática clínica nos distúrbios da comunicação, no sentido de investigações de fala com alterações e o trabalho para a evolução terapêutica, além de possibilidade de comparação de características de falantes no campo da Sociofonética (SCOBIE, *et al*, 2008; THOMAS, 2011) e do ensino do português como língua estrangeira.

Destaca-se ainda, que a ampliação de estudos dessa natureza em vogais nasais, também possibilitará novas contribuições teóricas e clínicas, especialmente se conjugada à análise articulatória que enfoque o segmento velofaríngeo.

A necessidade de correlacionar os dados acústicos aos ajustes articulatórios (fisiológicos) durante a produção de fala levantou a possibilidade do uso da técnica de USG para a descrição desses eventos.

No entanto, a falta de tradição da incorporação do instrumental de USG para a coleta específica de dados de fala em nosso meio revelou demandas e adaptações necessárias para a obtenção de resultados articulatórios.

Em consequência, o trabalho específico desta pesquisa focou-se no desenvolvimento de aparatos que permitissem o registro dos eventos articulatórios em tempo real. Dessa forma, o instrumental desenvolvido para a presente pesquisa, especialmente o suporte do transdutor, possibilitou a coleta de dados e uma investigação preliminar das particularidades das vogais orais, por meio das imagens de USG, bem como a reflexão sobre a necessidade de avanço no sentido de adaptação de instrumentos capazes de permitir a coleta de fala em um número ampliado de falantes.

O estudo de caso aqui apresentado revela as várias demandas e os desafios impostos aos pesquisadores que se aventuraram nos domínios da investigação em

fonética articulatória, e a dificuldade em se trabalhar com instrumentos ainda não adaptados para a realidade da coleta específica de fala. Este segmento do trabalho visa agregar vários profissionais para esta realidade e, ao mesmo tempo, defender a sua viabilidade e aplicabilidade.

As imagens de USG aqui apresentadas foram extraídas da gravação de um vídeo (com as sentenças-veículo do *corpus* apresentado). Ou seja, as imagens foram retiradas de uma sequência dinâmica de fala, o que é limitador em outras técnicas utilizadas em pesquisas anteriores que tiveram o mérito de considerar a análise articulatória como fator importante de descrição das vogais do PB (PINHO, 1988; MASTER, 1991; GREGIO, 2006).

Os dados de imagens de USG apresentaram, em corte sagital, o contorno da língua nos eixos vertical e horizontal, descritos na caracterização articulatória da produção das vogais, dados estes que são conjugados às informações de arredondamento labial, abertura/fechamento da mandíbula e segmento velofaríngeo, não propiciados pela técnica de USG (GREGIO, 2006; CAMARGO *et al*, 2008).

As correspondências entre dados acústicos – frequências formânticas – e dados articulatórios – imagens de USG do trato vocal supraglótico obtidas neste estudo, em uma população sem alterações de fala, foram evidentes para o deslocamento vertical da língua. A abordagem 1 favoreceu a identificação de vogais posteriores, enquanto as vogais anteriores foram melhor definidas pela abordagem 2, cabendo-se salientar que a vogal [i] foi corretamente identificada pelas duas abordagens.

Tais dados reforçam o potencial do uso da análise conjugada de dados acústicos e por imagens de USG e reforçam a demanda por aprimoramento das técnicas de análise das imagens, com planejamento específico para a avaliação de fala e correlação aos dados acústicos.

Destaca-se que investigações semelhantes de parâmetros acústicos que forneçam informações que possam caracterizar diferentes populações em outras faixas etárias, como crianças e idosos, são de muita importância para a descrição dos eventos de fala ao longo do crescimento e envelhecimento dos falantes (ANDRADE DAN, 2009; ANDRADE FLA, 2009; OLIVEIRA, 2011). Além disso, as características encontradas nessas populações podem servir como referência nos casos de alterações de fala e na consequente conduta clínica.

Quanto aos dados de imagem de USG, acreditamos que essa investigação possa estimular o aprimoramento da técnica em nosso meio, especialmente no que se refere ao aparato instrumental, para o desenvolvimento de recursos e configurações específicas para a coleta de dados de fala, além de eficientes mecanismos de análise das imagens coletadas.

Portanto, a descrição e a integração de achados acústicos e articulatórios podem revelar parte de um processo dinâmico e refinado que é a produção de fala.

## **6. CONCLUSÃO**

Do ponto de vista acústico, os valores das frequências formânticas para as sete vogais orais do PB para grupos de falantes adultos dos gêneros masculino e feminino foram caracterizados em termos da média, DP e IC dos valores de F1, F2 e F3.

A investigação da produção das vogais orais por imagens de USG revelou a característica articulatória do deslocamento do dorso de língua nos eixos vertical e horizontal (ântero-posterior). A técnica mostra-se promissora para a descrição de eventos de fala, do ponto de vista articulatório, e pode se estender como método complementar no diagnóstico de alterações de fala e ao ensino de fonética.

Os dados preliminares da correlação entre os achados acústicos e as imagens de USG revelaram e reforçaram a potencial aplicação no estudo das correlações acústico-articulatórias de vogais.

## 7. Anexos

### Anexo 1



**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO**  
**COMITÊ DE ÉTICA EM PESQUISA DA PUC-SP**  
**SEDE CAMPUS MONTE ALEGRE**

Protocolo de Pesquisa nº 224/2011

Faculdade de Filosofia, Comunicação, Letras e Artes  
Programa de Estudos Pós-Graduados em Linguística Aplicada: Estudo da Linguagem  
Orientador(a): Prof.(a). Dr.(a). Zuleica Antonia Camargo  
Autor(a): Maria Augusta Fone Svícero

**PARECER** sobre o Protocolo de Pesquisa, em nível de Dissertação de Mestrado, intitulado *Dados acústicos e de ultrassonografia da produção das vogais orais do português brasileiro*

**CONSIDERAÇÕES APROVADAS EM COLEGIADO**

Em conformidade com os dispositivos da Resolução nº 196 de 10 de outubro de 1996 e demais resoluções do Conselho Nacional de Saúde (CNS) do Ministério da Saúde (MS), em que os critérios da relevância social, da relação custo/benefício e da autonomia dos sujeitos da pesquisa pesquisados foram preenchidos.

O Termo de Consentimento Livre e Esclarecido permite ao sujeito compreender o significado, o alcance e os limites de sua participação nesta pesquisa.

A exposição do Projeto é clara e objetiva, feita de maneira concisa e fundamentada, permitindo concluir que o trabalho tem uma linha metodológica bem definida, na base do qual será possível retirar conclusões consistentes e, portanto, válidas.

No entendimento do CEP da PUC-SP, o Projeto em questão não apresenta qualquer risco ou dano ao ser humano do ponto de vista ético.

**CONCLUSÃO**

Faça ao parecer consubstanciado apensado ao Protocolo de Pesquisa, o Comitê de Ética em Pesquisa da Pontifícia Universidade Católica de São Paulo – PUC/SP – Sede Campus Monte Alegre, em Reunião Ordinária de **12/09/2011**, **APROVOU** o Protocolo de Pesquisa nº **224/2011**.

Cabe ao(s) pesquisador(es) elaborar e apresentar ao CEP da PUC-SP – Sede Campus Monte Alegre, os relatórios parcial e final sobre a pesquisa, conforme disposto na Resolução nº 196 de 10 de outubro de 1996, inciso IX.2, alínea "c", do Conselho Nacional de Saúde (CNS) do Ministério da Saúde (MS), bem como cumprir integralmente os comandos do referido texto legal e demais resoluções do Conselho Nacional de Saúde (CNS) do Ministério da Saúde (MS).

São Paulo, 12 de setembro de 2011.

  
Prof. Dr. Edgard de Assis Carvalho  
Coordenador do Comitê de Ética em Pesquisa da PUC-SP

## QUESTIONÁRIO

**Nome:**

**Idade:**

**Escolaridade:**

**Data:**

1. Você sente alguma dificuldade ao falar/produzir um som específico como uma consoante, por exemplo?
2. Já realizou alguma cirurgia/intervenção nas regiões da cabeça, pescoço e/ou face?
3. Apresenta alguma queixa em relação a audição?



**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

Nome do participante: \_\_\_\_\_ D/N: \_\_\_\_/\_\_\_\_/\_\_\_\_

Endereço: \_\_\_\_\_

Cidade: \_\_\_\_\_ Estado: \_\_\_\_ CEP: \_\_\_\_\_ Telefone: (\_\_\_\_) \_\_\_\_\_

RG: \_\_\_\_\_ CPF: \_\_\_\_\_

Nome da pesquisadora principal: Maria Augusta Forte Svicero

Instituições envolvidas: Laboratório Integrado de Análise Acústica e Cognição da Pontifícia Universidade Católica de São Paulo (LIAAC-PUCSP)

1. *Título do estudo:* Dados acústicos e de ultrassonografia da produção das vogais orais do Português Brasileiro.
2. *Propósito do estudo:* caracterizar acusticamente (frequências formânticas) as vogais orais do português brasileiro, variante paulistana, e apresentar dados preliminares da investigação de vogais por ultrassonografia, apontando possíveis correlações com a esfera acústica.
3. *Procedimentos:* Participarei de uma gravação em áudio e coleta de dados de ultrassonografia de amostra de fala composta da apresentação de sentenças veículos. Serei previamente orientado pelo pesquisador que efetuará a coleta.
4. *Riscos e desconfortos:* Nenhum.
5. *Benefícios:* A minha participação é voluntária e não trará qualquer benefício direto, mas proporcionará um melhor conhecimento a respeito
6. *Direitos do participante:* Eu posso retirar o(a) minha participação deste estudo a qualquer momento, sem sofrer nenhum prejuízo e tenho direito de acesso, em qualquer etapa do estudo, sobre qualquer esclarecimento de eventuais dúvidas.
7. *Compensação financeira:* Não existirão despesas ou compensações financeiras relacionadas à participação do(a) meu(minha) filho(a) em qualquer etapa do estudo, incluindo exames e consultas, exceto reembolso dos gastos com transporte até o local do estudo.
8. *Incorporação ao banco de dados do LIAAC-PUCSP:* Os dados obtidos com minha participação, serão incorporados ao banco de dados do laboratório referido, cujos responsáveis zelarão pelo uso e aplicabilidade das amostras exclusivamente para fins científicos, apenas consentindo o seu uso futuro em projetos que atestem pelo cumprimento dos preceitos éticos em pesquisas envolvendo seres humanos.
9. *Em caso de dúvida quanto ao item 8, posso entrar em contato com os responsáveis pelo banco de dados do LIAAC (Profa. Dra. Zuleica Antônia de Camargo e Prof. Mário Fontes) no telefone: (11)3670-8333.*
10. *Confidencialidade:* Compreendo que os resultados deste estudo poderão ser publicados em jornais profissionais ou apresentados em congressos profissionais, sem que a identidade do(a) meu(minha) filho(a) seja revelada.
11. *Se tiver dúvidas quanto à pesquisa descrita posso telefonar para a pesquisadora Maria Augusta Forte Svicero no número (11)9575-9613 a qualquer momento.*

Eu compreendo os meus direitos como um sujeito de pesquisa e voluntariamente consinto em autorizar a participação dele(a) neste estudo e em ceder os dados dele(a) para o banco de dados do LIAAC-PUCSP.

Compreendo sobre o que, como e porque este estudo está sendo feito. Receberei uma cópia assinada deste formulário de consentimento.

\_\_\_\_\_  
Assinatura do sujeito participante

\_\_\_\_\_  
Assinatura do pesquisador

São Paulo, \_\_\_\_\_

Data

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE DAN, V. *Análise de parâmetros espectrais clássicos da voz em crianças respiradoras orais de 5 a 10 anos*. In: Mestrado em Fonoaudiologia, Faculdade de Fonoaudiologia Universidade Veiga de Almeida. Rio de Janeiro: 2009.

ANDRADE FLA, V. *Análise de parâmetros espectrais da voz em crianças saudáveis de 4 a 8 anos*. In: Mestrado em Fonoaudiologia, Faculdade de Fonoaudiologia Universidade Veiga de Almeida. Rio de Janeiro: 2009.

BAUERMAN, G. Ultrassonografia. Disponível em: <[www.imagesurvey.com.br/2009/03/ultrassonografia/](http://www.imagesurvey.com.br/2009/03/ultrassonografia/)>. Acesso em 08 de maio de 2009.

BECK, J. M. The vocal tract as a resonator. In: HEWLETT, N.; JANET, J. M. In: *Introduction to the science of phonetics*. Lawrence Erlbaum Associates, Inc., 2006. p. 239-306.

BERTI, L. C. *Investigação da produção de fala a partir da ultrassonografia do movimento de língua*. In: 18º Congresso Brasileiro de Fonoaudiologia: Exercício Profissional: Bases Teóricas, Avanço e Realidade Nacional, Sociedade Brasileira de Fonoaudiologia, Curitiba: 2010. v. 1. p. 661-1.

BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer*. In: <[www.fon.hum.uva.nl/praat/](http://www.fon.hum.uva.nl/praat/)>. Amsterdam: 2002.

CAMARGO, Z. A; FONTES, M. A. S.; MADUREIRA, S. *Introdução ao estudo dos sons da fala*. [Apostila da disciplina de Fonética e Fonologia I do curso de Fonoaudiologia da Pontifícia Universidade Católica de São Paulo]. São Paulo, 2008.

CAMPOS, M. S. *O Alçamento das Vogais Posteriores Tônicas Do Português Falado No Amazonas*. In: 1º CIELLI – Colóquio Internacional de Estudos Lingüísticos e Literários, 4º CIELLI – Colóquio de Estudos Lingüísticos e Literários. Universidade Estadual de Maringá – UEM. Maringá: 2010.

CUKIER, S. *Qualidade vocal em indivíduos asmáticos com e sem disfunção paradoxal de pregas vocais: correlatos perceptivo-auditivos, acústicos e fisiológicos*. In: Mestrado em Linguística Aplicada e Estudos da Linguagem, PUC-SP. São Paulo: 2006.

DEMOLIN, D.; METENS, T.; SOQUET, A. *Real time MRI and articulatory coordinations in vowels*. Convention ARC “Dynamique des système phonologiques” n.226, 2006.

DEMOLIN, D.; MEDEIROS, B. R. *Vogais nasais do português brasileiro: um estudo de IRM*. Revista da ABRALIN, 2006. v.5, n.1 e 2, p. 131-42.

FANT, G. *Acoustic theory of speech production*. Mouton The Hague: 1960.

- FICKER, L. B. *Produção e percepção das plosivas do português brasileiro: estudo fonético-acústico da fala de um sujeito com deficiência auditiva*. In: Doutorado em Linguística Aplicada e Estudos da Linguagem, PUC-SP. São Paulo: 2003.
- FINGËR, T. C.; SILVA, A.; ZIMMERMANN, R. F.; JÚNIOR, A. T. *Avaliação ultrassonográfica do tamanho da língua em adultos*. Rev CEFAC, São Paulo: 2004. v.6, n.3, p. 272-6, jul-set.
- GREGIO, F. N. *Configuração do Trato Vocal supraglótico na produção das vogais do Português Brasileiro: dados de imagens de ressonância*. In: Mestrado em Linguística Aplicada e Estudos da Linguagem, PUC-SP. São Paulo: 2006.
- GREGIO, F. N.; GAMA-ROSSI, A.; MADUREIRA, S.; CAMARGO Z. *Modelos teóricos de produção e percepção da fala como um modelo dinâmico*. Rev. CEFAC: São Paulo, 2006. v. 8, n. 2, p. 244-247, abr-jun.
- JOHNSON, K. *Acoustic and Auditory Phonetics*. Blackwell Publishers. Malden: 2000.
- KELSEY, C. A.; MINIFIE, F. D.; HIXON, T. J. *Applications of Ultrasound in Speech Research*. J Speech Hear Res, 1969. v. 12, n. 3, p.564-575.
- KELSO, J. A. S. *Dynamic Patterns: the self-organization of brain and behavior*. Massachusetts Institute of Technology. Massachusetts: 1995.
- KENT, R. D.; READ, C. *The acoustic analysis of speech*. Singular Publishing Group, Inc. Califórnia: 1992.
- KOCJANCIC, T. *Tongue movements and syllable onset complexity: ultrasound study*. Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics, 28-30 August 2006, Athens, Greece: 2008.
- LADEFOGED, P; HARSHMAN, R. *Formant frequencies and movements of the tongue*. In: Frontiers of Speech Communication Research. Academic Press Inc, London: 1979.
- LIEBERMAN, P.; BLUMSTEIN, S. E. *Speech Physiology, speech perception and acoustic phonetics*. Cambridge University Press. Cambridge: 1988.
- MADUREIRA, S. Expressividade da fala. In: KYRILLOS, L. (org). *Expressividade – da teoria à prática*. Editora Revinter. Rio de Janeiro: 2004. p. 15-25.
- MAGRI, A.; STAMADO, T.; CAMARGO, Z. A. *Influência da largura de banda de formantes na qualidade vocal*. Rev. CEFAC, São Paulo, 2009. v.11, n. 2, jun.
- MASTER, S.; PONTES, P. A. L.; BEHLAU, M. S. *Configurações do trato vocal nas vogais nasais do português brasileiro*. Acta AWHO, 1991. v. 10, n. 2, p. 67-75.

MEIRELLES, V. A. G. *Estudo acústico das vogais tônicas em palavras paroxítonas do português falado no Rio Grande do Sul*. Revista de Letras da Universidade Católica de Brasília. Brasília: 2008. v.1, n.2, nov.

MENDES, B. C. A. *Estudo Fonético acústico das vogais do português brasileiro: dados da produção e percepção de fala de um sujeito deficiente auditivo*. In: Doutorado em Linguística Aplicada e Estudos da Linguagem, PUC-SP. São Paulo: 2003.

OLIVEIRA, L. R. *Análise acústica comparativa das vogais orais entre respiradores orais e nasais*. In: Mestrado em Linguística Aplicada e Estudos da Linguagem, PUC-SP. São Paulo: 2011.

PEREIRA, A. H. A. *Cerâmicas piezoelétricas: funcionamento e propriedades*. Application note RT-ATCP-01, 2010.

PINHO, S. M. R; PONTES, P. A. L; ABÍLIO, S. O; GANANÇA, M. M. *Configuração do trato vocal nas vogais orais do português*. Acta AWHO, 1988. v. 7, n.2, p. 124-36.

SANTOS, M. H. G. *Desenvolvimento de transdutores piezoelétricos de ultrassom para formação de imagens*. In: Mestrado na Escola Politécnica da Universidade de São Paulo, São Paulo: 2010.

SCOBIE, J. M.; STUART-SMITH, J.; LAWSON, E. *Looking variation and change in the mouth: developing the sociolinguistic potencial of Ultrasound Tongue Imaging*. Research Report for ESRC Project RES-000-22-2032, 2008.

STEVENS, K. N. *Acoustic Phonetics*. MIT Press. London: 2000.

THOMAS, E. R. *Sociophonetics: an introduction*. Basingstoke Palgrave Macmillan, 2011.

XU, Y. *In Defense of Lab Speech*. Journal of Phonetics. London: 2010. p. 329-336.

ZHARKOVA, N; HEWLETT, N; HARDCASTLE, W.J. *An Ultrasound Study of Lingual Coarticulation in Children and Adults*. Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics, 28-30 August 2006, Athens, Greece: 2008.