UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ESTUDOS DA LINGUAGEM

FRANCISCO DE OLIVEIRA MENESES

UMA VISÃO DINÂMICA DOS PROCESSOS DE APAGAMENTO DE VOGAIS NO PORTUGUÊS BRASILEIRO

CAMPINAS,
2016

FRANCISCO DE OLIVEIRA MENESES

UMA VISÃO DINÂMICA DOS PROCESSOS DE APAGAMENTO DE VOGAIS NO PORTUGUÊS BRASILEIRO

Tese de doutorado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para obtenção do título de Doutor em Linguística.

Orientador (a): Profa. Dra. Eleonora Cavalcante Albano

Este exemplar corresponde à versão final da Tese defendida pelo aluno Francisco de Oliveira Meneses e orientada pela Profa. Dra. Eleonora Cavalcante Albano.

CAMPINAS,
2016

Agência(s) de fomento e nº(s) de processo(s): CAPES

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Estudos da Linguagem
Crisllene Queiroz Custódio - CRB 8/8624

M524v

Meneses, Francisco, 1986-
Uma visão dinâmica dos processos de apagamento de vogais no português brasileiro / Francisco de Oliveira Meneses. – Campinas, SP : [s.n.], 2016.

Orientador: Eleonora Cavalcante Albano.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Língua portuguesa - Brasil. 2. Língua portuguesa - Vogais. 3. Língua portuguesa - Fonemas. 4. Gramática comparada e geral - Fonologia. 5. Sândi. I. Albano, Eleonora Cavalcante,1950-. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: A dynamic view of vowel deletion process in Brazilian portuguese
Palavras-chave em inglês:
Portuguese language - Brazil
Portuguese language - Vowels
Portuguese language - Phonemes
Grammar, Comparative and general - Phonology
Sandhi
Área de concentração: Linguística
Titulação: Doutor em Linguística
Banca examinadora:
Eleonora Cavalcante Albano [Orientador]
Didier Sheila Jean Marie Demolin
Ryan Keith Shosted
Larissa Cristina Berti
Wilmar da Rocha D'Angelis
Data de defesa: 26-08-2016
Programa de Pós-Graduação: Linguística

BANCA EXAMINADORA:

Eleonora Cavalcante Albano

Wilmar da Rocha D'Angelis

Didier Sheila Jean Marie Demolin

Larissa Cristina Berti

Ryan Keith Shosted

Vera Pacheco

Maria Filomena Spatti Sândalo

Antonio Carlos Silvano Pessotti

IEL/UNICAMP
2016

Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

É tudo por você, Morena...

Agradecimentos

À Profª. Drª. Eleonora Cavalcante Albano, que orientou meu trabalho nos últimos seis anos (mestrado e doutorado) e foi responsável por meu crescimento acadêmico. Agradeço por suas contribuições e paciência.

Também agradeço à Profª. Drª. Vera Pacheco, que deu início ao caminho que percorri até aqui. Agradeço imensamente, acima de tudo, o carinho, os conselhos e o apoio.

Aos membros da banca da tese, Prof. Dr. Ryan Shosted, Profª. Drª. Larissa Berti, Prof. Dr. Didier Demolin e Prof. Dr. Wilmar Rocha D'Angelis, agradeço a disposição e tempo. Em especial, agradeço ao Prof. Wilmar, que participou de todas minhas bancas e sempre teve leitura certeira em todas as ocasiões. Além dele, agradeço de coração o Prof. Ryan que, mais que um professor, é um grande amigo.

Aos sujeitos da pesquisa, agradeço a simpatia e a disposição que demonstraram durante as longas sessões de gravação.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

Aos demais professores e funcionários do Departamento de Linguística do IEL, que fizeram parte da minha formação durante os últimos anos.

Agradeço aos colegas do LAFAPE pela colaboração e companheirismo. Meu muito obrigado a todos que passaram pelo Laboratório, em especial, para Luciana Lessa, Denise Pozzani, Larissa Rinaldi, Antonio Pessoti, Mariana Hungria, Lucila Schieleman, Diego Jiqüilin Ramirez.

Agradeço a todos meus amigos que, direta ou indiretamente, fazem parte dessa conquista. Em especial, meu muito obrigado aos amigos Jamile (Kbça), Ana Amélia, Tati, Magnum, Melina, Vini . Ainda agradeço a Marcos Pires que fez importantes considerações na revisão dessa tese.

A meus pais, Norma e João, e a minha irmã, Bam, pelo incentivo e amor. Agradeço também aos meus avós, tios, primos e sogros pelo apoio de toda vida.

Por fim, meu agradecimento e todo meu amor incondicional a minha esposa, Juliana. Ela é minha mulher, amiga e companheira de todas as horas. Muito obrigado pela cumplicidade e pela paciência nas noites mal dormidas. É uma honra amar você.

RESUMO

A presente tese de doutoramento investiga dois processos de redução do Português Brasileiro: o desvozeamento das vogais altas finais e a elisão no sândi vocálico externo. Sobre o primeiro, sugere-se que as vogais átonas finais tendem a ser apagadas em posições fracas. Meneses (em preparação) e Meneses e Albano (2015) argumentam que não se trata de apócope, mas de um desvozeamento da vogal. Para o sândi, uma extensa bibliografia sugere um apagamento categórico da vogal átona em junção de palavras, enquanto Albano et al. (1998) sugere que a vogal ainda deixa pistas no sinal acústico na fala normal. Ambos os fenômenos estão relacionados ao acento e a velocidade de fala que alteraram toda a configuração dos articuladores. Os resultados mostram que as vogais altas prevalecem sobre as vogais baixas em contexto de sândi. No desvozeamento, as vogais altas perdem sua magnitude e são completamente sobrepostas pelo ruído do /s/. No contexto do desvozeamento, a velocidade de movimentação dos gestos supraglóticos parece inibir o vozeamento, ou seja, não há tempo suficiente para retomar o vozeamento da vogal em ambiente desvozeado. No sândi vocálico externo, é a velocidade de transição entre /a#i/ que impede a movimentação regular da mandíbula, inibindo a produção do /a/ na fala rápida, mas a vogal é sustentada na taxa normal. Diante dos resultados, é possível afirmar que ambos os processos parecem ser resultado de questões específicas de *timing* e *target* articulatório, sendo que, no desvozeamento, há uma alteração radical da aerodinâmica supraglótica e glótica. A sobreposição entre a consoante e a vogal no desvozeamento e no encontro vocálico no sândi geram “erros” de percepção que podem, eventualmente, ser a matéria-prima para a mudança sonora.

Palavras-chave: Redução vocálica, Desvozeamento, Sândi, Elisão, Mudança Sonora.

ABSTRACT

This thesis investigates two Brazilian Portuguese reduction processes: final vowel devoicing and e vocalic elision in sandhi. As to devoicing, it is suggested that the final post-stressed vowels tend to be deleted in weak positions. Meneses & Albano (2015) and Meneses (in preparation) have argued that this is not apocope, but an ongoing devoicing process. As to sandhi, an extensive bibliography suggests a categorical deletion of unstressed vowel in joining words, while Albano et al. (1998) and Meneses (in preparation) suggest that the vowel leaves clues in the acoustic signal. Both phenomena are related to the stress and the speech rate that changed the whole configuration of the articulators, and also related to speech planning and execution. The results show that the high vowels prevail over the low vowels in sandhi context, as they prevalent in devoicing, even without the voicing default. In devoicing, the velocity of the supraglottic gestures movement appears to inhibit the voicing, i.e., there is not enough time to resume vowel voicing in devoiced environment. In the external vocalic sandhi, the velocity of the transition between /a#i/ that prevents the regular jaw movement by inhibiting the production of the vowel /a/ in fast speech, but the vowel is produced in the normal speech rate. Given the results, we can say that both processes appear to be the result of specific issues of timing and articulatory target. As to devoicing, there is a radical change in the supraglottic and glottic aerodynamic. The overlap between consonant and vowel in devoicing and the combination of two adjacent vowel in generate perception "errors" that can possibly be the raw material for the sound change.

Keywords: Vowel Reduction, Devoicing, Sandhi, Elision, Sound Change

Lista de Figuras, Quadros e Tabelas

Figura 1 – Exemplos de sílabas com vogais vozeadas, parcialmente desvozeadas e totalmente desvozeadas, respectivamente, extraídas da palavra "lance" /laNsi/ 36
Figura 2 – Área de espaço vocálico de vogais seguidas de consoantes desvozeadas (A) e vozeadas (B) 37
Figura 3 – Diagrama de dispersão entre o VSA dos sujeitos e a duração das vogais em contexto desvozeado (R (24) = 0,53, *p* = 0,005) 38
Figura 4 – Ocorrência de vogais desvozeadas, parcialmente desvozeadas e vozeadas para todos os sujeitos analisados em Meneses (2012) 39
Figura 5 – Centroide da fricativa em sílabas "sem vogal" (/i/ desvozeado) e com vogal (/i/ vozeado) 40
Figura 6 – Duração do ruído do [s] com vogal vozeada, parcialmente desvozeada e totalmente desvozeada 41
Figura 7 – Duração da sílaba com vogal vozeada, parcialmente desvozeada e totalmente desvozeada 42
Figura 8 – Diagrama de dispersão entre a duração da vogal e da consoante normalizada pelo *z-score* 43
Figura 9 – Esboço de representação das hipóteses levantadas a partir dos resultados do Experimento 1 46
Figura 10 – Imagem do EVA2 e seus canais de captação 50
Figura 11 – Forma de onda (acima) e fluxo de ar oral (abaixo) de /s/ com vogal vozeada em "passe" (à esquerda) e /s/ de coda em "paz" (à direita) 51
Figura 12 – Forma de onda (acima) e fluxo de ar oral (abaixo) de /s/ com vogal desvozeada 52
Figura 13 – Esboço de representação das hipóteses e do fluxo de ar obtidos no Experimento 2 53
Figura 14 – Escala utilizada pelos informantes para o experimento-piloto 57
Figura 15 – Porcentagem de cada escolha em um teste de identificação forçada para vogais desvozeadas e parcialmente desvozeadas 58
Figura 16 – Porcentagem de cada escolha dos sujeitos na identificação das vogais desvozeadas em cada ponto da escala 59
Figura 17 – Diagrama de dispersão entre as médias das escolhas e o tempo de reação (R (58) = -0,64, p < 0,0001) 60
Figura 18 – Porcentagem de acertos e erros para a identificação de /s/ alongado, /s/ de duração média e /s/ de coda 63
Figura 19 – Porcentagem de acertos e erros de identificação por sujeito 64
Figura 20 – Média do escore ponderado pela confiança do ouvinte para/s/ alongado, de duração média e de coda 65
Figura 21 – Variação da média dos escores ponderados por sujeito 66
Figura 22 – Porcentagem de escolha de cada vogal no Experimento III 67
Figura 23 – Exemplos de palavras com sílabas com vogais desvozeadas acompanhadas de oclusivas para tato [ˈtatʊ̥] e baque [ˈbakɪ̥], respectivamente 71
Figura 24 – *Boxplots* para a velocidade do sensor da posição medial da língua para sete sujeitos do PB (em mm/s) para /i/[1] (vermelho) e /i/[2] (azul) 73
Figura 25 – Plano cartesiano com os escores de /i[1]/ (vermelho) e /i[2]/ (azul) em PC1 (x) e PC2 (y) 74
Figura 26 – *Boxplots* dos escores de sete sujeitos PC2 ordenados pelo grau de

anteorização 75
Figura 27 – Plano cartesiano de /i/[1] e /i/[2] para PC1 (x) e PC2 (y) dos sujeitos S9 e S10 76
Figura 28 – Imagem sagital do quadro correspondente à fricativa /s/ no sinal acústico para "paz" 78
Figura 29 – Imagem sagital do quadro correspondente à sílaba /si/ com vogal vozeada (à esquerda) e desvozeada (à direita) no sinal acústico para a palavra "passe" 78
Figura 30 – Imagem sagital do quadro correspondente à sílaba /su/ com vogal vozeada (à esquerda) e desvozeada (à direita) no sinal acústico para a palavra "passo" 79
Figura 31 – Linha da variação sincrônica e da variação diacrônica hipotética 81
Figura 32 – Espectros de LPC após a liberação do [R] em três pronúncias (lenta, moderada e rápida) 90
Figura 33 – Ausência de africação em /muitˈidade/ pronunciada por um dos informantes do experimento de Albano *et al*. (1998) 91
Figura 34 – Pauta gestual proposta por Albano (1999) para [mult'idad͡ʒi] 92
Figura 35 – Linha de tendência da duração de /a#i/, /i/ e /a/ durante a realização da tarefa de repetição 95
Figura 36 – Linha não suavizada da duração de /a#i/, /i/ e /a/ (à esquerda) e linha de duração da palavra (à direita) 96
Figura 37 – Área de espaço vocálico das vogais /a/ e /i/ 97
Figura 38 – Tendência das frequências de F1 (à esquerda) e F2 (à direita) durante a tarefa de repetição 97
Figura 39 – Forma de onda e espectrograma de palavras retiradas de dois momentos diferentes da tarefa, normal e rápida 98
Figura 40 – Porcentagem de ocorrências de possíveis apagamentos e realização em casos de encontro de vogais na fala normal (à esquerda) e rápida (à direita) 102
Figura 41 – Duração da sequência /a+i/ e /i/ em condição de taxa normal de fala 103
Figura 42 – Duração da sequência /a+i/ e /i/ em condição de taxa rápida de fala 103
Figura 43 – Valores de F1 (à esquerda) e F2 (à direita) para /a/ e /i/ em sequências /a+i/ em condição de taxa normal de fala 104
Figura 44 – Valores de F1 (à esquerda) e F2 (à direita) para /a/ e /i/ em sequências /a+i/ em condição de taxa rápida de fala 104
Figura 45 – Representação de hipótese do encontro de vogais em fronteira para a fala normal 106
Figura 46 – Representação de hipótese do encontro de vogais em fronteira para a fala rápida 107
Figura 47 – Forma de onda (acima) e fluxo de ar oral (abaixo) de "raridade" (à esquerda) e "rara idade" (à direita) na fala normal 110
Figura 48 – Forma de onda (acima) e fluxo de ar oral (abaixo) de "raridade" (esquerda) e "rara idade" (direita) na fala rápida 111
Figura 49 – Forma de onda (acima) e fluxo de ar oral (abaixo) de "pielite" (à esquerda) e "pia elite" (à direita) na fala rápida 112
Figura 50 – Porcentagem de acerto e erro para os estímulos de taxa normal e rápida .. 115
Figura 51 – Imagem sagital do quadro correspondente à /i/ na fala de velocidade normal (à esquerda) e fala de velocidade rápida (à direita) 117
Figura 52 – Imagem sagital do quadro correspondente à sequência /a+i/ na fala de velocidade normal (à esquerda) e fala de velocidade rápida (à direita) 118

Quadro 1 – Palavras-alvo usadas no Experimento 1 34
Quadro 2 – Quadro-resumo das hipóteses relacionadas à H1 e H2 49
Quadro 3 – Tríades de palavras usadas no Experimento 2 50
Quadro 4 – Lista de palavras usadas no experimento-piloto 56
Quadro 5 – Tríades de palavras usadas no Experimento 3 62
Quadro 6 – Palavras-alvo usadas no Experimento de Produção 2 100
Quadro 7 – Palavras-alvo usadas no Experimento Aerodinâmico 109

Tabela 1 – Classificação dos resultados do LDA, mostrando a separação de /i/[1] e /i/[2] em duas classes 73
Tabela 2 – Porcentagem de acerto por palavra com velocidade de fala normal 115
Tabela 3 – Porcentagem de acerto por palavra com velocidade de fala rápida 116

Lista de Abreviaturas e Siglas

ABS	Valor Absoluto
C	Consoante
EMA	*Electromagnetic Articulography*
FAR	Fonologia Articulatória
FFT	Transformada rápida de Fourier
FonGest	Fonologia Gestual
IEL	Instituto de Estudos da Linguagem
LAFAPE	Laboratório de Acústica e Psicolinguística
LDA	Linear Discriminante Analisys
LPC	linear predictive coding
PB	português brasileiro
PCA	Análise de Componentes Principais
PE	português europeu
UIUC	University of Illinois at Urbana-Champaign
Unicamp	Universidade Estadual de Campinas
V	Vogal
VSA	Área de Espaço Vocálico

Sumário

INTRODUÇÃO GERAL 16
PARTE 1 – DA REDUÇÃO À APÓCOPE: DESVOZEAMENTO DE VOGAIS PÓS-TÔNICAS DO PB 21
CAPÍTULO 1 – REVISÃO DA LITERATURA 22
1.0 Introdução: redução e desvozeamento 22
1.1 Sobre redução 23
1.2. Sobre o desvozeamento 25
1.3 Evidências da diacronia e variação sincrônica de vogais no PB 27
1.4 Mudança sonora em andamento 29
1.5 Enquadramento teórico e termos basilares 30
CAPÍTULO 2 – UMA NOVA VISÃO DOS RESULTADOS DE MENESES (2012) 33
2.0 Introdução 33
2.1 Experimento I: breve revisão da metodologia usada em Meneses (2012) ... 33
2.1.1 Materiais e participantes 33
2.1.2 Medidas 34
2.1.3 Análise acústica 34
2.1.4 Análise estatística 35
2.2 Revisão e releitura dos dados 35
2.2.1 Achados gerais 45
2.2.2 Centralização vocálica extrema 36
2.2.3 Centroide 40
2.2.4 Dados de duração 41
2.3 Discussões 44
CAPÍTULO 3 – NOVAS PERSPECTIVAS SOBRE O DESVOZEAMENTO VOCÁLICO NO PB 46
3.0 Alongamento compensatório ou um resquício de vogal? 46
3.1 Experimento II: dados aerodinâmicos 49
3.1.1 Participantes 49
3.1.2 Materiais 50
3.1.3 Resultados 51
3.1.4 Discussão 52
CAPÍTULO 4 – SOBRE A PERCEPÇÃO 55
4.0 Teste de percepção 55
4.1 Teste-piloto 55
4.1.1 Participantes e materiais 56
4.1.2 Procedimento 56
4.1.3 Medidas 57
4.1.4 Análise estatística 57
4.1.5 Resultados 57
4.1.6 Discussão dos resultados (e falhas do experimento) 60
4.2 Experimento III: tarefa de identificação forçada 61
4.2.1 Participantes 61
4.2.2 Materiais 61

4.2.3 Procedimento 62
4.2.4 Análise estatística 62
4.2.5 Resultados 62
4.2.6 Discussão 67
4.3 Resumo geral dos achados e discussão 68
CAPÍTULO 5 – DIRETRIZES PARA PESQUISAS FUTURAS DO DESVOZEAMENTO 70
5.0 Breve introdução aos dados 70
5.1 Oclusivas 70
5.2 EMA 71
5.2.1 Metodologia 71
5.2.2 Resultados 72
5.3 Ultrassom 77
5.3.1 Visão geral do método 77
5.3.2 Resultados iniciais 77
5.4 Conclusão 80
PARTE 2 – A ELISÃO NO SÂNDI OU A REDUÇÃO DA VOGAL BAIXA 82
CAPÍTULO 6 – REVISITANDO O PROCESSO DE SÂNDI VOCÁLICO NO PB . 83
6.0 Introdução 83
6.1 O sândi nas línguas naturais 84
6.2 A influência do acento no PB e o processo de sândi vocálico 88
6.3 O sândi sob o viés da FonGest: elisão de /a/ átono em juntura com vogal átona na visão de Albano *et al*. (1998) e Albano (1999) 90
CAPÍTULO 7 – EXPERIMENTO DE PRODUÇÃO 93
7.0 Introdução: experimentos de produção 93
7.1 Experimento 1: tarefa de repetição contínua 93
7.1.1 Metodologia 93
7.1.1.1 Materiais e participantes 93
7.1.1.2 Procedimento 93
7.1.1.3 Segmentação e medidas 94
7.1.1.4 Análise estatística 94
7.1.2 Resultados 94
7.1.2.1 Duração 94
7.1.2.2 Formantes 96
7.1.3 Discussão 98
7.2 Experimento de produção II 100
7.2.1 Metodologia 100
7.2.1.1 Materiais, participantes e procedimento 100
7.2.1.2 Segmentação e medidas 101
7.2.1.3 Análise estatística 101
7.2.2 Resultados 101
7.2.2.1 Ocorrências 101
7.2.2.2 Duração 102
7.2.2.3 Formantes 103
7.2.3 Discussão 105
7.2.3.1 Hipóteses 106
CAPÍTULO 8 – TESTE AERODINÂMICO 108
8.0 Teste aerodinâmico: fluxo de ar oral 108

8.1 Teste aerodinâmico 108
8.1.1 Metodologia 108
8.1.1.1 Materiais, participantes e procedimento 108
8.1.2 Materiais 109
8.1.3 Resultados 109
8.1.4 Discussões 112
CAPÍTULO 9 – DIRETRIZES PARA PESQUISAS FUTURAS NO SÂNDI POR ELISÃO 114
9.0 Breve visão dos dados articulatórios 114
9.1 Percepção: teste de discriminação 114
9.1.1 Breve descrição da metodologia 114
9.1.2 Resultados 115
9.2 Dados articulatórios (prévia de Meneses *et al.*, em preparação) 116
9.3 Conclusão 118
CONSIDERAÇÕES FINAIS 119
REFERÊNCIAS 123
ANEXOS 132

Introdução Geral

"Language is to the mind more than light is to the eye"

William Gibson

Na fala espontânea, a produção de vogais e consoantes muitas vezes carece de algumas das pistas acústicas que nós, pesquisadores ou estudantes, estamos acostumados a encontrar na fala de laboratório. Por exemplo, é muito comum oclusivas sonoras intervocálicas não apresentarem o seu período característico de silêncio, explosão e vozeamento (Lewis, 2001; Warner, 2011), enquanto as vogais podem ser produzidas com *undershoot* de frequências formânticas (Koopmans-van Beinum, 1985; Monn & Lindblom, 1994 *inter alia*). Nesse contexto, vogais podem ser produzidas de forma parcial ou completamente desvozeadas sob a influência coarticulatória das consoantes surdas adjacentes e, também, influenciadas pela posição mais fraca na palavra (Davidson, 2006; Torreira & Ernestus, 2010 *inter alia*). Em casos extremos de redução de magnitude, segundo a literatura, os segmentos podem, de fato, ser acusticamente apagados (Browman & Goldstein, 1990).[1]

Em um *corpus* de fala espontânea do inglês americano, Johnson (2004) relata o apagamento de algum segmento em pelo menos 20% das palavras. Além disso, o autor reporta a ocorrência de algum tipo de redução (consonantal ou vocálica) em 60% das palavras. Embora a difusão de tais reduções na fala cotidiana seja, agora, amplamente reconhecida (como atestam os trabalhos, por exemplo, de Johnson, 1997; Ernestus, 2000; Kohler, 2000 *inter alia*), pouco ainda se sabe sobre sua exata natureza acústica, perceptual e aerodinâmica (sobre as duas últimas, as questões ainda são mais numerosas do que as respostas).

Na verdade, descrições detalhadas dos fenômenos de enfraquecimento na fala espontânea são escassas, especialmente fora do âmbito das línguas germânicas, como é o caso do português. Se tratarmos mais especificamente da redução da vogal, apesar de sua alta incidência na fala corrente do português brasileiro (doravante PB), o processo não tem

[1] Outro viés dos estudos sobre redução trata da redução de palavras. Por exemplo, a palavra "*ordinary*", em inglês, pode ser pronunciada como "*onry*", "*computer*" simplesmente como "*puter*" e "*yesterday*" como "*yeshay*". Para o português brasileiro, logo nos vem à cabeça expressões cristalizadas como "opaió", que é redução de "olha para issaí, olha", ou "pópôpó" para "pode pôr pó".

recebido a devida atenção nas literaturas fonético-fonológicas e, portanto, oferece um campo fértil de investigação. De fato, empiricamente, faltam estudos que relacionam a redução à mudança fônica, principalmente, as relacionadas ao suposto apagamento e eventual queda de vogais e abordagem dinâmicas sobre tais eventos. Esta tese pretende estender a discussão e preencher essa lacuna. Aqui, o objetivo é iluminar a redução sincrônica e diacrônica através de uma perspectiva dinânica de produção da fala. Para isso, investigaremos as características e condicionantes da redução que afetam a vogal em dois contextos diferentes e levantaremos hipóteses acerca da dinâmica de produção das vogais reduzidas no PB.

A chamada apócope de vogais altas, que resulta da ausência desse segmento na forma de onda de palavras, como [bafu] - [baf], [fasu] - [fas] (cf. Rolo, 2010), é de particular interesse para nós e se constitui no primeiro objeto de estudo da tese. À luz da Sociolinguística Variacionista, Rolo (2010) e Rolo & Mota (2012) sugeriram que as vogais altas finais [i] e [u] são apagadas. Viegas & Oliveira (2008) também argumentam na mesma direção.

À luz de Browman & Goldstein (1990), os fenômenos de redução de fala casual ocorrem, principalmente, devido a dois processos gradientes: redução na magnitude e aumento na sobreposição gestual. A redução da magnitude gestual ocorre quando uma constrição consonântica não é totalmente atigida, como por exemplo, quando uma aproximante velar é produzida no lugar de uma plosiva velar (esse mecanismo pode ser pensado para a maioria dos fenômenos de *undershoot*[2]). A sobreposição gestual, por sua vez, envolve o tempo entre movimentos articulatórios e pode levar a diferentes tipos de redução, incluindo assimilações segmentares aparentes, como também a apagamentos aparentes; por exemplo, os gestos articulatórios em grupos de consoantes consistindo de consoantes coronais e bilabiais (por exemplo, /tm/ na sentença "*perfect memory*" ou /dp/ em "*hundred pounds*"[3]) são, por vezes, sobrepostos de tal maneira que a constrição coronal é acusticamente mascarada pelo *offset* bilabial.

A visão de que os fenômenos como a redução são, sobretudo, gradientes tem sido apoiada por inúmeros estudos. Davidson (2006), a título de exemplo, mostra que o

[2] O termo *undershoot* refere-se à situação na qual os articuladores não atingem o alvo específico da vogal devido a restrições temporais. O *undershoot* foi encontrado para a maioria dos tipos de fenômenos de redução de vogais e provoca uma retração geral do espaço vocálico.

[3] "*Perfect memory*" e "*hundred pounds*" são os exemplos utilizados por Browman & Goldstein (1990) para exemplificar processos de *undershoot* e sobreposição. Por se tratar de uma revisão, achamos por bem manter os exemplos reportados pelos autores.

apagamento do schwa pré-tônico em inglês (como nas palavras "*potato*" ou "*support*"), que tinha sido tradicionalmente analisado como uma regra de eliminação fonológica (Zwicky, 1972 *apud* Davidson, 2006), é o ponto de extremo de um *continuum* de redução fonética e um fenômeno gradiante.

Em relação ao PB, Meneses (2012) e Meneses e Albano (2015) mostraram que as vogais não se apagam completamente, mas são desvozeadas. O desvozeamento é recorrente e segue o padrão descrito na literatura internacional: as vogais altas desvozeam, centralizam-se e persem amplitude. Mais do que apontar a existência de vogais desvozeadas no PB[4], o referido trabalho mostrou uma grande variabilidade entre falantes nos dados analisados, além de variação na produção da vogal: elas eram reduzidas, parcialmente desvozeadas ou totalmente desvozeadas. No entanto, as conclusões de Meneses (2012) ainda estavam aquém da complexidade envolvida no processo de desvozeamento.

Outros fenômenos da fala corrente também se mostram gradientes em vez de categóricos, como nosso outro objeto de interesse: o sândi externo de vogais. O processo de sândi vocálico foi amplamente estudado por Bisol (1992, 1996) e foi caracterizado pela autora como um processo de ressilabificação em qualquer uma das suas três manifestações: elisão, ditongação e degeminação.

A elisão, mais especificamente, diz respeito ao apagamento da vogal /a/ em posição não acentuada de final de palavra quando a palavra seguinte começa por vogal de qualidade diferente. Conforme Bisol (1992), o sândi externo é, portanto, um processo de desestruturação[5] silábica que apaga uma sílaba e deixa unidades flutuantes. Na visão da autora, o ponto de partida do processo de sândi por elisão provém de uma das sensibilidades métricas do português: a rejeição à sequência imediata de dois núcleos silábicos de vocábulos diferentes, ou seja, de hiatos.

Até então descrito como um processo categórico, o fenômeno foi reinterpretado por Albano *et al*. (1998) e Albano (1999, 2001). A autora aponta para o fato de que, em casos em que a transcrição de oitiva trata o sândi como elisão, a vogal supostamente elidida deixa um rastro claro no sinal acústico. Assim, a pronúncia de "cara idoso", por exemplo, tem, pelo menos, três versões: uma com encontro vocálico nítido, uma muito próxima de "caridoso" e

[4] Chitoran & Marsico (2010) já consideravam o PB como uma língua com vogais desvozeadas antes da existência de um estudo acústico dessa língua. Análises que tinham outro foco, como Battiste & Hermans (2003), também apontaram a possibilidade de que vogais perdessem suas características por conta da palatalização. Lemle (1966) também reportou a ocorrência de vogais surdas na fala sussurrada.

[5] O termo "desestruturação" aqui está relacionado à quebra da sílaba e à possível formação de outra.

outra intermediária, em que os valores de F1 e F2 logo após a liberação do [R] evidenciam um breve resquício da vogal /a/.

Como veremos, reduções vocálicas, como as citadas acima, são muitas vezes generalizadas como consequência da interação entre princípios da economia de esforço e restrições perceptuais (Lindblom, 1990), sem se levar em conta a coarticulação e a recorrência dos enfraquecimentos na história das línguas (não à toa, dezenas de trabalhos se apoiam nesse ponto de vista, em grande medida, para explicar fenômenos de enfraquecimento de consoantes e vogais). Há razões para acreditarmos que, pelo menos, alguns aspectos dos padrões de redução presentes na fala são específicos de cada idioma, mas também condicionados pela aerodinâmica da produção da fala e pela dinâmica dos gestos articulatórios.

Um alerta dado por Albano (1999, p. 44) é a base de toda a argumentação e da linha teórica sobre a qual discorreremos nos capítulos que seguem: "um tratamento sério dos padrões fônicos do PB passa por uma tentativa de compreender a relação entre processos gradientes e seus análogos categóricos". Portanto, na nossa visão, modelos que estabelecem um fosso entre a Fonética e a Fonologia estão fadados a tratar versões distintas do mesmo processo como radicalmente desassociadas.

Aqui, tentaremos unificar fenômenos aparentemente distintos, explicando-os a partir da ideia de que "a diferença entre o discreto e o contínuo está apenas no número de graus de liberdade" (Albano, 1999, p. 40). No tocante às vogais átonas do PB, a discussão que segue é uma primeira tentativa nessa direção.

Com o propósito descrito acima, esta tese está dividida em duas grandes partes e suas subseções. A Parte 1, que tratar da redução e do consequente desvozeamento das vogais pós-tônicas finais, é composta de três capítulos, nos quais discutimos (i) fatos e conceitos relacionados ao fenômeno estudado e à Fonologia Gestual (FonGest), modelo fonológico que sustenta toda a discussão, (ii) a releitura dos dados publicados em Meneses (2012) e (iii) novos dados aerodinâmicos e de percepção. No fim, apresentaremos dados preliminares de uma análise articulatória que podem esclarecer mais particularidades do desvozeamento vocálico.

Na segunda parte, abordamos o processo de sândi por elisão da vogal. Composta também de três capítulos, na Parte 2 apresentamos (i) a revisão dos trabalhos de cunho categórico e dinâmico, (ii) os dados acústicos e, finalmente, (iii) os dados da investigação

aerodinâmica e de percepção. Aqui, também como parte das considerações finais, mostramos resultados de uma análise inicial de uma investigação articulatória. Por fim, uma conclusão geral irá relacionar todos os resultados do ponto de vista teórico e apontar os desafios a serem resolvidos.

PARTE 1

DA REDUÇÃO À APÓCOPE:
DESVOZEAMENTO
DE VOGAIS PÓS-TÔNICAS DO PB

Capítulo 1

"*Revisão da literatura*"

1.0 Introdução: redução e desvozeamento

A primeira parte desta tese analisa o processo do PB que tem sido comumente descrito como apagamento de vogais em posição átona final. Diacronicamente, diversos processos fonológicos geraram mudanças ao longo da história do português, de forma que a apócope vocálica está entre as mais recorrentes, como por exemplo, na formação de palavras como *amare* > amar, *legale* > leal, *mense* > mês. Esse tipo de processo é, de fato, difundido em línguas românicas, dada sua origem latina.

Recentemente, vários estudos têm afirmado que as vogais finais tendem a ser apagadas em posições fracas em PB (Pagel, 1993; Viegas & Oliveira, 2008; Rolo & Mota, 2012). No entanto, pesquisas semelhantes em outros idiomas, como o inglês, o turco, o francês e o japonês, têm argumentado de forma convincente na direção de que não se trata de apócope, mas de desvozeamento (Jannedy, 1995; Tsuchida, 1997; Smith, 2003; Davidson, 2006).

Apesar de algumas especulações baseadas em registros históricos (Lemle, 1966; Viaro, 2005), só recentemente o desvozeamento foi reconhecido como tal no PB (Meneses, 2012; Meneses & Albano, 2015), assim como há também relatos de uma mudança de som mais avançada no português europeu (PE) (Fernandes, 2007; Cunha, 2015).

Diante desses fatos, o primeiro objetivo deste estudo é mostrar que há um processo de variação sincrônica na produção das vogais pós-tônicas finais do PB, que se inicia na redução e se completa com o seu desvozeamento total. Tal variação leva a ambiguidades que, por sua vez, podem resultar, diacronicamente, na apócope baseada no "erro" de percepção do ouvinte (Ohala, 1981). Um segundo objetivo desta tese é discutir os passos dessa mudança à luz da FonGest.

A FonGest explica vários fenômenos fonológicos por meio de três processos dinâmicos: a sobreposição gestual, o *blending*[1] gestual e a redução de magnitude do gesto

[1] O termo "*blending*" será usado nesta tese por ser o termo mais utilizado na literatura, na maioria das vezes, sem tradução equivalente para o português, assim como o termo "*timing*".

(Saltzman & Munhall, 1989; Browman & Goldstein, 1990; Romero, 1996; Iskarous, McDonough & Whalen, 2012). [2] A redução extrema que leva ao desvozeamento e, diacronicamente, à apócope é interpretada aqui, inicialmente, como uma diminuição da magnitude gestual condicionada pela prosódia e acompanhada pela coprodução entre o gesto da vogal e da consoante precedente.

Neste primeiro capítulo, percorrem-se alguns pontos fundamentais para o entendimento de como o trabalho trata o fenômeno em questão: a seção 1.1 apresenta o fenômeno da redução acústica com mais detalhes e, em seguida, em 1.2, como a redução está relacionada ao desvozeamento da vogal; nos tópicos 1.3 e 1.4, discutimos as evidências de uma variação da produção de vogais no PB; finalmente, no tópico 1.5 apresentamos o quadro teórico que fundamenta esta análise. Vale lembrar que este capítulo segue os achados do trabalho de Meneses & Albano (2015).

1.1 Sobre a redução

A redução vocálica é um fenômeno comum em vários idiomas e tem sido bem documentada desde Lindblom (1963).[3] Ela tem sido discutida principalmente em termos de percepção, uma vez que os ouvintes têm estratégias para recuperar a identidade das vogais, apesar de suas alterações acústicas[4] (Beckman *et al*., 1992; van Bergem, 1995). De acordo com Lindblom (1963), tais alterações incluem a centralização de F1/F2 e menor duração da vogal. A redução também está relacionada ao acento, velocidade de fala e ambiente linguístico, além do contexto pragmático, variando da hipo à hiperarticulação, segundo a teoria de Lindblom (1990)[5].

[2] Na seção 1.4, a sobreposição gestual, o *blending* e a redução de magnitude do gesto serão discutidos com mais profundidade.

[3] De forma geral, é difícil imaginar línguas nas quais não há redução vocálica. No entanto, sabe-se que, em algumas línguas, são as consoantes que sofrem a redução com mais recorrência do que vogais, como o espanhol, por exemplo.

[4] Não é o foco desta tese discutir os vários aspectos deste fenômeno, mas é uma questão intrigante saber como ouvintes superam as distorções para reconhecer as palavras apesar de desvios da pronúncia canônica. Aparentemente, os ouvintes têm diferentes maneiras de lidar com a redução. Por exemplo, eles podem compreender o contexto mais amplo e o contexto fonológico mais estreito envolvido no enunciado; ou, além disso, eles podem extrair informações sobre a forma reduzida por meio do conhecimento lexical ou pela frequência das palavras; finalmente, os ouvintes podem fazer uso de um processo auditório fino através do conhecimento fonético-fonológico da língua.

[5] A teoria de Lindblom (1990) observou que as palavras podem ser pronunciadas ao longo de um *continuum* de hiperarticulação para hipoarticulação. A hiperarticulação envolve a produção mais clara de palavras e está associada com várias características acústico-fonéticas avançadas e do esforço dos falantes, como períodos mais

Albano (1999) já alertava para a complexidade envolvida nos fenômenos de posição pós-tônica do PB e os encarava como um desafio para a FonGest, uma teoria que ainda se estabelecia à época (naquela altura, Fonologia Articulatória (FAR)). Segundo a autora, é importante atentar para a distinção entre neutralização e redução, uma vez que a última afeta o /a/, vogal que não tem par no inventário tônico do PB, não podendo, portanto, participar de processos de neutralização. Outra distinção, igualmente forte, é que a redução – mais diretamente observável no /a/ pós-tônico – é sensível a fatores gradientes, tais como estilo, taxa de elocução e força das fronteiras prosódicas.

Em Albano *et al*. (1998), os autores apresentam dados em que F1 e F2 de [ɐ] pós-tônico variam, na mesma palavra, de acordo com a força das fronteiras prosódicas (numa fronteira forte – final de enunciado –, F1 é mais baixo e F2 é mais alto, comparados a uma fronteira fraca – interior a um sintagma nominal), o que evidencia graus distintos de redução. Segundo os autores, isso permite argumentar que a redução é um processo contínuo, já que a força das fronteiras se cruza com os demais fatores gradientes mencionados, criando um leque muito amplo de possibilidades.

Já que a redução afeta também o /i/ e o /u/ pós-tônicos, sobrepondo-se ao efeito da neutralização, é evidente que [ɪ] e [ʊ] reduzidos ou [i̥] e [u̥] desvozeados não podem ser gerados exatamente pelo mesmo mecanismo de produção que [i] e [u]. Segundo a análise de Albano (1999), além da neutralização, é preciso, então, postular um processo adicional – **a redução** – que explique por que [ɪ] e [ʊ] se dispersam, aproximando-se, com isso, de outra vogal ou, até mesmo, aproximando-se da queda.

Demonstrado o caráter gradiente da redução e sustentada a hipótese da sua existência para todas as vogais pós-tônicas, o desafio está em explicá-la. Para tanto, a FonGest dispõe de um mecanismo muito apropriado: trata-se, como já mencionado, da redução da magnitude dos gestos articulatórios. Nas vogais reduzidas, encurta-se a diferença entre a posição inicial da variável do trato envolvida e o movimento dos osciladores – por meio da qual se modela o sistema dinâmico atualmente. Com isso, o gesto perde amplitude e não chega a alcançar o alvo planejado. A redução da magnitude do gesto articulatório prevê uma centralização da vogal tal como observada na redução vocálica tomando-se [i], [a] e [u], mas não é suficiente para explicar uma redução tão radical que suprima características básicas das vogais. Trataremos disso no tópico seguinte.

---

longos e maiores espaços vocálicos; a hipoarticulação, por outro lado, envolve a pronúncia menos clara de palavras, com durações mais curtas, espaços vocálicos reduzidos e queda de segmentos.

1.2 Sobre o desvozeamento

Enquanto os mecanismos de redução vocálica parecem ser razoavelmente compreendidos, o desvozeamento ainda precisa de mais esclarecimento. Como a desobstrução da cavidade oral é parte inalienável da definição de vogal, todas as vogais deveriam ser soantes, isto é, possuir o que Fant (1960) denomina vozeamento espontâneo. Ocorre que muitas línguas, dentre as quais se encontram o PB (Meneses, 2012), o búlgaro (Andreeva & Koreman, 2008), o francês (Smith, 2003), o turco (Jannedy, 1995) e, notadamente, o japonês (Tsuchida, 1994), apresentam processos de desvozeamento vocálico em ambientes surdos ou em posição final absoluta. Além das v ogais orais, as vogais nasais também estão sujeitas ao desvozeamento (SHOSTED, 2006, 2008).

Sabe-se que o desvozeamento elimina a maioria das propriedades que são o padrão esperado para as vogais, como o vozeamento, a forma de onda periódica e a estrutura formântica regular (Han, 1962; Tsuchida, 1997). Quando precedidas por uma fricativa, as vogais não vozeadas parecem uma continuação acústica da consoante precedente (Kondo, 1997, 2005).

Já é estabelecido que, em vez de um simples produto de restrições biomecânicas universais, o desvozeamento é um complexo processo da fala corrente, pois está ligado a propriedades intrínsecas das vogais e ao contexto segmental e prosódico[6]. Toda essa possibilidade de variação pode ser responsável por realizações que abrangem uma gama de possibilidades de vozeamento, desde o completo até a vogal vozeada. Esse gradiente de produção já foi mostrado em Meneses (2012).

Outros autores da literatura internacional também levantaram hipóteses sobre o desvozeamento. Entre elas, merece destaque Gordon (1998), que apresentou um levantamento de 55 línguas em que ocorrem as vogais denominadas pelo autor de não modais. Uma vogal não modal é definida, de acordo com Gordon, como qualquer vogal produzida fora dos padrões. Com base na análise de fenômenos fonéticos e fonológicos que afetam as línguas inventariadas, Gordon elenca uma série de características comuns que levam as vogais a não serem produzidas com vozeamento. Assim, os fatores que desencadeiam o desvozeamento são: articulatórios (altura, duração, segmentos adjacentes) e caráter prosódico (fronteira prosódica, a posição do acento). Tais aspectos afetam, majoritariamente, as vogais átonas.

[6] Os contextos podem ser segmentos adjacentes, atonicidade da sílaba, posição no enunciado, entre outros fatores.

Chitoran & Marsico (2010) também apresentam um estudo sobre os fatores que favorecem a vogal desvozeada em 39 idiomas, entre eles o PB. Os dados citados pelos autores vão ao encontro do levantamento feito por Gordon (1998), uma vez que ambos os trabalhos sugerem a existência de restrições ao desvozeamento, dependendo do tipo de segmentos em proximidade da vogal.

Para Kager (1997), a gradiente eliminação da vogal é fonologicamente incompleta, ocorrendo em variação livre com a sua redução. Esse processo tende a preservar a silabicidade da vogal "suprimida", o que pode ser marcado por sinais fonéticos. Para o autor, a eliminação categórica da vogal não tem correspondência com a redução da vogal, destruindo, assim, a silabicidade da vogal apagada. Na mesma linha, Kondo (1997) considera o desvozeamento mais do que uma simples perda da voz, e passa a interpretá-lo como uma das etapas do que ela chama de "enfraquecimento da vogal", a saber, redução, ensurdecimento e, eventualmente, eliminação.

Torreira & Ernestus (2010) investigaram formas reduzidas (segmentos e palavras reduzidas) em diversas línguas, incluindo o desvozeamento, indicando que o grau de redução que um determinado tipo de segmento pode sofrer no discurso coloquial pode variar significativamente entre línguas. Por exemplo, em termos de grau de constrição e vozeamento, o espanhol e o francês parecem ter mecanismos diferentes no que se refere à redução de segmentos. No espanhol, as oclusivas desvozeadas intervocálicas parecem ser coarticuladas às vogais adjacentes em um grau maior do que oclusivas do francês. No francês, por outro lado, as vogais mostram mais sinais da influência coarticulatória da consoante vizinha do que as vogais espanholas – esse parece ser também o caso do PB, no qual as vogais parecem ser altamente coproduzidas e, com a falta do acento tônico, mais suscetíveis à coarticulação (coprodução).

No âmbito da FonGest, a literatura sobre o desvozeamento de vogais produziu até agora um único estudo (Chitoran & Iskarous, 2008)[7], muito preliminar. Com base em dados acústicos, os autores argumentam que o desvozeamento em lezgui (língua falada no sul da Rússia e no Arzebaijão) resulta de alterações gestuais que levam à sobreposição entre o gesto consonantal adjacente e os gestos articulatórios vocálicos. Eles explicam o desvozeamento em termos de alterações na magnitude e *timing* de tais gestos.

[7] O artigo de Meneses & Albano (2015), que é fruto do trabalho realizado nesta tese de doutorado, também segue o mesmo modelo.

Apesar de ter feito avanços na descrição da redução de vogal e do desvozeamento como fenômenos independentes, a literatura atual não foi além da proposta inicial de Kondo (1997) e (Chitoran & Iskarous, 2008), nem mesmo tentou unificar o seu tratamento teórico. Adiantamos que nossa investigação tenta preencher essa lacuna.

1.3 Evidências da diacronia e variação sincrônica de vogais no PB

Como dito no tópico inicial, o português já sofreu mudanças por meio de diversos processos fonológicos durante sua história, sendo que a apócope é um dos fenômenos recorrentes. Esse tipo de processo é, de fato, difundido em línguas românicas devido às suas raízes no latim vulgar (Taylor, 1994; Wheeler, 2007). O apagamento tende a acontecer primeiramente para as sílabas pós-tônicas que têm mantido o contorno acentual assimétrico do latim vulgar (uma reinterpretação da regra do acento latino, na qual a intensidade cai abruptamente depois do acento tônico). Mais tarde, espalhou-se para sílabas pretônicas, como síncope (Taylor, 1994; Wheeler, 2007; Cunha, 2015), em línguas que tendem ao ritmo de *stress-timing*, como o catalão e o PE (Ramus, Dupoux & Mehler, 2003). No entanto, os atuais contornos acentuais mais conservadores em PB, galego e espanhol são mais susceptíveis à apócope. Essa tendência é observada, por exemplo, na evolução do latim vulgar para o castelhano e o português, na qual podemos observar a apócope de /e/, como demostram os exemplos abaixo:

(1) a. *amare* (latim) – amar (castelhano e português)

b. *hospitale* (latim) – hospital (castelhano e português)

Se tomarmos também o inglês como exemplo (cf. Campbell, 1959), é possível perceber que houve uma tendência única e óbvia que se aplica não apenas ao inglês antigo, mas à história do inglês como um todo:

> *"(...) sound tended to be reduced, so that, for example, long vowels become short, short vowels lost their distinctive phonetic characteristics and merged, eventually as the reduced vowel schwa, and reduced vowels were lost. Thus if we take the Germanic word *namani 'name', this developed in Old English through the stage *namani to*

> *naman. If we move to Middle English and to the accusative singular form we then see the development naman > nama > name (=[na:mə]) > (=[na:m]). (...) we shall consider the exemplification of these trends in two specific areas, namely the reduction in variety of unstressed vowels and the loss of unstressed vowels.*"[8] (Hogg, 1992, p. 109)

Outro fator diacrônico relevante para a nossa discussão diz respeito à natureza da vogal que sofre apócope. O processo no inglês parece ocorrer historicamente com as vogais /i/ e /u/, como por exemplo, em *fe:ti '*feet*' > fe:t > fi:t. Não causa estranheza que tais vogais sejam as mais suscetíveis também à redução e ao desvozeamento, uma vez que elas têm mais ruído intrínseco e menos intensidade intrínseca se comparadas com vogais mais baixas. A relação entre a redução e o desvozeamento tem boas pistas em gramáticas históricas, que registram abundantemente a instabilidade das vogais em ambientes nos quais elas desapareceram tardiamente.

Do ponto de vista sincrônico, uma das características do PB é a diferença entre os seus sistemas vocálicos tônico e átono. Quando não recebem acento, as vogais geralmente diminuem a sua duração relativa e sua intensidade. Essa modificação de qualidade vocálica deixa as vogais não acentuadas mais vulneráveis aos fenômenos de redução e de apagamento. Para entender como esses processos podem ter afetado o PB, duas considerações são necessárias.

A primeira diz respeito à redução do inventário vocálico dependendo do acento. De acordo com a Câmara Jr. (1953, 1970), as vogais do PB sofrem neutralização da seguinte forma: há sete vogais – / a, ɛ, e, i, ɔ, o, u / – na posição acentuada, cinco vogais – / a, e, i, o, u / – na posição pretônica, quatro vogais – / a, i, o, u / – na posição postônica não final e apenas três vogais – / a, i, u / – na posição pós-tônica final. Isso contribui para a natureza assimétrica do contorno acentual do PB, o qual se inicia relativamente alto, aumenta suavemente até a vogal acentuada e, em seguida, cai abruptamente.

[8] Tradução nossa: *"(...) os sons tendem a ser reduzidos, de modo que, por exemplo, vogais longas tornam-se curtas, vogais curtas perdem suas características fonéticas distintivas e mescladas, eventualmente como o schwa reduzido e vogais reduzidas foram perdidos. Assim, se tomarmos a palavra germânica *namani 'nome', desenvolvido em Inglês Antigo através do estágio *namani para naman. Se nos movermos para o Inglês Médio e para a forma singular do acusativo vemos, então, o desenvolvimento de naman > nama > name (= [na:mə])> (=[na:m]). (...) Vamos considerar exemplos dessas tendências em duas áreas específicas, a saber, a redução da variedade de vogais átonas e à perda de vogais átonas*"

O segundo ponto a considerar é a extrema fraqueza da posição pós-tônica, que, segundo Câmara Jr. (1970), carrega esforço zero em uma escala de zero a três. Por conseguinte, essa posição é a mais suscetível à redução radical e ao desvozeamento do tipo examinado neste e em outros trabalhos do PB (Aquino, 1997; Ferreira, 2007; Napoleão, 2012; Dubiela, 2013).

É por isso que a posição pós-tônica final foi o foco do estudo de Meneses (2012), que visava a investigar o lado gradiente da "apócope" vocálica. O argumento é que essas vogais não são completamente apagadas, mas apenas parcialmente ou totalmente desvozeadas. Aqui vamos partir das descobertas de Meneses para mostrar que a redução da vogal em posição pós-tônica constitui um *continuum* e, provavelmente, está relacionada às etapas do apagamento da vogal proposta por Kondo (2005). No entanto, ao contrário de Kondo, e em linha com Chitoran & Iskarous (2008), vamos tentar tratar o processo em questão como um fenômeno sincrônico gradiente, que pode resultar em mudança, e dar-lhe uma interpretação teórica dinâmica.

1.4 Mudança sonora em andamento

Partindo do pressuposto de que há uma variação sonora em curso no PB, há pelo menos três perspectivas que podem contribuir para o entendimento dos passos envolvidos nas mudanças que levam da redução/desvozeamento à apócope.

A primeira é filiada à sociolinguística variacionista, que estabelece que a mudança de som emerge da variabilidade sincrônica (Labov, 1972) e concebe a mudança como gradual, evoluindo de geração em geração, em resposta a fatores sociolinguísticos como idade, escolaridade, *status* socioeconômico etc.

A segunda perspectiva vê a mudança como um processo mais abrupto e alega que a percepção de um som diferente daquele a que se destina pode surgir quando um ouvinte não pode/consegue compensar os efeitos do contexto (Ohala, 1981, 1993). De acordo com Ohala (1993, p. 1),

> "diachronic variation emerges for the most part from a synchronic set of possibilities of variation forms thus: universal and timeless physical constraints on speech production and perception leads listeners to misapprehend the speech signal, this "mistake" is potentially the beginning of sound change. As affirmed by Ohala, it

> affects the way the listener himself speaks and this changed pronunciation may be imitated by other speakers. Speech changes are perceptual errors which give hints as what cues listeners use when speech perception is successful".[9]

É importante salientar que, nessa visão, a variação diacrônica também emerge de um conjunto de possibilidades de variações sincrônicas formadas a partir de restrições físicas sobre a produção e percepção da fala. Sob essa luz, qualquer equívoco ou *misparsing* pode ser o início de uma mudança de som.

A terceira visão também considera a mudança gradual, mas assume que ela pode ocorrer não só entre falantes e gerações, mas também para um mesmo falante (Harrington, Fletcher & Roberts, 1995; Harrington 2012). Esses estudos são, na verdade, um desenvolvimento da visão gestual da mudança sonora proposta por Browman & Goldstein (1989, 1990), que afirma que a sobreposição e o *blending* gestual desempenham papel fundamental na variação e estabilização de padrões fonológicos. Em boa parte desta tese, vamos explorar os fenômenos relacionados tanto à sobreposição quanto ao *blending*. Além disso, também vamos explorar a ideia de que o "erro perceptual" pode facilitar a mudança sonora.

1.5 Enquadramento teórico e termos basilares

Em relação ao enquadramento teórico desta pesquisa, tem sido um elo comum e característica recorrente das últimas teses defendidas junto ao Laboratório de Acústica e Psicolinguística (LAFAPE) do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (IEL/Unicamp) uma defesa – como um tutorial – da FonGest enquanto Modelo Fonológico. Uma vez que tal modelo já se mostra bem estabelecido e tem vasta produção acadêmica, esta tese não se estenderá na tarefa de detalhá-lo.

---

[9] Tradução nossa: *"a variação diacrônica emerge para a maior parte de um conjunto sincrônicas de possibilidades de formas de variação como: limitações físicas universais e intemporais na produção e percepção da fala leva os ouvintes ao erro de compreensão do sinal da fala, este "erro" é potencialmente o início da mudança sonora. Como afirmado por Ohala, isso afeta a maneira como o próprio ouvinte fala e essa alteração de pronúncia pode ser imitada por outros falantes. Mudanças de fala são erros perceptuais que dão sugestões como pistas para os ouvintes quando a percepção é bem sucedida."*

Nesta seção, por conseguinte, apresentamos apenas conceitos considerados fundamentais na interpretação do fenômeno em estudo. Para aqueles que buscam uma extensa descrição da FonGest, bem como de seu primitivo fonológico – o gesto –, sugerimos os trabalhos de Albano (1999, 2001), Silva (2002), Berti (2006), Pozzani (2011) e outras teses defendidas junto ao LAFAPE. Sobre as bases teóricas modelares, como Sistemas Dinâmicos e outros conceitos importantes, sugerimos fortemente o número especial da Revista da ABRALIN de 2012, organizado pela professora Eleonora Albano, sobre a FonGest (Albano, 2012).

A noção de que processos fonéticos/fonológicos[10] não são categóricos, mas gradientes, é o coração da FonGest (Browman & Goldstein 1987, 1989, 1990, 1992). Gestos, nesse contexto, são oscilações dos articuladores no trato vocal durante a produção da fala e são especificados espaciotemporalmente – essa é a principal diferença desse modelo para outros modelos fonológicos. Tomando-se por base tais premissas, fenômenos fonológicos tais quais substituições, apagamentos e epênteses podem ser explicados como resultados de três processos fundamentais: redução da magnitude gestual, sobreposição de gestos e o *blending*.

A redução em magnitude gestual ocorre quando uma constrição especificada não é totalmente alcançada, como por exemplo, quando uma aproximante velar é produzida no lugar de uma plosiva velar.

A magnitude e o tempo de ativação dos gestos são diretamente responsáveis pela sobreposição. De forma mais clara, a sobreposição é uma relação de fase na qual um gesto se inicia antes do final do outro. Essa é uma noção de grande importância na FonGest e pode explicar variações alofônicas, fenômenos de coarticulação, entre outros processos. Browman & Goldstein (1992) afirmam que sequências como [mp] – ou [mb] –, por exemplo, são causadas pela sobreposição do gesto labial do /p/ sobre o abaixamento do véu associado com o /n/. O gesto do /n/ não desaparece, mas é sobreposto para gesto labial.

No modelo inicial da FonGest, Browman & Goldstein (1989, 1990) diferenciam duas camadas, ou grupos de filiação dos gestos: a camada articulatória e a camada funcional. As definições de cada camada ficam claras no texto de Romero (1996, p. 101):

> "Articulatory tiers are defined on the basis of articulatory independence and correspond to the major articulator that is activated during the production of a

[10] O modelo da FonGest é uma tentativa de integrar os domínios da fonética e fonologia. Nesse modelo dinâmico de produção da fala, os gestos articulatórios surgem como unidades fonológicas e, ao mesmo tempo, unidades articulatórias (físicas). As unidades na FonGest são unidades de ação e, como consequência, não estáticas e potencialmente sobrepostas.

> particular gesture. Major articulators include the lips, the tongue tip, and the tongue dorsum. Functional tiers, on the other hand, separate consonants and vowels into two different levels. This allows for more accurate specification of the particular phasing relationships between consonants and vowels.".[11]

Quando a sobreposição ocorre na mesma camada articulatória, os gestos não podem se sobrepor sem afetar as respectivas trajetórias. Nesses casos, o resultado da sobreposição é o *blending* dos gestos envolvidos (Romero, 1996). Trata-se de uma forma de sobreposição extrema[12] na qual dois gestos compartilham o mesmo articulador. No *blending,* um segmento tende a mascarar o outro (Harrington, Fletcher & Roberts, 1995; Harrington, 2012). De acordo com Zsiga (2000) e Iskarous, McDonough & Whalen (2012), uma sobreposição extensa entre segmentos tão diferentes como uma obstruinte e uma vogal (ou uma aproximante) sempre leva ao *blending* devido à necessidade de ajuste do grau de constrição.

Como se sabe, modelos dinâmicos têm sucesso em explicar processos fônicos da fala rápida justamente por serem capazes de mostrar a existência de gradientes fônicos que não são bem representados pelas análises tradicionais.

A fim de aprofundar o estudo sobre o desvozeamento no PB, vamos primeiro apresentar e rever os achados acústicos de Meneses (2012) em termos do mecanismo gestual da vogal desvozeada. Vamos tentar demonstrar que existe uma continuidade de redução da vogal manifestada sincronicamente como uma extrema centralização da vogal e sobreposição parcial ou total pela consoante precedente, o que resulta em vogais extremamente curtas, parcial ou totalmente desvozeadas, as quais, mais para frente, chamaremos de "vogais fricatizadas". Essa variação pode, eventualmente, constituir a matéria-prima para as fases subsequentes da mudança sonora, resultando em uma apócope. Enquanto as fases iniciais de tal mudança tendem a ser relacionadas com o *timing* de produção, a sua última etapa é, muitas vezes, gerada por uma releitura perceptual do sinal acústico (Ohala, 1981).

---

[11] Tradução nossa: *"As camadas articulatórias são definidas com base na independência articulatória e correspondem ao articulador principal que é ativado durante a produção de um gesto particular. Os articuladores principais incluem os lábios, a ponta da língua, e o dorso da língua. As camadas funcionais, por outro lado, separam consoantes e vogais em dois níveis diferentes. Isto permite a especificação mais precisa das relações particulares de faseamento entre consoantes e vogais."*

[12] Isso é diferente de uma simples coordenação em fase, na qual a vogal atinge o seu destino após o deslocamento consoante. No *blending*, a vogal atinge o seu alvo antes do *offset* da consoante.

Capítulo 2

"*Uma nova visão dos resultados de Meneses (2012)*"

2.0 Introdução

A proposta do estudo realizado por Meneses (2012) foi averiguar o processo de desvozeamento vocálico partindo de um modelo dinâmico de produção da fala. Em contraposição aos modelos tradicionais, procurou-se explicar como as vogais desvozeadas são produzidas e como são gradualmente aceitas como variantes. Além disso, buscou-se sintetizar as pistas acústicas encontradas e relacioná-las aos achados de outras línguas. De forma geral e direta, o resultado mais importante de Meneses foi mostrar, no sinal das chamadas sílabas desvozeadas, rastros da vogal que impossibilitam uma hipótese de apócope vocálica.

Neste capítulo, antes de mostrar os resultados (seção 2.3) e lançar novas questões sobre os dados de Meneses, revisaremos a metodologia aplicada (seção 2.1) e os dados (seção 2.2) da referida dissertação de mestrado.

2.1 Experimento I: breve revisão da metodologia usada em Meneses (2012)

2.1.1 Materiais e participantes

Para explorar a natureza das vogais pós-tônicas finais do PB, foram gravadas sequências de palavras /'(C)V.CV/. Dez palavras-alvo foram escolhidas para exemplificar cada uma das três vogais diferentes, a saber, /a, i, u/ precedidas de /s/, como pode ser visto no Quadro 1 a seguir. A frase-veículo foi produzida com a mesma estrutura e número de sílabas para cada palavra-alvo, como por exemplo, "O X parecia perigoso". Cada sujeito leu as frases cinco vezes. Para comparar áreas de espaço vocálico, as mesmas palavras foram gravadas em um ambiente sonoro, como "Disse X baixinho".

Quadro 1 – Palavras-alvo usadas no Experimento 1

Vogal-alvo	Palavra-alvo		
/a/	caça	"*hunting*"	/'kasa/
	massa	"*mass*"	/'masa/
	taça	"*cup"*	/'tasa/
/i/	lance	"*throw*"	/'lãnsi/
	passe	"*pass*"	/'pasi/
	tosse	"*cough*"	/'tɔsi/
/u/	aço	"*steel*"	/'asu/
	baço	"*spleen*"	/'basu/
	maço	"*pack*"	/'masu/

Fonte: Meneses (2012, p. 91)

Os participantes foram seis indivíduos do sexo feminino ($S1_1$…$S6_1$[13]) de Vitória da Conquista, uma cidade no estado da Bahia, onde /s/ final não é palatalizado como em outros dialetos do Nordeste. O gênero dos participantes foi o mesmo para facilitar a normalização. Faixa etária (25-40 anos), escolaridade (alunos de graduação) e nível socioeconômico (classe média) também foram os mesmos.

2.1.2 Medidas

Várias medidas acústicas foram obtidas na primeira análise realizada em Meneses (2012). Entre elas, as mais importantes foram o centroide do ruído de /s/ e as durações da sílaba-alvo, do ruído e da vogal. Para essa análise, a fim de comparar os contextos que favoreciam e desfavoreciam o desvozeamento, a Área de Espaço Vocálico (VSA)[14] foi calculada como um índice de dispersão vocálica. A correlação entre o VSA e a duração da vogal foi computada como medida quantificadora da centralização. Essa última medida, especificamente, não foi publicada em Meneses, mas é proveniente do conjunto de dados colhidos para a dissertação.

2.1.3 Análise acústica

O desvozeamento foi classificado como parcial quando a vogal sobreposta pela consoante tinha uma porção periódica curta com alguns pulsos glóticos (média = 4) e duração

[13] Os números subscritos referem-se ao experimento em questão.
[14] Do inglês, *Vowel Space Area*.

reduzida (média = 20ms). A vogal foi classificada como completamente desvozeada quando apresentou uma porção completamente aperiódica sem pulso glótico e com configuração formântica indefinida. É importante salientar que tal configuração não é igualmente passível de medições acústicas de duração. Assim, no desvozeamento total, considera-se a duração do ruído como a totalidade da sílaba, uma vez que consoante e vogal estão totalmente fundidas.

Na análise espectral, as medidas foram realizadas automaticamente por um *script* PRAAT concebido para: (i) gerar uma janela FFT; (ii) calcular o comprimento do segmento rotulado e definir o seu centro; (iii) estimar o primeiro momento espectral do ruído. Como a sobreposição entre vogais e consoantes tende a afetar as medidas do centroide e tais efeitos são muitas vezes localizados em regiões relativamente altas do espectro, as medições foram realizadas com um filtro acima de 4 KHz. Essa banda foi escolhida porque fricativas alveolares geralmente têm pouca energia abaixo desse ponto (cf. Vaissière, 2009).

O VSA, conforme apresentado na subseção anterior, avalia as distâncias relativas das vogais de um sujeito e é utilizado para quantificar e representar a centralização vocálica (Kent & Kim, 2003). Os valores VSA foram obtidos pela fórmula: VSA = ABS ((F1i * (F2a - F2u) + F1a * (F2u-F2i) + F1u * (F2i-F2a)) / 2). Esses espaços foram estimados por regras de triangulação, de acordo com a geometria plana. Assim, a centralização é indicada por valores baixos do VSA.

2.1.4 Análise estatística

O teste t pareado foi utilizado para comparar valores dos centroides dos ruídos. Em seguida, o coeficiente de Pearson foi usado para correlacionar o VSA com a duração da vogal. As diferenças de duração foram avaliadas pelo teste de Kruskal-Wallis. Os valores da duração foram normalizados pela transformação de *z-score*.

2.2 Revisão e releitura dos dados

2.2.1 Achados gerais

Em geral, os dados de Meneses (2012) confirmaram os padrões acústicos descritos na literatura (por exemplo Jannedy, 1995; Tsuchida, 1997 *inter alia*). Vogais altas perdem vozeamento (por vezes de forma incompleta), periodicidade e estrutura formântica, como mostrado na Figura 1. Observemos abaixo que, como apontado por Han (1962), Kondo

(2005), Jannedy (1995) e Tsuchida (1997), o padrão formântico não está claro no caso de ensurdecimento (C) em comparação com o desvozeamento parcial (B) e com a vogal vozeada em (A), a qual apresenta limites bem definidos entre a fricativa e a vogal.

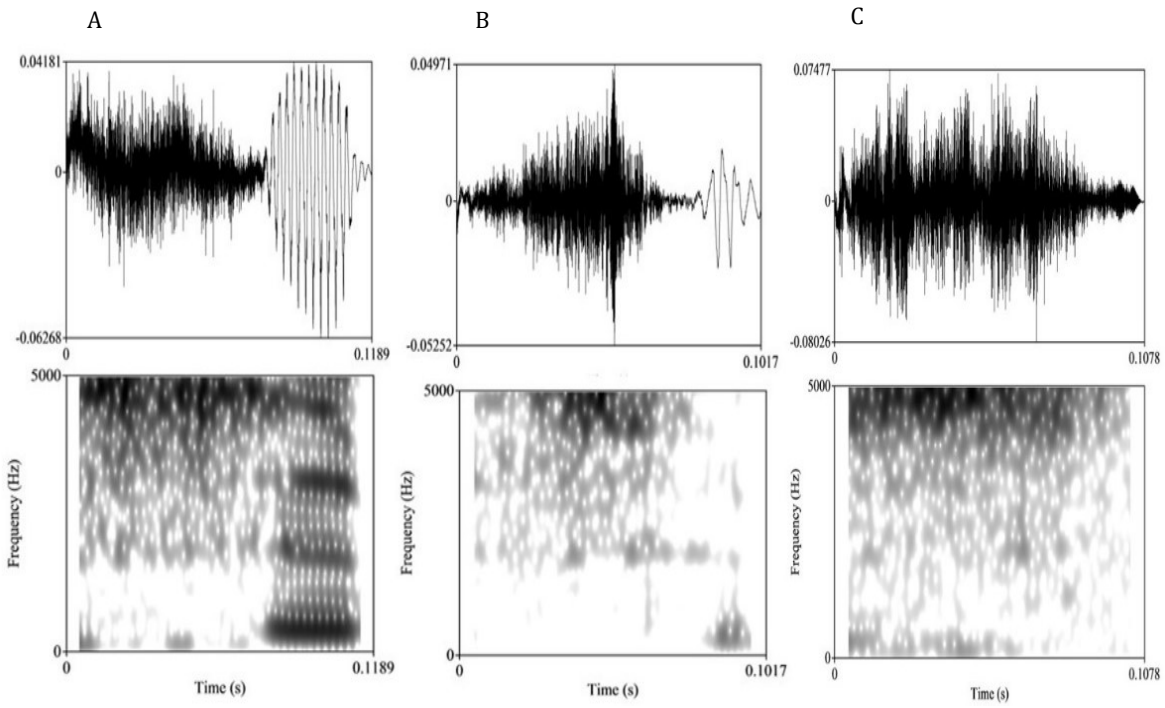


Figura 1 – Exemplos de sílabas com vogais vozeadas, parcialmente desvozeadas e totalmente desvozeadas, respectivamente, extraídas da palavra "lance" /laNsi/[15]. Na coluna A, os limites estão bem definidos entre a consoante e a vogal; na coluna B, é possível ver a diminuição radical da magnitude da vogal, que tem duração bem reduzida e poucos pulsos; na coluna C, observa-se apenas o ruído que representa a sílaba desvozeada.

As proporções de vogais desvozeadas de diferentes alturas reportadas estão também em consonância[16] com a literatura. As porcentagens médias totais de desvozeamento são de 46% para /u/, 42% para /i/ e 12% para /a/. A taxa de desvozeamento parcial também é alta para /u/ e /i/ e baixa para /a/. Essa variação pode ser constatada se analisarmos os dados gerais ou cada sujeito separadamente (como veremos adiante na Figura 4).

2.2.2 Centralização vocálica extrema

[15] A literatura fonológica do PB comumente utiliza um arquifonema nasal /N/, postulado inicialmente por Câmara Jr. (1953), para transcrições de consoantes homorgânicas como em /laNsi/.

[16] É consenso que as vogais altas são as que mais sofrem a redução extrema. No entanto, entre /i/ e /u/, é reportada maior recorrência de desvozeamento da vogal alta /i/. Nesse ponto, os dados apresentados divergem, mesmo que minimamente, da literatura.

Já foi mencionado que a redução da vogal é favorecida na posição pós-tônica no PB e em outras línguas. Nossos resultados mostram que a redução é ainda maior do que o esperado em contextos que provocam desvozeamento. A Figura 2 abaixo mostra as áreas de espaço de vogais átonas finais precedidas de /s/ em duas condições: seguidas de consoantes desvozeadas e seguidas de consoantes vozeadas. O espaço definido por vogais seguidas de consoantes desvozeadas é mais comprimido do que o espaço definido pela vogais seguidas de consoantes sonoras.

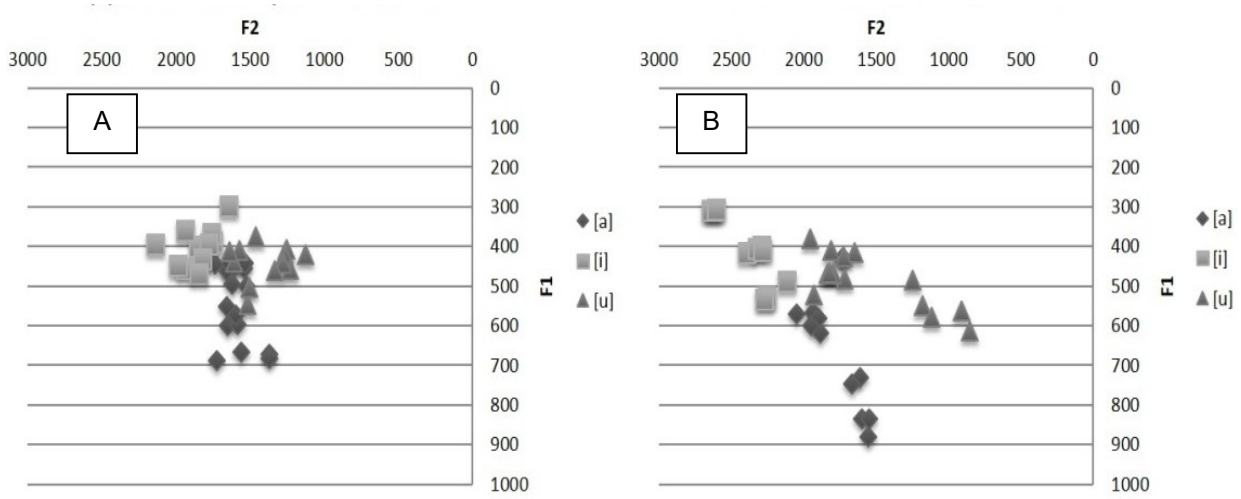


Figura 2 – Área de espaço vocálico de vogais seguidas de consoantes desvozeadas (A) e vozeadas (B)
Fonte: Meneses (2012, p. 86)

A Figura 3, a seguir, mostra a correlação entre os VSAs dos sujeitos e a duração do conjunto das vogais parcialmente desvozeadas e vozeadas. Há uma tendência clara de que as vogais mais centralizadas sejam mais curtas do que vogais menos centralizados. A correlação positiva (**R (24) = 0,53, *p* = 0,005**) indica que a centralização tende a ser diretamente proporcional ao encurtamento da vogal. Esse dado sugere fortemente que a redução da vogal pós-tônica é ainda mais proeminente nas sílabas que começam com consoantes surdas.

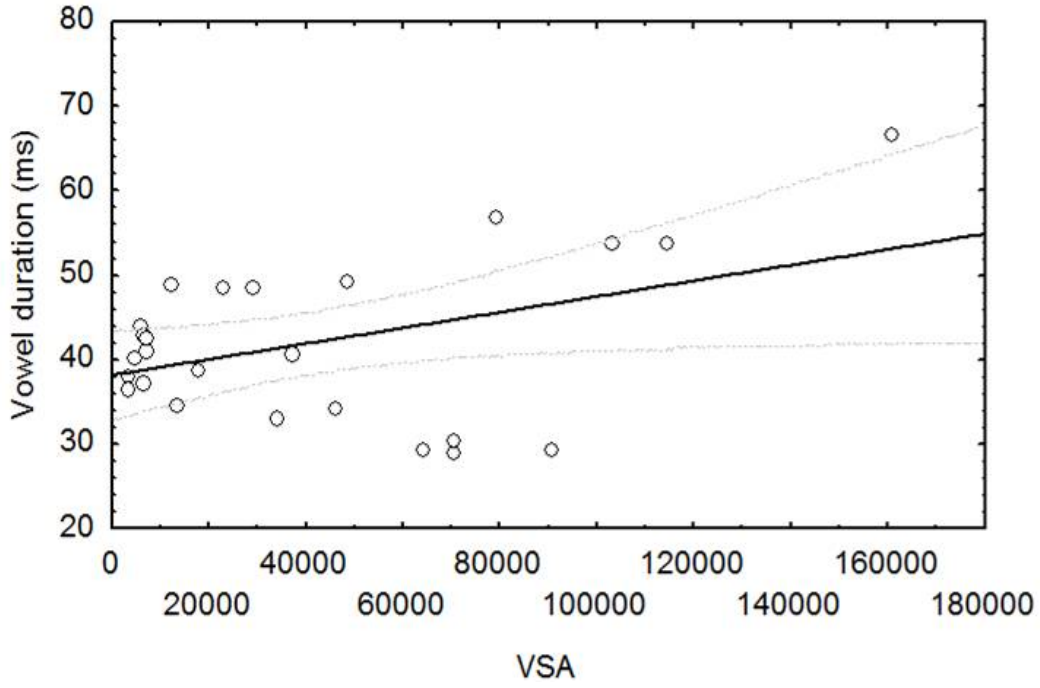


Figura 3 – Diagrama de dispersão entre o VSA dos sujeitos e a duração das vogais em contexto desvozeado (R (24) = 0,53, *p* = 0,005)
Fonte: Meneses (2015, p. 68)

A Figura 4 a seguir mostra as diferenças quanto à ocorrência de desvozeamento entre os sujeitos analisados. Se tormarmos, por exemplo, os $S1_1$ e $S5_1$, os sujeitos extremos quanto à variação do vozeamento, é possível entender melhor a relação entre a centralização e a redução da Figura 3. Enquanto $S1_1$ (à esquerda) desvozeia praticamente todas as vogais altas, $S5_1$ (à direita) tem uma baixa ocorrência de desvozeamento, produzindo mais vogais vozeadas ou parcialmente desvozeadas. As vogais de $S1_1$ tendem a ser mais curtas (em média, 69 ms) e mais centralizadas (VSA médio = 17,02). As vogais de $S5_1$, por sua vez, tendem a ser mais longas (em média, 91 ms) e menos centralizadas (VSA = 103,21). Assim, $S1_1$ parece ser mais avançada do que $S5_1$ no que diz respeito à mudança de som em questão.

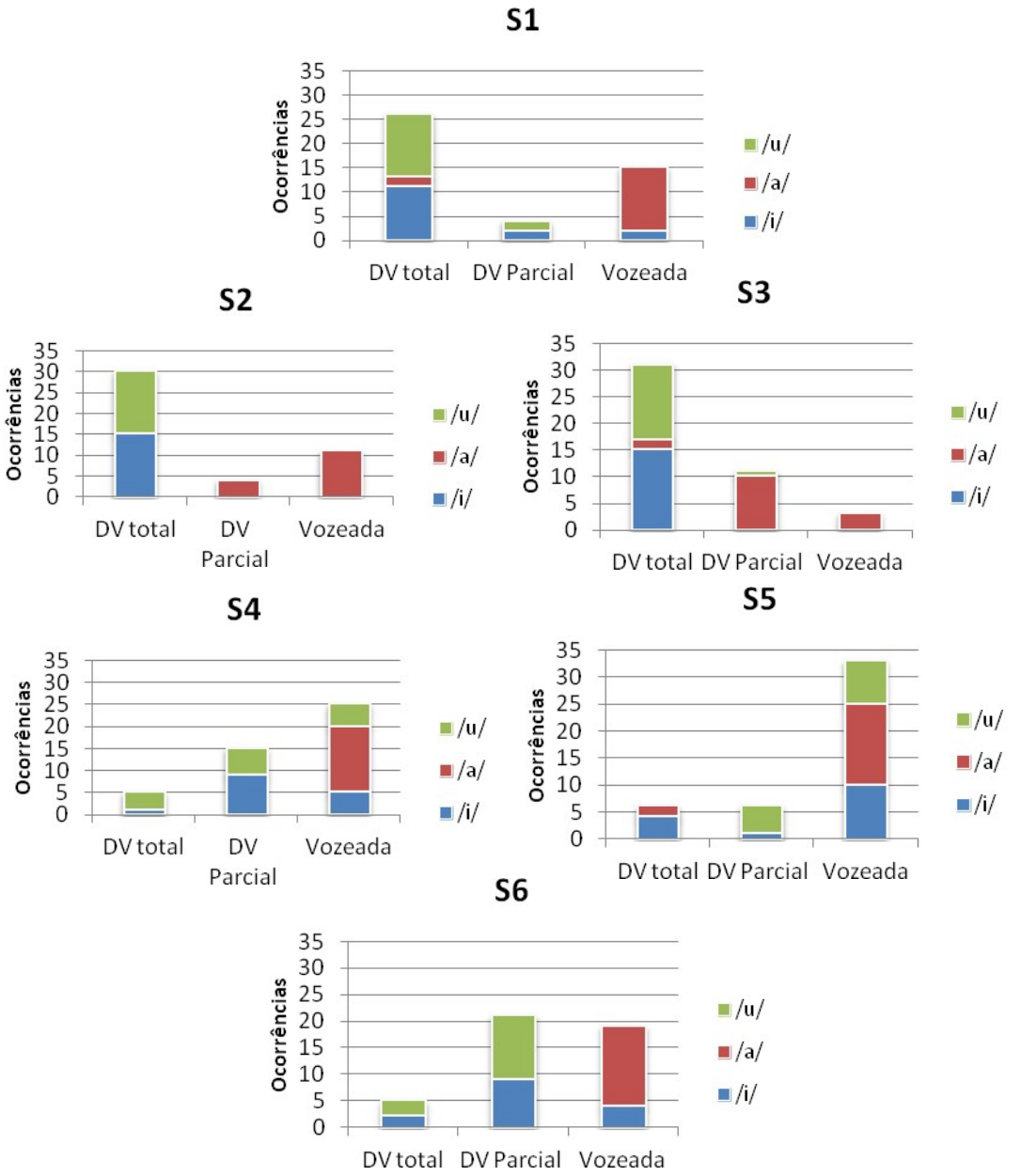


Figura 4 – Ocorrência de vogais desvozeadas, parcialmente desvozeadas e vozeadas para todos os sujeitos analisados em Meneses (2012)
Fonte: Meneses (2012, p. 45)

Está claro que as vogais átonas finais seguidas de consoantes surdas sofrem redução ainda mais radical em sujeitos tais como $S1_1$. Assim como ela, $S2_1$ e $S3_1$ também são extremos se comparadas à S4, S5 e S6.

2.2.3 Centroide

A Figura 5 mostra a diferença entre os centroides do ruído de /s/, de /i/ desvozeado e de / i / vozeado. Essa diferença é estatisticamente significativa (t (2) = -6,64 , p < 0:02).

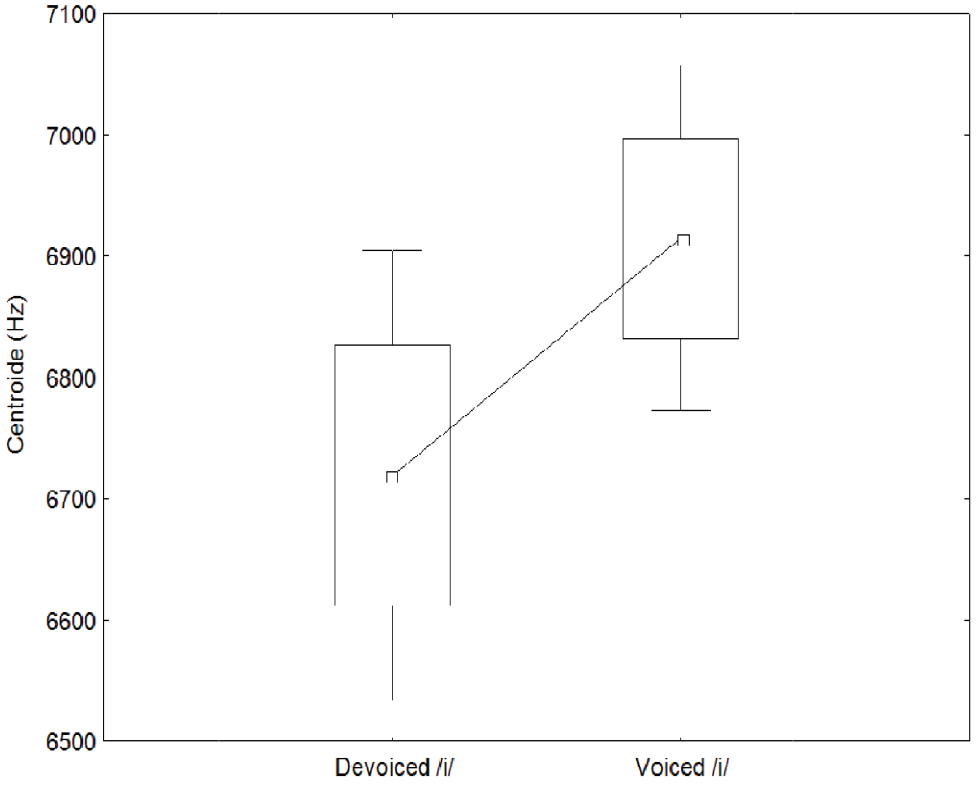


Figura 5 – Centroide da fricativa em sílabas "sem vogal" (/i/ desvozeado) e com vogal (/i/ vozeado)
Fonte: Meneses (2012, p. 60)

Em palavras que terminam em /i/, a diferença entre o centroide do ruído na presença de uma vogal visível e o centroide do ruído na sua ausência é de aproximadamente 200 Hz. O valor mais baixo do centroide indica que o ruído contém pistas acústicas da vogal quando ela parece ausente. É importante ressaltar que a elevada taxa de desvozeamento no *corpus* analisado e a variabilidade intra e intersujeitos para a vogal alta /u/ impediu que essa análise fosse realizada para ela. Tal diferença vai ao encontro dos resultados reportados por Tsuchida (1994) e Torreira & Ernestus (2010), que afirmam, respectivamente, que centros de gravidade mais baixos fornecem pistas de presença de vogal e que o aumento da coarticulação

de /s + V/ na sílaba desvozeada resulta em um deslocamento dos parâmetros espectrais para baixo se comparados a /s/ em sílaba de vogal vozeada.

2.2.4 Dados de duração

Geralmente, o desvozeamento da vogal é acompanhado pela compressão temporal, devido à pesada coarticulação envolvida (Smith, 2003; Fagyal & Moisset, 1999). Os dados a seguir, sobre a duração das vogais pós-tônicas finais, apoiam essa interpretação e introduzem um novo achado, a saber: a duração do ruído nos casos de desvozeamento é maior do que nos demais ambientes. As medidas são resumidas na Figura 6, que exibe as durações médias dos ruídos de /s/ para todos os três casos: desvozeamento total, desvozeamento parcial e vozeamento.

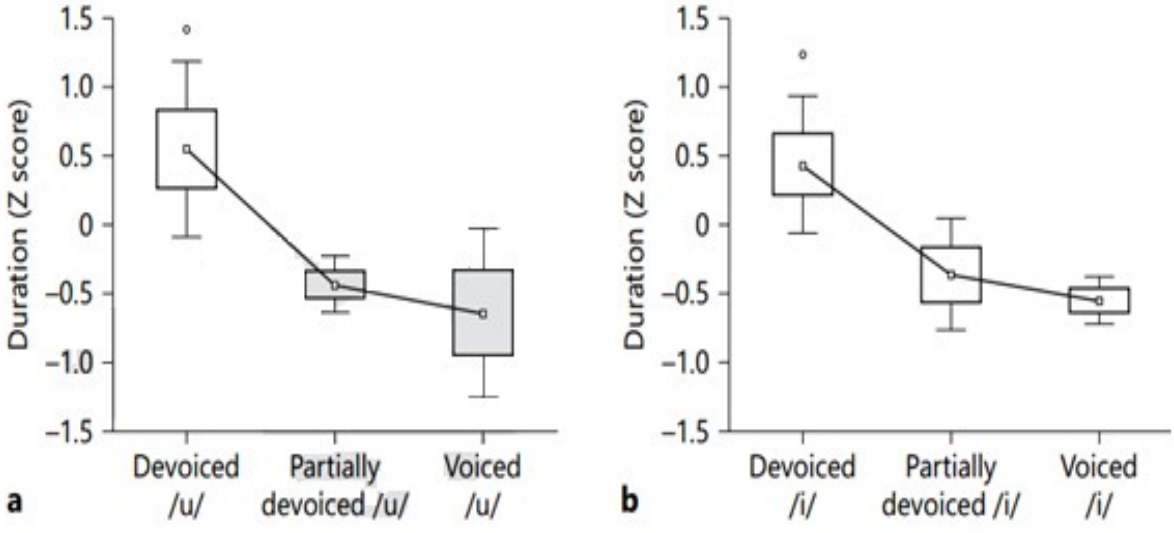


Figura 6 – Duração do ruído do [s] com vogal vozeada, parcialmente desvozeada e totalmente desvozeada
Fonte: Meneses (2012, p. 52-53)

A Figura 6 acima mostra que o /s/, seguido por uma vogal desvozeada, tem duração maior do que nos outros contextos, tanto para /i/ (à direita) quanto para /u/ (à esquerda). O teste de Kruskal-Wallis indica que as durações médias do ruído diferem ($H = 8{,}12$, $d.f. = 2$, $p < 0{,}02$). A Figura 7 a seguir mostra que a tendência para a duração das sílabas é exatamente a oposta: as sílabas que contêm vogais vozeados são mais longas do que aquelas que apresentam vogais parcialmente desvozeadas ou totalmente desvozeadas.

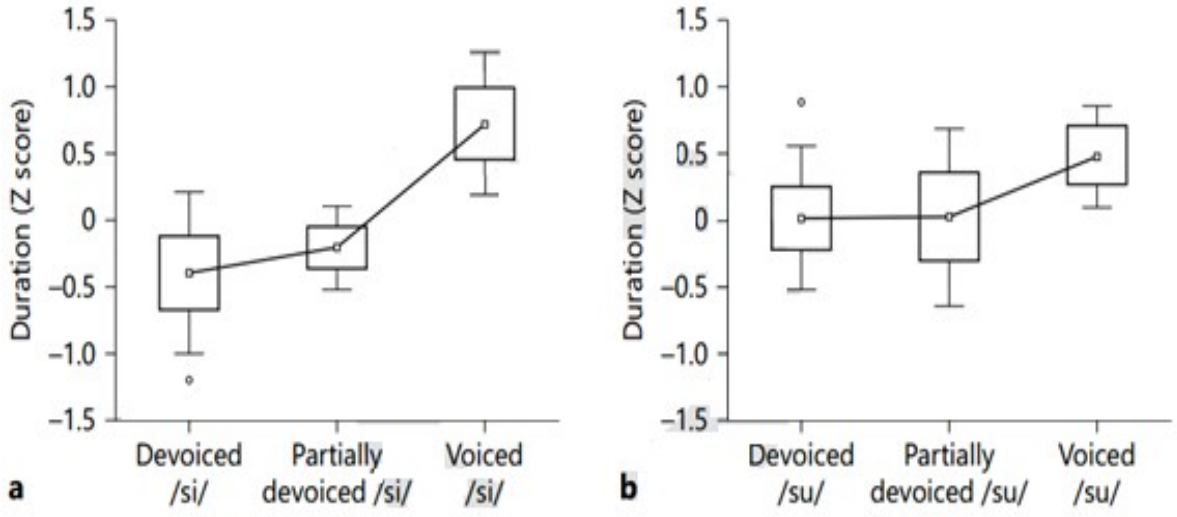


Figura 7 – Duração da sílaba com vogal vozeada, parcialmente desvozeada e totalmente desvozeada
Fonte: Meneses (2012, p. 55-56)

Para /i/, novamente, o teste de Kruskal-Wallis indica que as durações médias das sílabas diferem ($H = 6,95$, d.f. $= 2$, $p < 0,03$). Para /u/, não houve diferença significativa, apesar da diferença aparente no painel à direita da Figura 7. Isso se deve ao tamanho pequeno da amostra, levando-se em conta a baixa taxa de vozeamento de /u/.

A mudança da duração também é observada nas oclusivas. Albano & Meneses (2015), em um estudo preliminar, testaram o desvozeamento em vogais precedidas por oclusivas. Os resultados preliminares mostraram a presença de um *burst* mais longo na ausência de uma vogal visível. Assim, os padrões são aparentemente semelhantes às fricativas seguintes.

As Figuras 6 e 7 apresentadas anteriormente ilustram a tendência de vogais e consoantes se alongarem ou encurtarem dependendo da quantidade de vozeamento. Isso é confirmado na Figura 8 a seguir, que representa graficamente a correlação entre as durações de consoantes e vogais em casos nos quais o desvozeamento não era completo. A correlação é negativa e significativa (r de Pearson $(45) = -0,46$, $p = 0,001$), confirmando que a duração da vogal e da consoante são inversamente proporcionais.

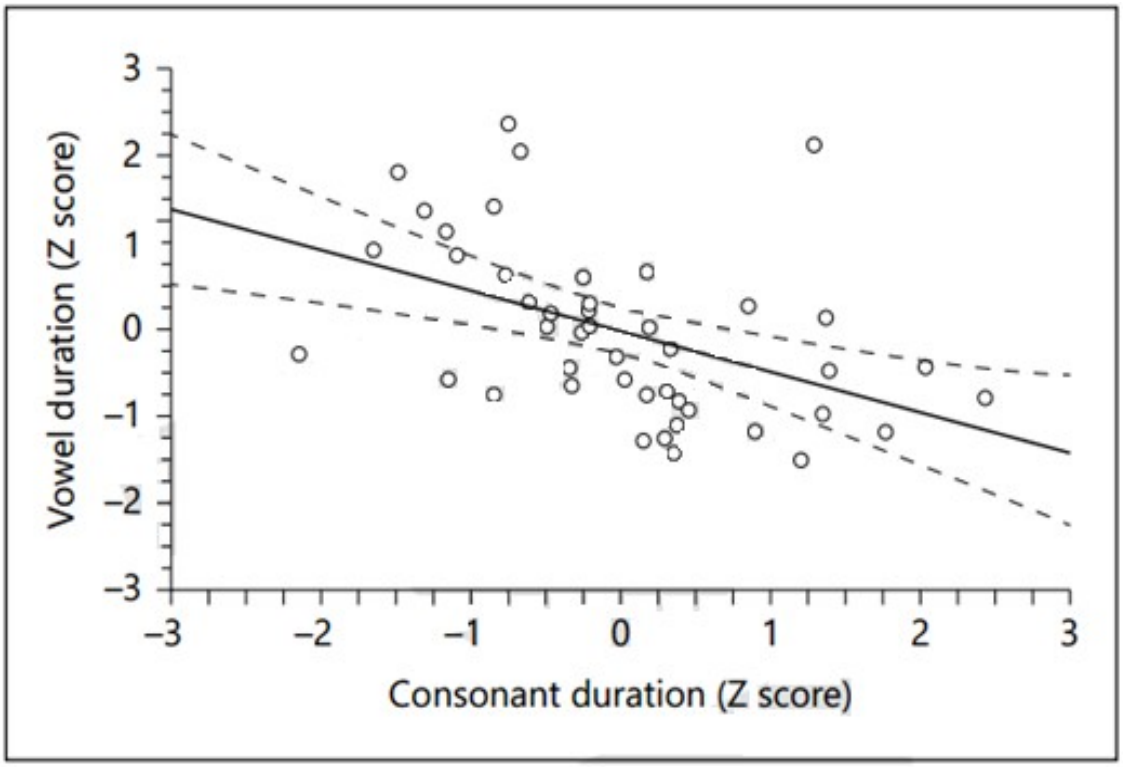


Figura 8 – Diagrama de dispersão entre a duração da vogal e da consoante normalizada pelo *z-score*[17]
Fonte: Meneses & Albano (2015, p.10)

Fica claro que, se de fato a vogal sofresse um processo de apócope, o esperado seria que o /s/ não apresentasse diferenças de duração nas três condições analisadas. Parece, na verdade, que o desvozeamento é resultado de uma redução do tempo da parte oral da sílaba, em vez da sílaba como um todo. O aumento na coarticulação observado em ocorrências de desvozeamento, como mostram os dados de duração, indica que a sobreposição articulatória de consoantes sobre vogais pode ser a principal causa desse encurtamento da vogal e consequente alongamento da consoante, resultando no desvozeamento vocálico.

Até aqui, parece implausível falar na queda da vogal nas condições analisadas; os resultados respaldam essa conclusão, uma vez que a queda implicaria inexistência do gesto vocálico. O desvozeamento, assim, seria o último estágio de uma escala de redução da magnitude vocálica.

[17] Esse dado específico não fazia parte dos resultados de Meneses (2012), mas é consequência da referida dissertação e está reportado em Meneses & Albano (2015).

2.3 Discussões

Em resumo, os achados expostos na seção anterior se referem a três pontos fundamentais:

(i) o alongamento do ruído em sílabas de vogais desvozeadas;

(ii) o deslocamento do centroide, que diferencia sílabas com presença de vogal vozeada daquelas com vogal desvozeada;

(iii) a tendência à redução das vogais (redução da magnitude) em contexto favorecedor de desvozeamento vocálico.

Os resultados da seção 2.2 mostram que os dados acústicos de vogais finais radicalmente reduzidas no PB contradizem a literatura tradicional, mostrando que não há apagamento nessa posição. Na verdade, eles mostram que a redução de vogais pós-tônicas finais do PB varia de moderada a extrema, o que implica um grau variável de desvozeamento e, na medida em que se aproxima a redução crítica, implica desvozeamento total. Essa reflexão já está presente em Meneses (2012). O que não aparecia no texto supracitado é a ideia de que esse tipo de variação, recorrente entre os indivíduos, parece ser a matéria-prima para a mudança de som em progresso, como concebido por Ohala (1981). Para o autor, variações sincrônicas compõem um "conjunto de variações"[18] que podem desencadear uma mudança sonora baseada em um lapso auditivo dos ouvintes. Se a variante da vogal desvozeada prevalecer entre os falantes, o caminho tende a ser o mesmo que os dados históricos mostram: a queda das vogais.

Aos olhos do modelo da FonGest, na produção da vogal vozeada, passando pela vogal parcialmente desvozeada até a vogal desvozeada, há um declínio da duração, uma redução do centroide e uma redução do espaço vocálico, que revelam o mecanismo escalar de redução da magnitude dos gestos articulatórios. A vogal desvozeada é, então, o último estágio da redução vocálica na posição tônica do PB.

[18] Do inglês, "*pool of variation*".

Outros fatores dos resultados acima merecem destaque: quando precedidas de /s/ e seguidas por uma consoante desvozeada, vogais vozeadas são muito curtas e altamente centralizadas, como indicam a diminuição da duração da vogal e do VSA. Quando a vogal está totalmente desvozeada, o gesto consonantal domina. Essas vogais tendem a encurtar gradualmente e deslizar sob a consoante precedente. Duas pistas acústicas apontam para a presença de gestos vocálicos: em primeiro lugar, um centroide reduzido (ii) sinaliza a vogal sobreposta; em segundo lugar, o alongamento da consoante (i) no desvozeamento total aponta para a radicalidade de tal sobreposição.

O resultado de (i) – a duração – é o dado mais inquietante entre os citados. Sobre ele, Meneses (2012, p. 69) afirma que na apócope:

> o esperado seria que o /s/ não apresentasse diferenças de duração nas três condições analisadas. Parece, na verdade, que ensurdecimento é resultado de uma redução temporal da parte oral da sílaba, em vez de a sílaba como um todo. O aumento observado na coarticulação em ocorrências de desvozeamento, como os dados de duração mostram, indica que a sobreposição articulatória de C sobre V pode ser a principal causa deste encurtamento da vogal.

É verdade e parece claro que o encurtamento da vogal é causado pelo aumento da sobreposição. Se voltarmos à Figura 8, veremos vogais e consoantes se alongando e encurtando proporcionalmente. No entanto, só a sobreposição não explica exatamente o alongamento da consoante no nível mais crítico de redução: o desvozeamento total. Duas hipóteses nos parecem plausíveis aqui: o alongamento da consoante resulta da manutenção (simulação) da duração da sílaba ou o ruído carrega uma porção vocálica. É com essa discussão que iniciaremos o próximo capítulo.

Capítulo 3

"*Novas perspectivas sobre o desvozeamento vocálico no PB*"

3.0 Alongamento compensatório ou um resquício de vogal?

Os dados revistos no capítulo anterior apontam para duas hipóteses básicas: (H1) há um alongamento compensatório da fricativa; (H2) existe uma porção vocálica análoga ao ruído da fricativa, uma espécie de vogal fricatizada (esse tipo de vogal mais ruidosa já foi investigada, por exemplo, por Feng, 2007).

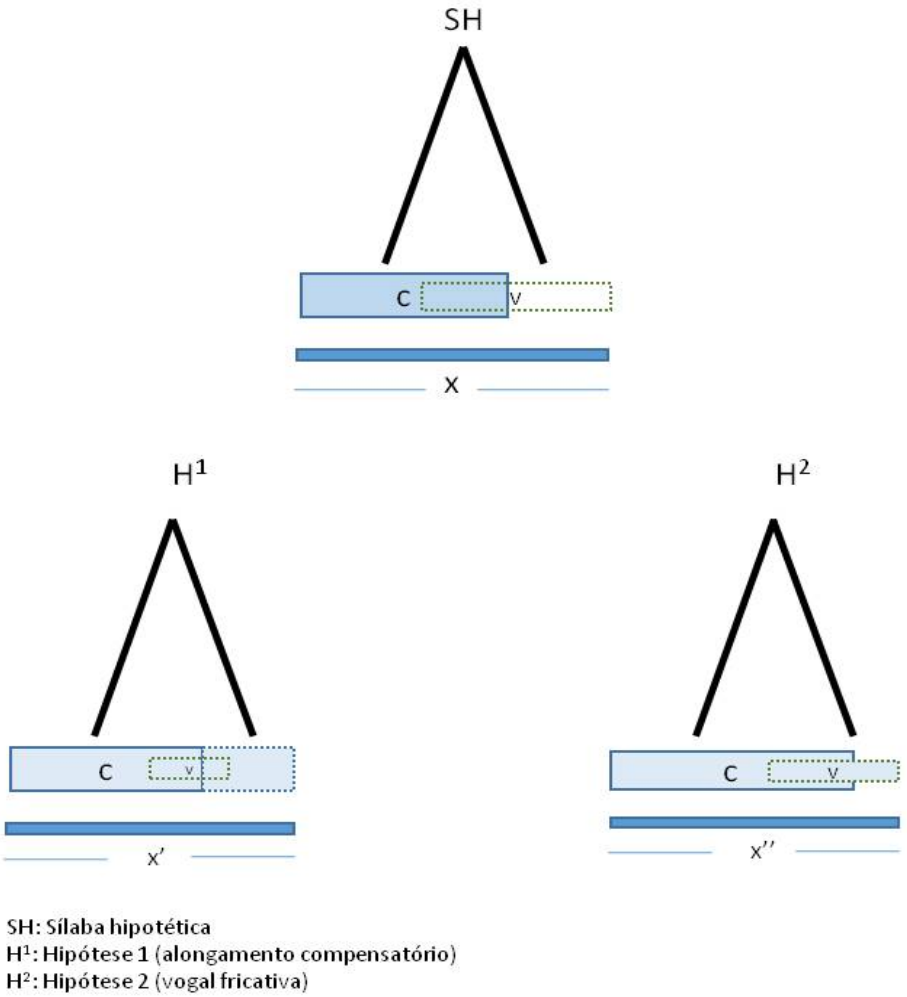


Figura 9 – Esboço de representação das hipóteses levantadas a partir dos resultados do Experimento 1[1]

[1] Na Figura 9, SH representa uma sílaba padrão com a sobreposição entre consoante e vogal; H1 mostra o alongamento compensatório e a redução significativa das pistas vocálicas; e H2 representa a sequência consoante-vogal, de modo que a vogal tem características consonantais.

A Figura 9 acima é uma tentativa de representação das hipóteses compatíveis com os resultados do Capítulo 2.

Na primeira hipótese, o alongamento do ruído do /s/ com desvozeamento completo é interpretado literalmente. Lembremos que a duração do ruído do /s/ de vogais não vozeadas é significativamente diferente do que ocorre com vogais parcialmente vozeadas ou vozeadas (rever Figura 8 no capítulo anterior).

Levando em consideração H1, é possível esboçar uma interpretação teórica a partir dos dados de Byrd & Saltzman (1998) e Saltzman *et al*. (2008). Esses estudos abordam como o contexto prosódico modula a dinâmica da fala para dentro de uma hierarquia de ritmo que é postulada através da adição de várias camadas prosódicas de acordo com Browman & Goldstein (1990) para a FonGest.

Segundo os autores, mudanças relacionadas à constrição, como o nosso processo de redução, poderiam ser explicadas por perturbações e efeitos nos osciladores de planejamento subjacentes à estrutura silábica – os "*clocks*" (cf. Goldstein, Chitoran & Selkirk, 2007; Goldstein *et al*. 2009; Nam, Goldstein & Saltzman, 2009). Aqui, os padrões de duração consonantal podem ser concebidos como um efeito prosódico extremo resultante da falta do acento e da posição fraca. Assim, enquanto o oscilador da vogal perde a força de seu acoplamento com o oscilador da sílaba, o acoplamento com o oscilador da consoante é reforçado. Além disso, a vogal desvozeada quebra o contorno de sonoridade, uma vez que a vogal perde a sua periodicidade devido à sobreposição. Então, a consoante é alongada para compensar essa perda e manter silabicidade. O próximo passo seria a ressilabificação do /s/ com a sílaba tônica anterior.

Em resumo, a hipótese é que a consoante se torna maior em magnitude e aumenta a sua duração para preservar a força da sílaba, uma vez que o gesto vocálico está inteiramente mascarado. Esse tipo de alongamento da consoante pode ser um tipo de modulação local e temporal, melhor explicado pela **desaceleração** do gesto, possivelmente em reação à redução extrema da vogal, como podemos observar na representação de H1 na Figura 9.

Vale ressaltar que a hipótese de um alongamento compensatório também foi levantada por Whang (2014), baseada nos trabalhos de Kondo (2005), que mostram que a redução da vogal é acompanhada por uma duração maior do *burst*/ruído (/s, k, t/) da consoante. No entanto, pesa contra a hipótese H1 o fato de, no melhor de nosso

conhecimento, o alongamento compensatório de consoante ser um fenômeno raro (ou inexistente[2]) entre as línguas naturais, segundo a fonologia tradicional. O alongamento compensatório é um tipo de processo fonológico sincrônico que ocorre quando um segmento, geralmente na coda silábica, é apagado (ou desocupa a posição em uma mora) e é substituído pelo alongamento do segmento anterior (Hayes, 1989). Em grego antigo, por exemplo, quando um /s/, em posição de coda, é apagado, a mora sobrevive pelo alongamento de um segmento vizinho, a vogal precedente na maioria dos casos.

Em contrapartida, tomando H2 como verdadeira, os dados de duração podem sugerir que o ruído alongado é, na verdade, uma porção final de vogal que é demasiadamente curta para retomar o vozeamento e se parece com uma vogal fricatizada, diante de uma **aceleração** na produção da vogal. A vogal mantém a silabicidade, mas perde periodicidade e ganha uma "cara" mais ruidosa. Vogais fricativas são relatadas em muitas línguas e são caracterizadas como mais ruidosas do que uma vogal típica (Feng, 2007). Aqui, a vogal fricatizada pode ser mais um passo no processo de redução, muito provavelmente, o último antes da queda da vogal.

Dados da literatura revelam uma tendência de a redução articulatória e a sobreposição gestual afetarem segmentos de diferentes maneiras (Byrd, 1996; Recasens & Mira, 2014). A alta complexidade articulatória envolvida na produção das vogais finais pode explicar por que elas podem sofrer redução gestual considerável e misturarem-se radicalmente com as fricativas. Esse cenário nos permite indagar se uma vogal que não deixa rastros acústicos tão óbvios e não é percebida no sinal visual (ou seja, /i, u/ em nosso estudo) pode ainda ser realizada foneticamente e, portanto, deixar um resíduo gestual no sinal articulatório. Exemplos de gestos perceptualmente assimilados, mas, mesmo assim, presentes no sinal acústico, já foram reportados na literatura, como a alveolar nasal (Browman & Goldstein, 1989, 1992). Nesse sentido, H2 mostra-se bastante plausível.

A seguir, o Quadro 2 apresenta um resumo das hipóteses desenvolvidas até aqui.

[2] Em comunicação pessoal, o pesquisador Didier Demolin apontou a existência de línguas com alongamento compensatório de consoantes.

Quadro 2 – Quadro-resumo das hipóteses relacionadas à H1 e H2

H1	H2
Espalhamento da consoante (alongamento compensatório)	Vogal assimila características da consoante (vogal fricatizada)
Consoante simula duração silábica	Vogal mantém duração e posição silábica
Pistas vocálicas precárias (somente acústicas)	Pistas vocálicas mais robustas (além da acústica)

Ante o exposto, ambas as hipóteses encontram embasamento nos dados acústicos e merecem a devida investigação. Assim, as seções seguintes apresentam os resultados de três experimentos que investigam especificamente o alongamento do ruído atestado no Experimento 1: no primeiro, apresentamos nossos dados aerodinâmicos que podem mostrar pistas do que de fato acontece na porção final da sílaba desvozeada; depois, investigamos se o ruído tem alguma influência sobre a detecção da vogal em casos de desvozeamento total. Os dados que serão apresentados a seguir podem iluminar nossas interpretações e indicar qual das hipóteses acima parece estar mais relacionada à articulação das vogais desvozeadas no PB.

3.1 Experimento II: dados aerodinâmicos

3.1.1 Participantes

O teste aerodinâmico foi realizado com três dos participantes (chamaremos aqui $S1_2$, $S2_2$, $S3_2$) do experimento anterior. Foram gravadas tríades de palavras que podem ser vistas no Quadro 3 a seguir. Novamente, apenas participantes do sexo feminino foram gravadas para manter a comparabilidade com o experimento anterior e os demais experimentos. Cada sujeito leu as frases-alvo três vezes. Todas as palavras-alvo foram gravadas com as frases veículos “Digo X paciente” e “Digo X baixinho”.

Quadro 3 – Tríades de palavras usadas no Experimento 2

	/i/		/u/		Sem vogal	
Palavras	ace	/'a.si/	Aço	/'a.su/	ás	/'as/
	canse	/'kan.si/	Canso	/'kan.su/	cãs	/'kans/
	coce	/'kɔ.si/	Coço	/'kɔ.su/	cós	/'kɔs/
	face	/'fa.si/	Faço	/'fa.su/	faz	/'fas/
	lance	/'lan.si/	Lanço	/'lan.su/	lãs	/'lãs/
	passe	/'pa.si/	Passo	/'pa.su/	paz	/'pas/
	trace	/'tra.si/	Traço	/'tra.su/	traz	/'tras/

3.1.2 Materiais

Para a captação dos dados, usamos a estação de trabalho portátil EVA2, que faz medidas aerodinâmicas e acústicas (Ghio & Teston, 2004). O EVA2 possui dois canais de fluxo de ar e dois canais de pressão estão disponíveis para medir o fluxo de ar nasal e oral junto com a pressão da faringe. Na Figura 10, as setas mostram o canal de captação do fluxo de ar e o microfone. Na figura à esquerda, é possível ver a máscara acoplada ao aparelho. No experimento, apenas o canal de fluxo de ar oral foi utilizado.

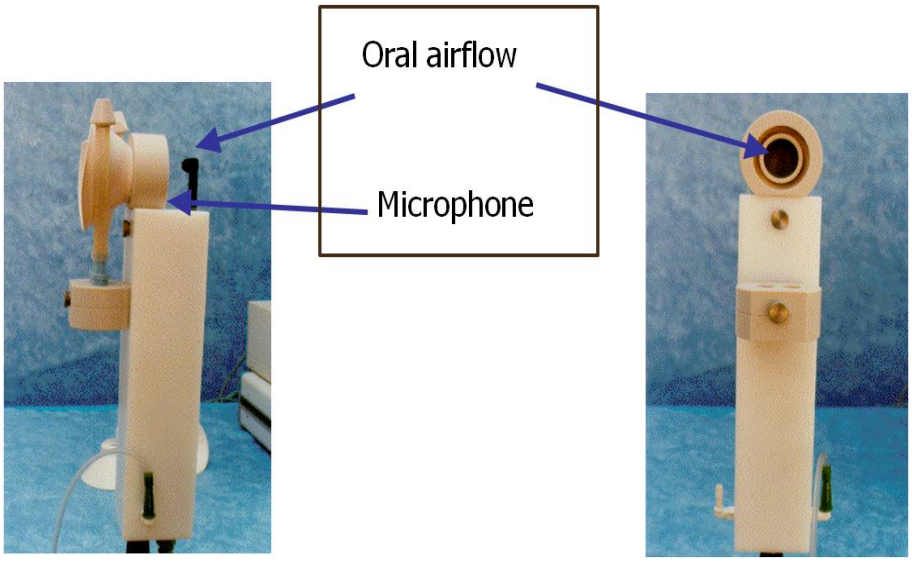


Figura 10 – Imagem do EVA2 e seus canais de captação

A gravação simultânea do som e do fluxo de ar exige a utilização de um "bocal" ligado ao suporte mecânico do microfone. Uma máscara de silicone flexível foi utilizada para vedar qualquer escapamento oral de ar, de modo a obter as medições mais fiáveis possíveis do fluxo de ar por via oral. Foi para esse fim que os sensores de pressão foram manualmente calibrados e foram verificados para cada sujeito e entre todas as repetições. O software Phonedit foi usado para a gravação e também para a análise e processamento de dados.

3.1.3 Resultados

A análise dos sinais sincronizados da forma de onda e do fluxo de ar mostra que o fluxo oral da porção /s/ + vogal vozeada apresenta dois picos de fluxo de ar, tal como mostrado na Figura 11 (à esquerda). O primeiro pico refere-se à abertura da glote. Depois, é possível observar um vale que é correlato da liberação contínua do fluxo durante a produção da fricativa. Por fim, o segundo pico refere-se ao abrandamento da constrição e abaixamento da língua para o início da produção a vogal.

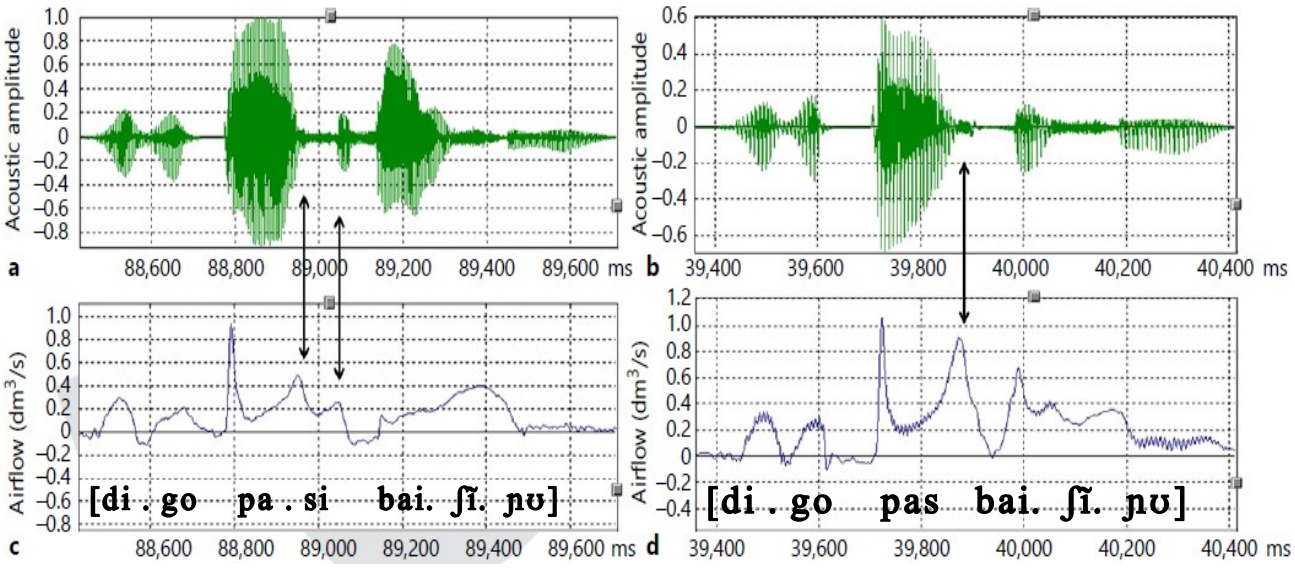


Figura 11 – Forma de onda (acima) e fluxo de ar oral (abaixo) de /s/ com vogal vozeada em "passe" (à esquerda) e /s/ de coda em "paz" (à direita). Em **c**, é possível perceber dois picos de fluxo de ar na realização da sílaba /si/, enquanto em **d** há apenas um pico de fluxo de ar na produção do /s/.

No painel à direita da Figura 11, no /s/ de coda há apenas um pico de fluxo de ar oral, que está relacionado com a abertura da glote na produção de fricativas. Depois, há uma

queda abrupta do fluxo de ar oral, uma vez que o /s/ está na posição de coda (isto é, não há uma vogal seguinte). Se compararmos a Figura 11 e a Figura 12 a seguir, vemos que há uma clara semelhança entre as vogais vozeadas e vogais desvozeadas: o ruído que ocorre nos casos de vogais desvozeadas também apresenta, em primeiro lugar, um pico de abertura da glote e, em segundo lugar, um segundo pico análogo ao mostrado na Figura 11 (c). Por fim, há um contraste com o /s/ de coda da Figura 11 (d), que tem apenas um pico de fluxo de ar.

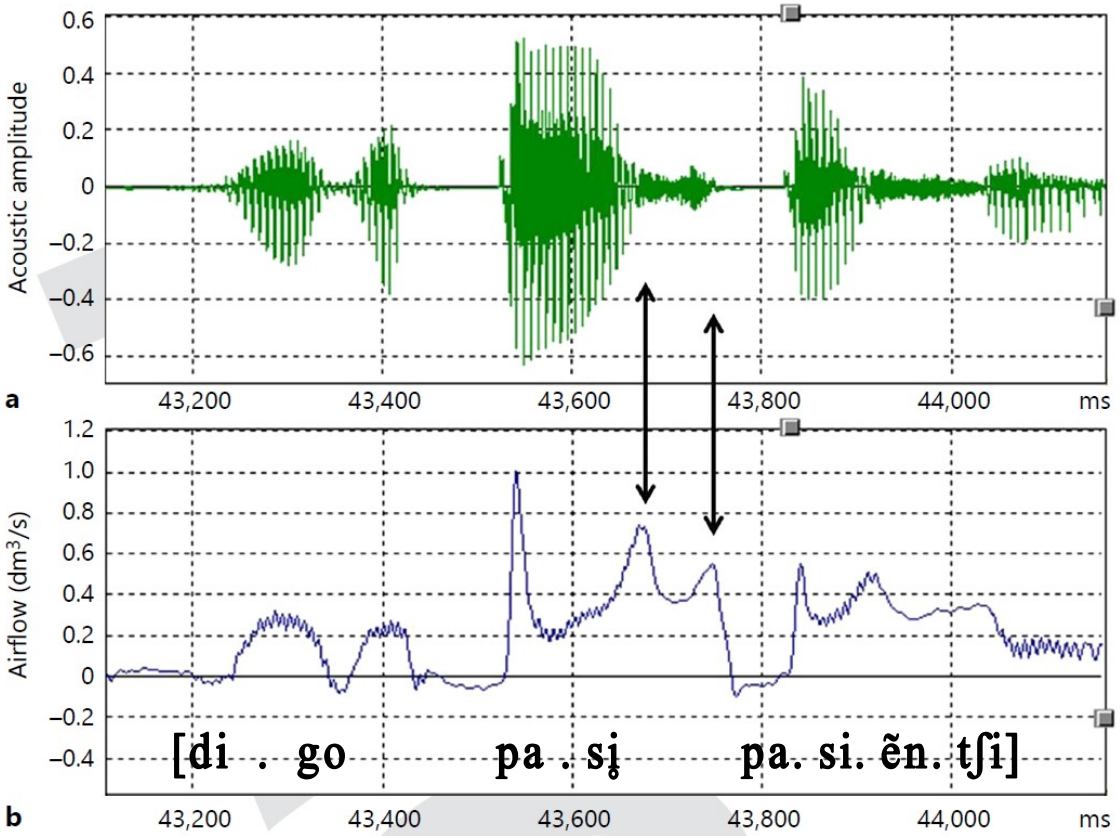


Figura 12 – Forma de onda (acima) e fluxo de ar oral (abaixo) de /s/ com vogal desvozeada

Tomando os dados expostos na Figura 12 como referência, sugerimos que nem o tempo e nem a corrente de ar são suficientes para a produção de uma vogal vozeada canônica quando essa é seguida por uma consoante desvozeada. Aparentemente, o vozeamento é inibido pelo domínio do gesto da fricativa, como previsto em H2, na qual há uma mistura entre os gestos da vogal e da consoante.

3.1.4 Discussão

Os dados aerodinâmicos do Experimento 2 nos dão mais evidências sobre a extrema sobreposição entre consoante e vogal. Ao retomarmos a representação da Figura 9 e acrescentarmos um esboço dos dados aerodinâmicos, temos o seguinte esquema apresentado a seguir na Figura 13.

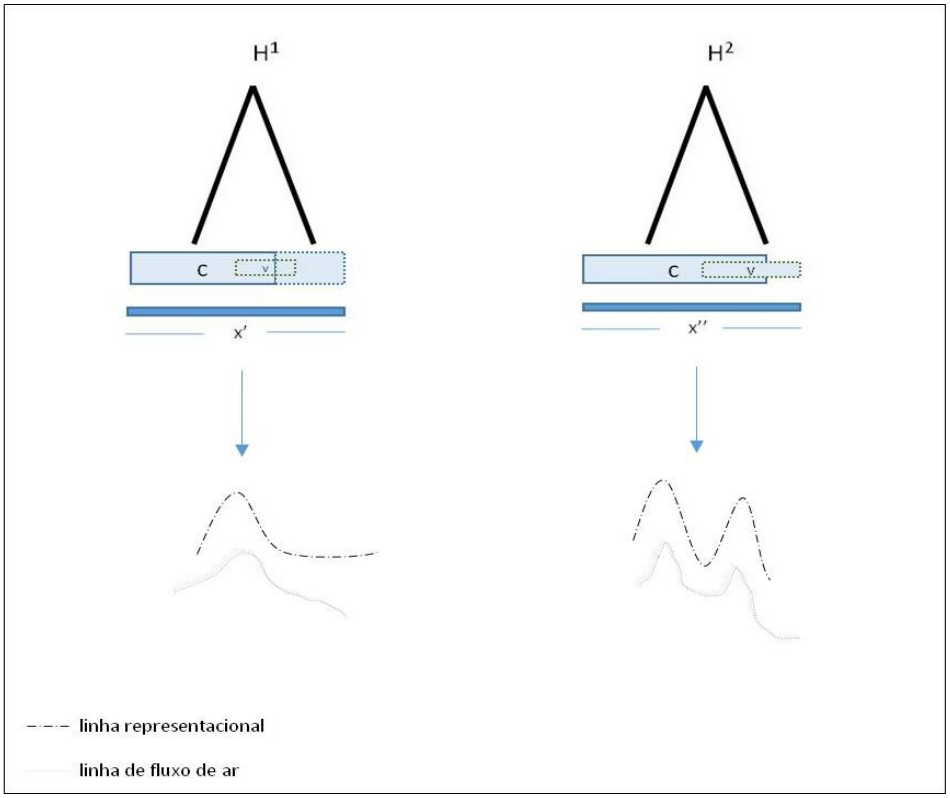


Figura 13 – Esboço de representação das hipóteses e do fluxo de ar obtidos no Experimento 2. A linha pontilhada em H1 representa o fluxo de ar esperando no caso do alongamento compensatório; em H2, a linha representacional mostra o fluxo de ar encontrado (mais relacionado com a existência de uma porção vocálica no ruído).

Se o alongamento compensatório fosse resultado de uma emulação da fricativa, como postulado em H1, o resultado do fluxo de ar deveria ser semelhante ao encontrado nas fricativas de coda: um pico resultante da constrição fricativa e depois a queda e estabilização do fluxo do resultado da fricção, como representado acima. Uma vez que esse não tenha sido o comportamento do fluxo de ar, H2 parece, portanto, mais respaldada pelos resultados aerodinâmicos. Os dados mostram que o fluxo de ar da porção final da constrição é consistente com a presença de uma porção vocálica no fim do ruído (H2). Dado que as sílabas com vogais desvozeadas apresentam dois picos de fluxo de ar, assim como as sílabas com

vogais vozeadas, parece mais consistente afirmar que temos sim uma porção vocálica que, por conta das características restritivas da posição átona, não retoma o vozeamento.

Possivelmente, isso se dá porque a configuração do trato vocal para vogais desvozeadas finais é mais estreita do que em vogais vozeadas, uma vez que não há tempo para alcançar a abertura necessária para uma queda significativa na pressão supraglótica.

Esses estudos levantam questões sobre à percepção: (i) se as sílabas desvozeadas – alongadas – carregam em seu ruído uma porção vocálica, como elas são percebidas?; (ii) as vogais são indistintamente percebidas quando alongadas e com duração abaixo da média? As perguntas acima serão respondidas nos experimentos a seguir.

Capítulo 4

"*Sobre a percepção*"

4.0 Teste de percepção

4.1 Teste-piloto

Como visto no capítulo anterior, nossos resultados, até agora, apontavam para duas direções distintas: (i) há o alongamento aparente compensatório da fricativa; (ii) há uma porção vocálica dentro do ruído fricativo, uma vogal-fricativa. Essa última parece-nos mais respaldada pelos dados aerodinâmicos de fluxo de ar. Diante disso, a presente seção apresenta os resultados de dois experimentos de percepção que investigam se o "alongamento" do ruído do /s/ tem alguma influência sobre a detecção de vogal em casos de desvozeamento total.

O experimento-piloto foi desenhado com o objetivo de testar nossas hipóteses em um *design* mais simples, no qual o sujeito tivesse a oportunidade de dar respostas mais detalhadas do que duas categorias de acerto e erro, ou seja, os possíveis gradientes do processo foram levados em consideração. Nossa questão no teste-piloto era básica: há diferença na percepção de vogais desvozeadas e parcialmente desvozeadas? É importante ressaltar que esses dois tipos de vogais empregados nos estímulos do teste foram dois grupos previamente delimitamos já no experimento de produção. Lembramos, então, que o desvozeamento foi classificado como parcial quando a vogal sobreposta pelo consoante teve uma porção periódica curta com alguns pulsos glóticos (média = 4) e duração reduzida (média = 20 ms). A vogal foi classificada como completamente desvozeada quando apresentou uma porção completamente aperiódica sem pulso glótico e com configuração formântica totalmente indefinida.

Sendo assim, antes da realização do teste, esperava-se que as parcialmente desvozeadas tivessem as maiores taxas de reconhecimento e que as desvozeadas totais tivessem taxas de reconhecimento, obviamente, menores. A metodologia e os resultados estão dispostos nas subseções seguintes.

4.1.1 Participantes e materiais

Quatro ouvintes do sexo feminino ($S1_3$ ... $S4_3$), falantes de dialetos que apresentam /s/ em posição de coda e desvozeamento, foram selecionadas para este experimento. Mais uma vez, foram mantidos constantes a faixa etária (25-40 anos), a escolaridade (estudantes de graduação, neste caso) e o nível socioeconômico (classe média).

Para o teste propriamente dito, pares de palavras com vogais parcial e completamente desvozeadas foram gravadas. As palavras selecionadas são pares mínimos separados pela existência ou não da vogal final e seguindo as estimativas de pulsos e duração já referidas acima, como arrolado no Quadro 4 abaixo.

Quadro 4 – Lista de palavras usadas no experimento-piloto

	/i/		**/u/**	
Palavra	ace	/'a.si/	aço	/'a.su/
	canse	/'kan.si/	canso	/'kan.su/
	coce	/'kɔ.si/	coço	/'kɔ.su/
	face	/'fa.si/	faço	/'fa.su/
	lance	/'lan.si/	lanço	/'lan.su/
	passe	/'pa.si/	passo	/'pa.su/
	trace	/'tra.si/	traço	/'tra.su/

4.1.2 Procedimento

Os dois grupos de estímulos – palavras com vogais desvozeadas e parcialmente desvozeadas – foram extraídos de nossas sentenças-alvo originais e tocados de forma isolada do contexto. Para o teste piloto, foram realizadas 3 repetições e um total de 14 estímulos.. Assim, foram analisados 168 dados (14 estímulos x 4 sujeitos x 3 repetições) para este experimento piloto.

Cada sujeito foi instruído a ouvir e escolher em uma escala de 1 a 5 pontos, no quais 1 (sem vogal) e 5 (com vogal) seriam os extremos de escolha. Inicialmente, consideraríamos os 5 pontos de análise. No entanto, a dificuldade de diferenciar pontos próximos fez com que os pontos 1 e 2 fossem considerados como “sem vogal”, o ponto 3

fosse considerado intermediário (dúvida) e os pontos 4 e 5 fossem considerados “com vogal”. A escala criada pode ser vista na Figura 14 a seguir.

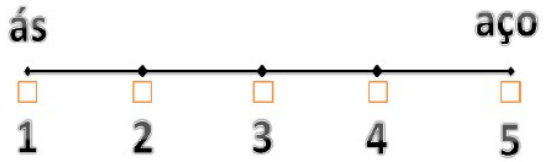


Figura 14 – Escala utilizada pelos informantes para o experimento-piloto

4.1.3 Medidas

As medidas obtidas nesse experimento foram as seguintes: (i) a porcentagem dos escores dados por cada sujeito (geral e por sujeito); e (ii) o tempo de reação. A correlação entre classificação do sujeito e o tempo de reação foi utilizada para estabelecer um limiar para a presença de dúvida na escolha do sujeito.

4.1.4 Análise estatística

Para averiguar relação entre a taxa de acerto de cada escolha e o tempo de reação de cada sujeito, além da estatística descritiva dos dados, a correlação de Spearman foi utilizada.

4.1.5 Resultados

Os dados apresentados na Figura 15 a seguir mostram os resultados dos dois tipos de estímulos aos quais os sujeitos foram submetidos. Como podemos ver, apenas desvozeamento extremo perturba a percepção, uma vez que o desvozeamento parcial é suficiente para a recuperação da vogal em 100% dos casos com ótimo percentual de certeza. Como esperado, os sujeitos variaram muito na tarefa de escolha quando ouviram palavras com vogais desvozeados. As vogais totalmente desvozeadas levaram a uma taxa de erro de identificação de 70%.

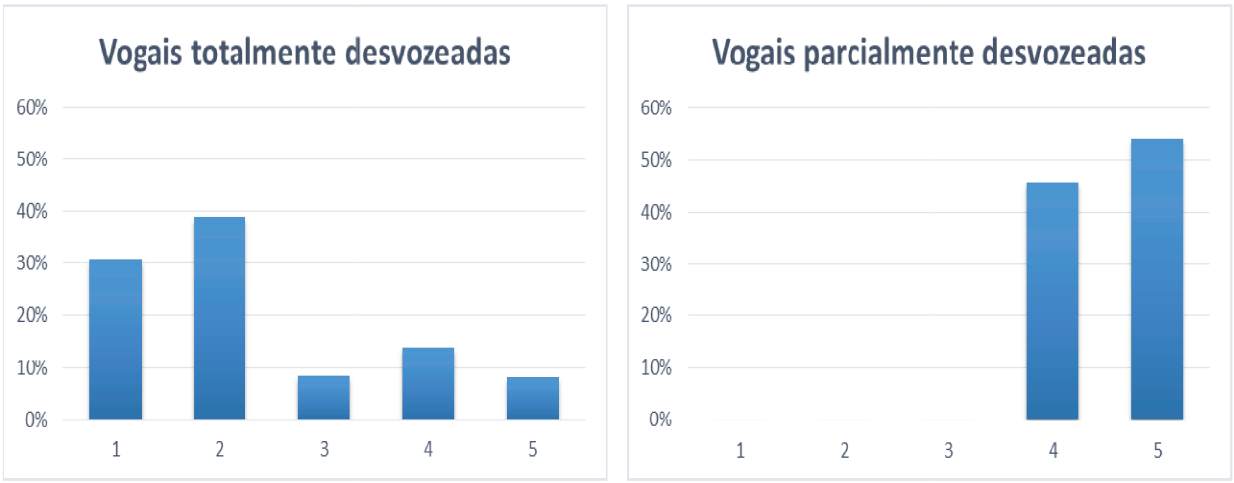


Figura 15 – Porcentagem de cada escolha em um teste de identificação forçada para vogais desvozeadas e parcialmente desvozeadas

Se olharmos os sujeitos do teste-piloto separadamente, focando os casos de desvozeamento completo, é possível perceber que S2 e S3 são aqueles que mais identificam as vogais desvozeadas apresentadas no experimento. S1 e S4, por outro lado, têm as taxas de erro mais altas.

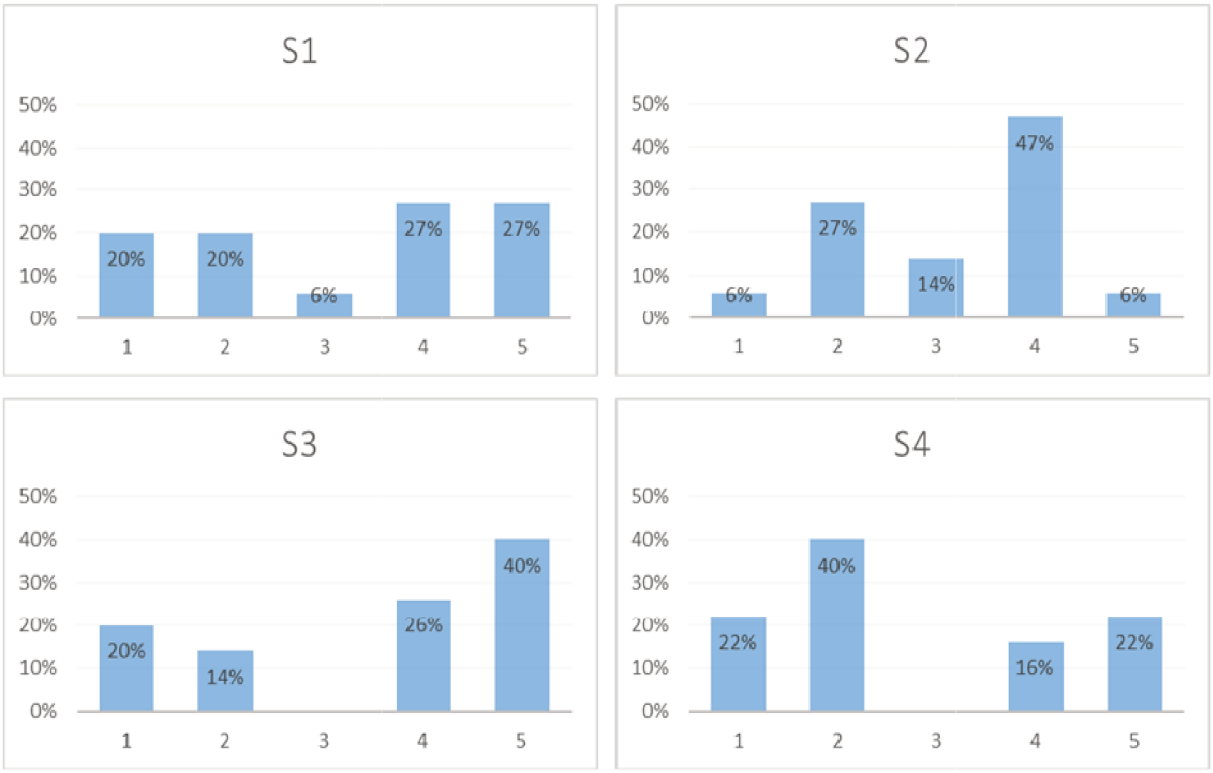


Figura 16 – Porcentagem de cada escolha dos sujeitos na identificação das vogais desvozeadas em cada ponto da escala. 1 e 2 indicam escolha pelo /s/ de coda; 3, dúvida; 4 e 5, escolha pela vogal.

A seguir, na Figura 17, apresentamos os dados da correlação (*Spearman correlation test*) entre o tempo de reação e a classificação dada por cada sujeito.

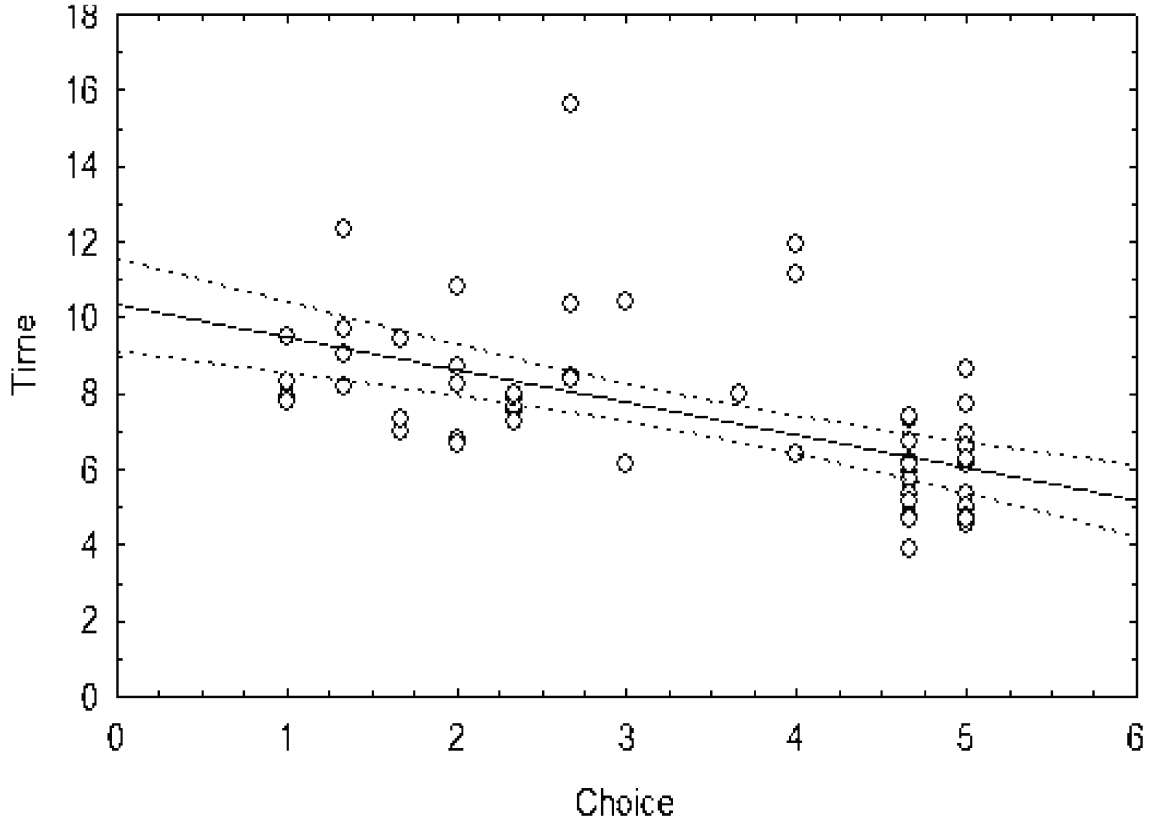


Figura 17 – Diagrama de dispersão entre as médias das escolhas e o tempo de reação (R (58) = -0,64, p < 0,0001)

O resultado da correlação (**R (58) = -0,64, p < 0,0001**) mostra que a relação entre os escores e o tempo de reação é proporcionalmente inversa, ou seja, os indivíduos levam mais tempo para identificar os estímulos com vogais desvozeadas, se comparados com os estímulos de vogais parcialmente desvozeadas.

A partir dos resultados da Figura 15, a correlação apresentada na Figura 17 já era esperada, uma vez que o mínimo de vozeamento já é suficiente para a identificação das vogais.

4.1.6 Discussão dos resultados (e falhas do experimento)

Diante dos objetivos, isto é, investigar a percepção das vogais desvozeadas e dar aos sujeitos uma tarefa na qual os gradientes de produção fossem levados em consideração, os resultados do teste-piloto estiveram dentro do esperado. Quatro resultados devem ser destacados: (i) as classificações das vogais totalmente desvozeadas são bem variáveis; (ii) há

um viés definido por palavras sem vogais (aproximadamente 70% dos casos); (ii) o vozeamento parcial é suficiente para a recuperação da vogal (100% dos casos); (iv) o desvozeamento completo leva a erros de identificação.

O teste deixa claro que o vozeamento, mesmo que mínimo, como no caso das parcialmente desvozeadas, é suficiente para o reconhecimento das vogais átonas finais.

Das variáveis analisadas, percebemos que o ponto 3 da escala de escores, aquele que significaria mais dúvida, foi preterido pelos sujeitos. Esse escore, aparentemente, não foi considerado relevante para uma parte dos sujeitos. Além disso, como foi dito na metodologia, o escore de 5 pontos foi, na realidade, reduzido a três: sem vogal (1), no meio (2) e com vogal (3). Sendo assim, escores em cinco pontos não transpareceram bem a escolha dos sujeitos.

Uma vez que o desvozeamento parcial não gera dúvidas nos sujeitos, o Experimento III visa a analisar o que acontece quando apresentamos somente palavras com vogais desvozeadas em três contextos. Para o próximo experimento, escores de acerto e erro foram levados em consideração e uma nova escala gradiente foi criada com o objetivo de medir a confiança dos sujeitos nas repostas dadas. Mais detalhes da metodologia e os resultados são apresentados na próxima seção.

4.2 Experimento III: tarefa de identificação forçada

4.2.1 Participantes

Foi realizado um teste de identificação forçada com seis participantes do sexo feminino ($P1_3$ ... $P6_3$). Mais uma vez, a faixa etária (25-40 anos), a escolaridade (estudantes de graduação, neste caso) e o nível socioeconômico (classe média) foram mantidos constantes.

4.2.2 Materiais

Os estímulos foram selecionados por meio dos seguintes passos: (i) tríades de palavras, como /fas/, /fa.si/ e /fa.su/ (ver Quadro 5), foram registradas por um sujeito com uma alta taxa de desvozeamento. Uma frase veículo foi criada para cada palavra-alvo; (ii) foram selecionadas ocorrências típicas de desvozeamento; (iii) a duração do /s/ precedente foi medida e normalizada por *z-score*; (iv) os *tokens* selecionados foram classificados em três categorias: alongado, com duração média (ou seja, *z-score* igual a zero) ou de coda.

Quadro 5 – Tríades de palavras usadas no Experimento 3

	/i/		/u/		Sem vogal	
Palavras	ace	/'a.si/	aço	/'a.su/	ás	/'as/
	canse	/'kan.si/	canso	/'kan.su/	cãs	/'kans/
	coce	/'kɔ.si/	coço	/'kɔ.su/	cós	/'kɔs/
	face	/'fa.si/	faço	/'fa.su/	faz	/'fas/
	lance	/'lan.si/	lanço	/'lan.su/	lãs	/'lãs/
	passe	/'pa.si/	passo	/'pa.su/	paz	/'pas/
	trace	/'tra.si/	traço	/'tra.su/	traz	/'tras/

4.2.3 Procedimento

O teste de percepção consistiu na apresentação de 35 estímulos para cada sujeito. Cada sujeito ouviu os estímulos cinco vezes. No fim, foram analisados 1.050 dados (35 estímulos x 6 sujeitos x 6 repetições) para este experimento. Os participantes foram convidados a avaliar as suas respostas, novamente em uma escala de cinco pontos, variando da máxima para a mínima confiança. A identificação incorreta da vogal (por exemplo, /pa.si/ para /pa.su/) foi contada como um sucesso, uma vez que o principal problema foi a detecção de vogal. Após a identificação, os indivíduos indicaram o nível de confiança de sua resposta, e essa foi usada para ponderar as pontuações. Assim, foi possível mensurar o gradiente de percepção entre os sujeitos.

4.2.4 Análise estatística

O teste de Kruskall-Wallis foi utilizado para avaliar a diferença entre as pontuações ponderadas nas três condições.

4.2.5 Resultados

A Figura 18 a seguir mostra as taxas de acerto e erro para todos os estímulos do teste. O /s/ alongado apresentou mais acertos, ou seja, 76%. Se compararmos o /s/ alongado com o /s/ médio, a taxa de erro foi relativamente baixa nas alongadas, com 24%. O /s/ médio

também teve mais acertos, mas apresentou também uma taxa de erros relevante, 42%. Apesar da confusão causada pela escolha forçada, o /s/ de coda tem uma maior taxa de sucesso do que de erros.

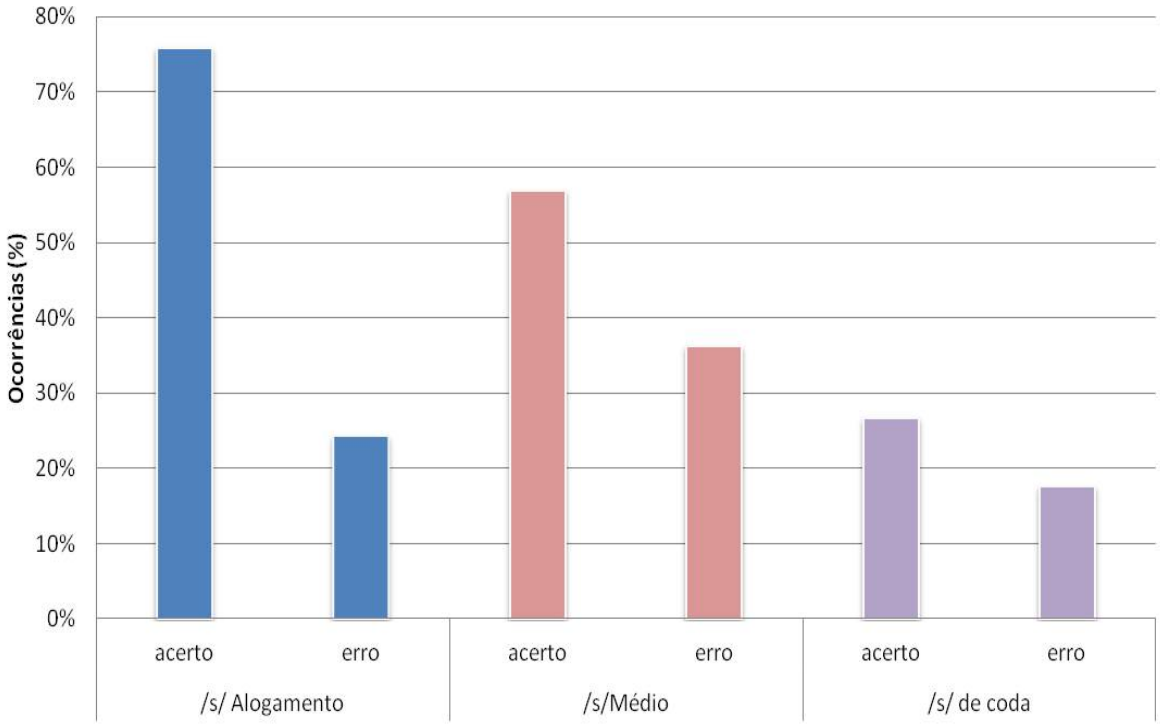


Figura 18 – Porcentagem de acertos e erros para a identificação de /s/ alongado, /s/ de duração média e /s/ de coda

A Figura 19 a seguir destaca a variação individual da percepção das vogais desvozeadas, que é análoga àquela encontrada no Experimento 1 e também pôde ser observada no nosso experimento de identificação. Aqui, os casos extremos são $P1_3$ e $P3_3$: $P1_3$ tem menos de 5% de erros para /s/ alongado; $P3_{3,}$ por outro lado, tem 15% de erros para as alongadas e mais de 20% para o /s/ médio. A diferença entre $P1_3$ e $P3_3$ sugere que há pessoas que são mais propensas a erro de percepção e mais sensíveis à redução extrema da vogal.

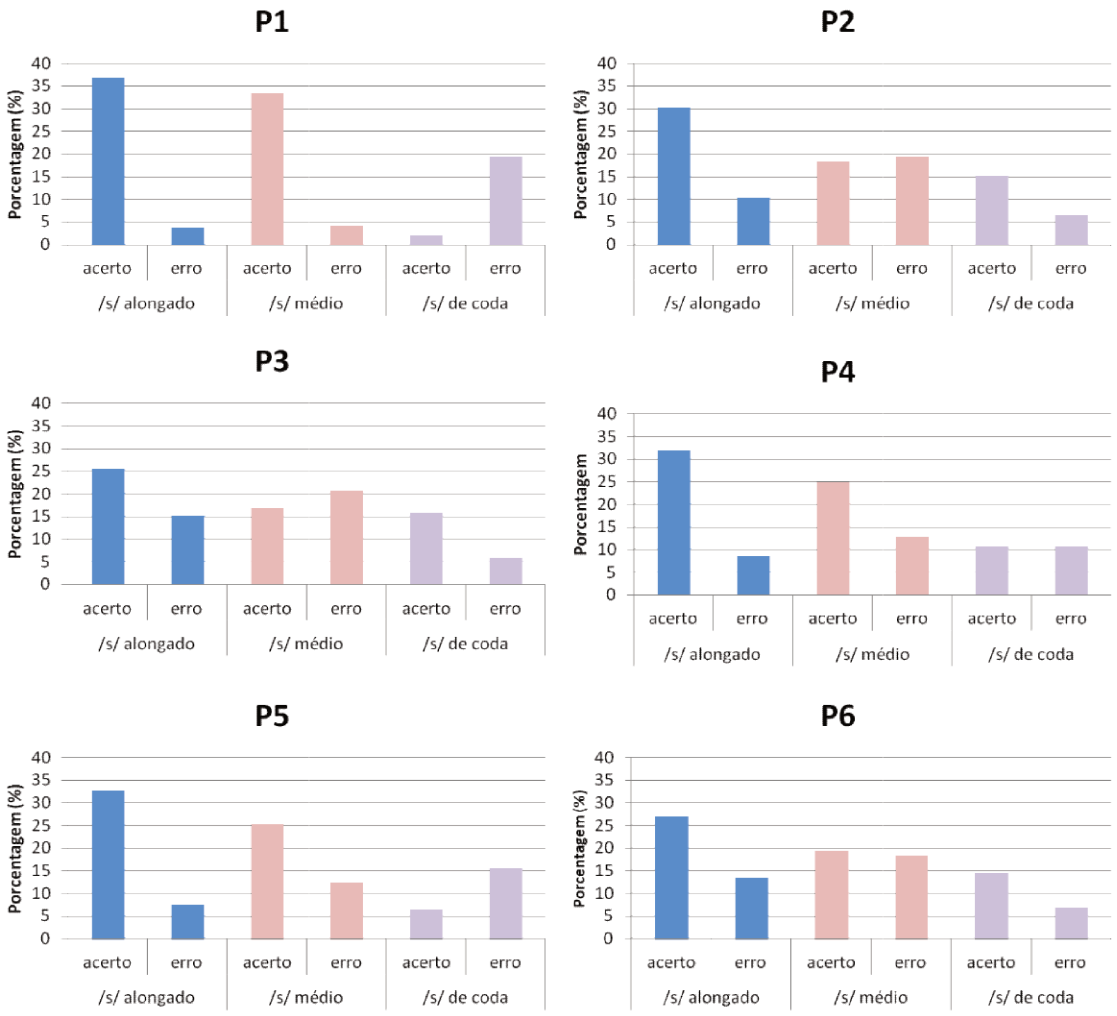


Figura 19 – Porcentagem de acertos e erros de identificação por sujeito

Quando consideramos as pontuações ponderadas na Figura 20 a seguir, ou seja, o resultado da ponderação de acertos e erros por nível de confiança, o teste de Kruskal-Wallis indica que o ruído alongado é uma pista decisiva para a detecção vogal ($H = 6{,}33$, $d.f. = 2$, $p < 0{,}04$).

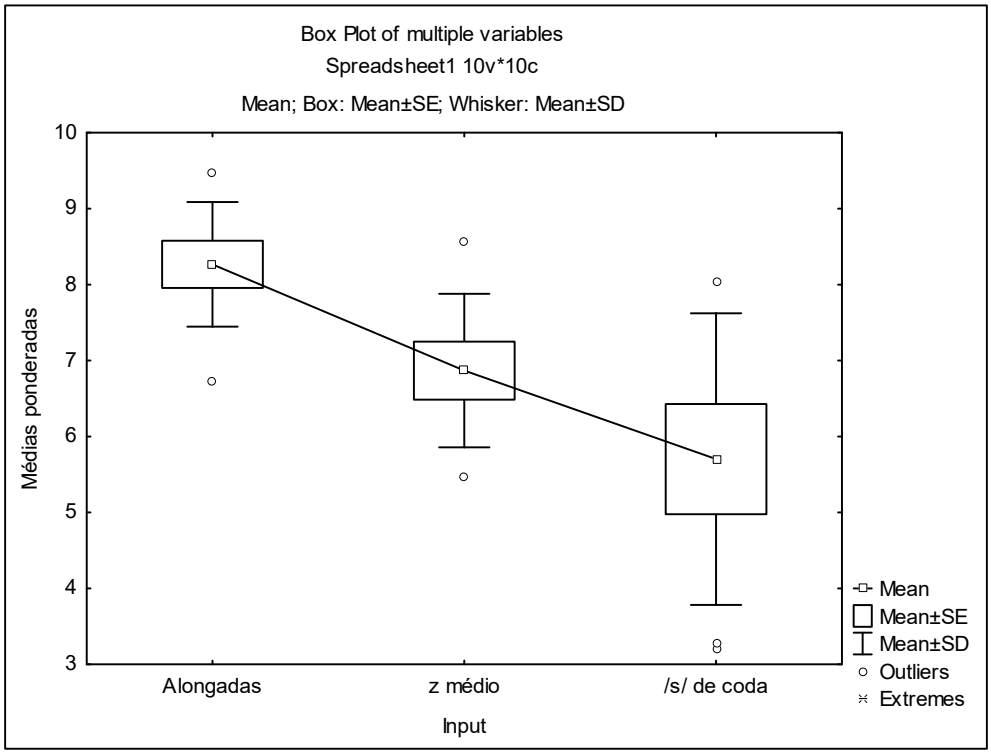


Figura 20 – Média do escore ponderado pela confiança do ouvinte para/s/ alongado, de duração média e de coda

O /s/ de duração média e o /s/ de coda apresentaram mais erros e menos confiança dos sujeitos no teste de identificação forçada. Em contraste, o /s/ alongado rendeu mais acertos e mais confiança. O ruído alongado, por conseguinte, parece transportar sinais importantes para a recuperação da vogal desvozeada. Tal resultado é coerente com os dados acústicos e aerodinâmicos e explicado em termos de alterações de *timing* e diferentes graus de sobreposição gestual – que, em caso extremo, resulta em *blending*.

Se, por um lado, as diferenças estatísticas nos mostram que há mesmo uma preferência para as alongadas, ainda há muita variação entre os sujeitos estudados, o que parece ser uma constante quando tratamos do desvozeamento.

Na Figura 21 a seguir, mostramos a variação dos escores ponderados por sujeito. As alongadas mostram o mesmo efeito da Figura 20 acima, isto é, mostram sempre mais acerto e mais confiança (valores mais altos). Logo abaixo, o /s/ de duração média também tem a mesma tendência entre os sujeitos. O /s/ de coda, por outro lado, é aquele que causou mais confusão, como já foi dito. Isso fica claro na variação da linha verde apresentada na Figura 21.

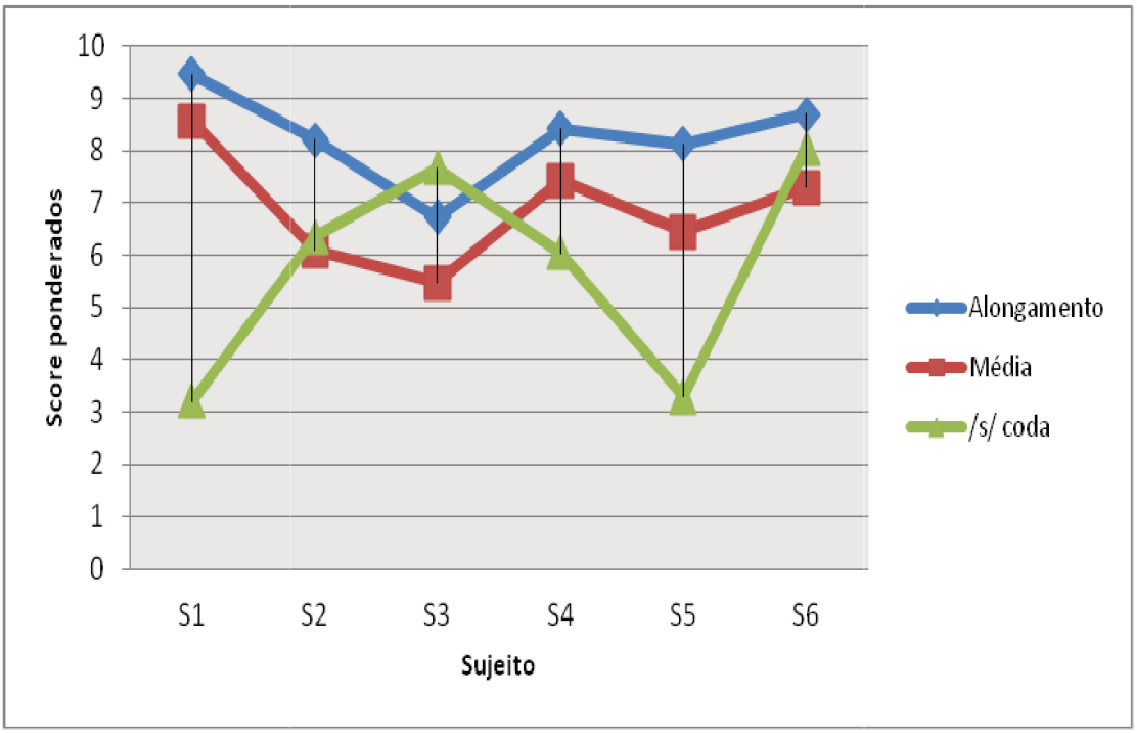


Figura 21 – Variação da média dos escores ponderados por sujeito. Note a variação para o /s/ de coda (indicada pela linha verde); alongadas e /s/ de duração média mostram menos variação.

Outro resultado que merece nossa atenção é o viés direcionado para /i/ nas escolhas de nossos informantes. Desconsiderando o erro ou acerto dos sujeitos, a vogal /i/ foi a preferida: o viés foi de aproximadamente 40% para as alongadas e para o /s/ de coda e foi maior para o /s/ de duração média, conforme a Figura 22 abaixo.

Em nossa visão, o viés voltado para /i/ é resultado da proximidade existente para os articuladores do /i/ e da consoante alveolar /s/. Mais uma vez, as alongadas, que tiveram maior número de acertos, parecem carregar mais pistas que ajudam no reconhecimento da vogal.

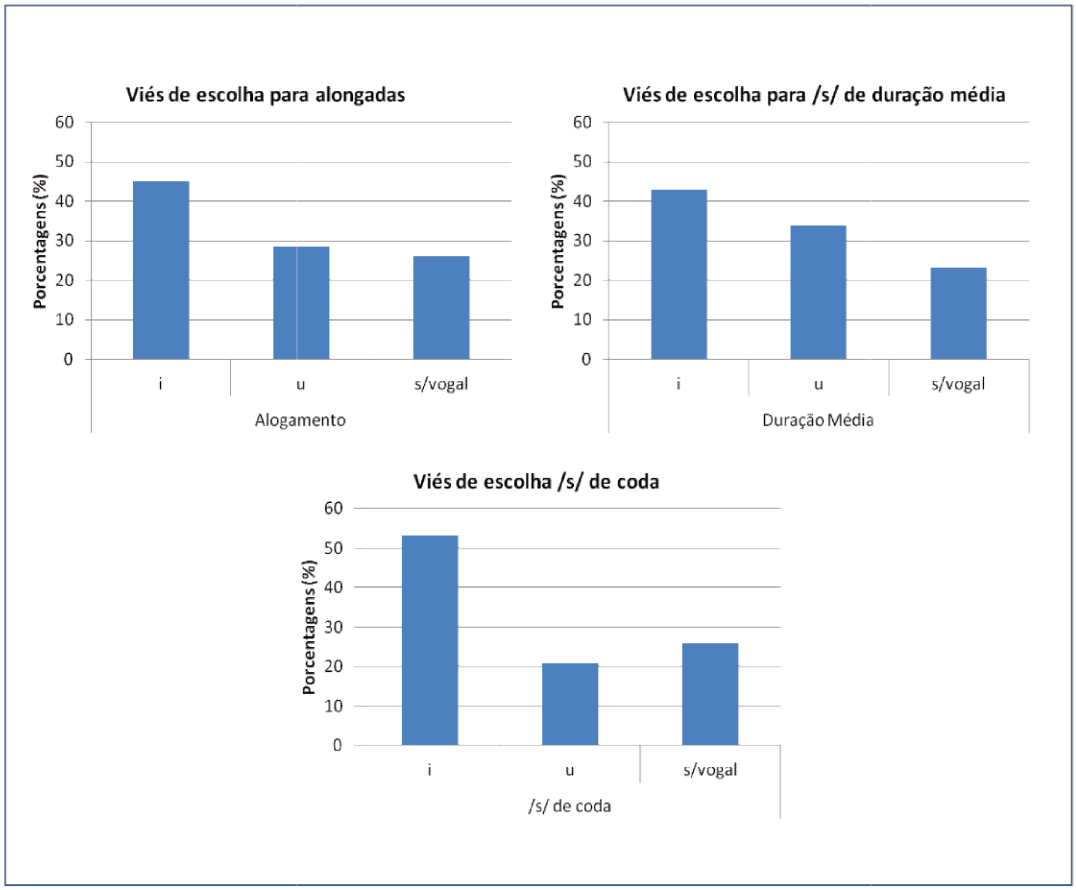


Figura 22 – Porcentagem de escolha de cada vogal no Experimento III. Em todos os casos, o /i/ foi a escolha mais recorrente.[1]

4.2.6 Discussão

Diferentemente do que aconteceu com o teste-piloto, os resultados do Experimento III nos dão mais possibilidades de discussão dos dados.

O primeiro ponto a se considerar é: os resultados mostram que o ruído do /s/ carrega pistas importantes para a identificação da vogal pelo ouvinte. A taxa de erro, relativamente baixa nessa condição em comparação com o ruído médio e o /s/ de coda, atesta essa afirmação.

Outro ponto do experimento merece destaque: a troca de escores com categorias intermediárias (como no teste-piloto) para escores com níveis de confiança mostrou-se uma opção mais fiável. A comparação entre os escores ponderados pelo nível de confiança mostra, estatisticamente, que os ouvintes têm mais acerto e mais confiança quando ouvem o ruído

[1] Cf. no Capítulo 5 uma hipótese embasada em dados articulatórios (preliminares) para esses resultados.

alongado em detrimento dos demais casos. Um experimento de discriminação ainda se faz necessário, uma vez que é complementar ao de identificação.

A variação entre sujeitos, que é recorrente no caso do desvozeamento, também é digna de discussão. Essa variação foi mais evidente quando o /s/ de coda era o estímulo. Espera-se que, em uma tarefa de discriminação, o /s/ de coda seja diferenciado com mais precisão.

Mesmo mostrando uma intensa variação entre os sujeitos, os resultados são bem consistentes. O teste de identificação tem seus resultados satisfatoriamente explicados pelas características acústicas, uma vez que as consoantes alongadas são decisivas para a detecção da presença de vogais. Além disso, os resultados apresentados acima se encaixam bem com os resultados do Experimento Aerodinâmico: a porção de vogal fricatizada aumenta a sensibilidade dos ouvintes para o reconhecimento dos estímulos.

4.3 Resumo geral dos achados e discussão

Os resultados deste trabalho mostram, até aqui, que a acústica, a aerodinâmica e a percepção de vogais finais radicalmente reduzidas no PB realmente contradizem a literatura fonológica mostrando que não há um apagamento nessa posição.

Dando seguimento às conclusões de Meneses (2012), os dados acima mostram que a redução da vogal pós-tônica final varia de moderada a extrema, o que implica desvozeamento na medida em que se aproxima da constrição crítica que leva ao desvozeamento total em um ambiente desvozeado. Em resumo, a redução de vogal final é expressa em cinco maneiras em nossos dados, a saber: (i) vogais parcialmente vozeadas são altamente centralizadas e mais curtas, se seguidas de uma consoante desvozeada; (ii) quando a vogal é aparentemente apagada, o centroide evidencia diferenças entre sobreposição regular e extrema; (iii) o ruído do /s/ aparentemente se alonga quando a vogal é totalmente desvozeada; (iv) o fluxo de ar do ruído alongado das desvozeadas é análogo ao da vogal vozeada, como revelam os testes aerodinâmicos; (v) esse alongamento facilita a recuperação vogal.

A seguir, vamos tentar integrar a interpretação dos fatos acima.

Quando precedidas de /s/ e seguidas por uma consoante desvozeada, vogais vozeadas são muito curtas e altamente centralizadas, como indicado pela diminuição tanto da duração da vogal quanto do VSA. Quando a vogal é totalmente desvozeada, o gesto da

consoante domina. As vogais tendem a encurtar de forma gradual e deslizar sob a consoante. Quando a sobreposição é tão extrema que produz uma vogal fricatizada, duas pistas acústicas apontam para a presença de um gesto vocálico. Em primeiro lugar, um centroide reduzido sinaliza a vogal sobreposta; em segundo lugar, o alongamento do ruído aponta para a sobreposição máxima e o consequente desvozeamento total. A vogal resultante desse processo é muito curta e comprimida para ser capaz de iniciar o vozeamento. Como dito, a vogal fricatizada parece ser um passo adicional no processo de redução.

Os dados aerodinâmicos do Experimento II também fornecem mais conhecimento sobre a extrema sobreposição entre a consoante e a vogal. Os dados mostram que o fluxo de ar da porção final da constrição é consistente com a presença de uma vogal fricatizada. Isso porque a configuração do trato vocal é mais estreita do que em vogais vozeadas e o tempo é insuficiente para a queda da pressão supraglotal. Esses resultados explicam bem os resultados do teste de identificação, nos quais a vogal fricatizada aumenta a sensibilidade dos ouvintes para as pistas de vogal do ruído.

Como visto acima, o desvozeamento vocálico é diretamente proporcional à sobreposição no PB. Assim, quanto maior for a sobreposição entre a vogal e consoante, mais a vogal é desvozeada. Em outras palavras, o *timing* de produção dos gestos de *onset* e núcleo na posição átona final determina a quantidade de vozeamento. Portanto, como previsto pela FonGest, a dimensão temporal de articulação parece ser o componente crucial do desvozeamento na variação sincrônica e na mudança de som.

Capítulo 5

"*Diretrizes para pesquisas futuras do desvozeamento*"

5.0 Breve introdução aos dados

Antes da conclusão propriamente dita, dados de desvozeamento para vogais seguidas de oclusivas e dois *corpora* de dados articulatórios serão tratados neste tópico de forma excepcional: trata-se de dados que já são sequelas do presente trabalho, estão em análise, publicados ou em processo de publicação. Falando mais especificamente sobre os dados articulatórios, o primeiro diz respeito ao trabalho apresentado no LabPhon 2016[1], que apresentou dados articulatórios sobre a velocidade dos gestos envolvidas no desvozeamento/redução de vogais em posição fracas da sentença por meio do *Electromagnetic Articulography* (EMA)[2]. Além desses, apresentamos brevemente dados de ultrassom[3] ainda em processo de preparação, fruto da parceria entre o LAFAPE-Unicamp e o Illinois Phonetics & Phonology Lab[4].

5.1 Oclusivas

Dados iniciais, ainda em análise, mostram que o mesmo padrão encontrado nas fricativas se repete no que diz respeito às oclusivas, ou seja, também há desvozeamento da vogal quando ela é acompanhada por oclusivas surdas.

Na Figura 23 a seguir, mostramos a forma de onda e o espectrograma de dois exemplos retirados do referido experimento.

[1] O resumo apresentado pode ser visto no Anexo 1 ao fim da tese.

[2] EMA é a técnica que aplica campos magnéticos alternados para controlar o movimento de sensores fixados nos articuladores e é utilizado para avaliar a velocidade e precisão dos movimentos da língua e dos lábios.

[3] Os referidos dados não podem compor uma seção especifica da tese por terem sido obtidos em parceria com o *Phonetics and Phonology Lab* da University of Illinois at Urbana-Champaign (UIUC) e por serem de autoria conjunta.

[4] O projeto em questão foi financiado pelo Lemann Institute for Brazilian Studies – UIUC.

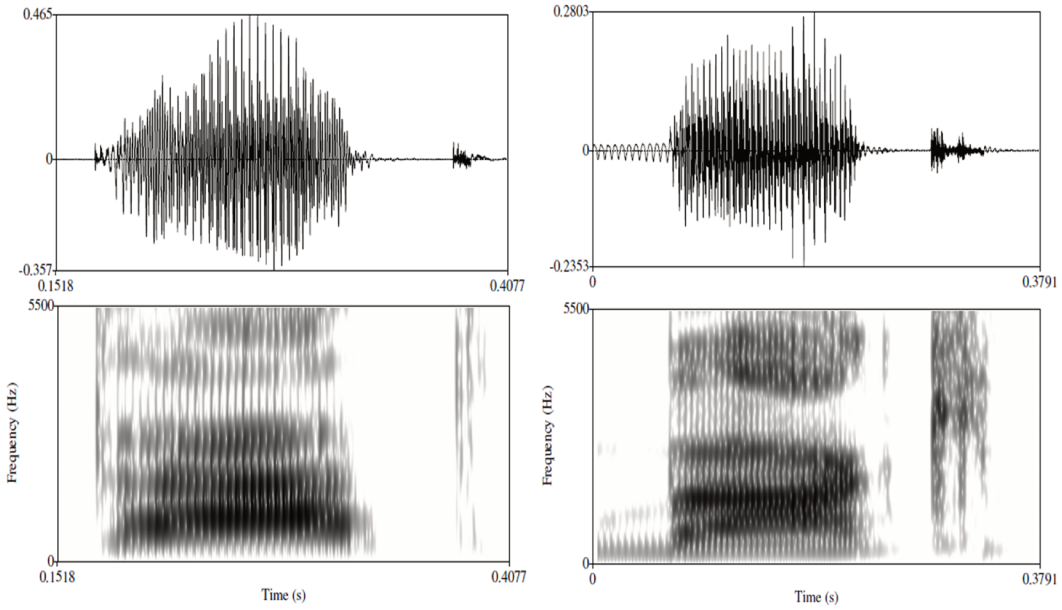


Figura 23 – Exemplos de palavras com sílabas com vogais desvozeadas acompanhadas de oclusivas para tato [ˈtatʊ̥] e baque [ˈbakɪ̥], respectivamente

Assim como nas fricativas, as vogais altas perdem vozeamento, periodicidade e estrutura formântica. Além disso, não é possível identificar qualquer padrão formântico caracteristico das vogais. As análises iniciais ainda mostram que há uma diferança de duração no estouro das oclusivas. Se confirmados, os dados de fricativas e ocluivas podem atestar que, com o encolhimento do gesto vocálico, a porção da vogal assíncrona com a consoante torna-se curta e fraca demais para sustentar a aerodinâmica da produção do vozeamento. A fusão (sobreposição extrema) dos gestos da consoante e da vogal, como mostramos para as fricativas, também parece ter padrão semelhante nas oclusivas.

5.2 EMA

5.2.1 Metodologia

Em Meneses *et al*. (2016), os autores apresentam dados de EMA para ajudar a esclarecer a articulação das vogais pós-tônicas finais desvozeados no PB e explicar como suas características supralaríngeas diferem em relação aos seus pares vozeados. Embora os dados apresentados pelos autores tenham sido recolhidos no contexto de outro experimento, a análise ajudou a gerar hipóteses para futuras pesquisas.

No referido trabalho, dados de posição da língua foram coletados usando-se o sistema NDI Wave. Três sensores[5] foram colados em intervalos de 1 cm na porção anterior da língua (ponta: TT; meio: TM; dorso: TB). Outros sensores foram colocados na mandíbula/queixo (JC) e nos lábios superior e inferior (UL, LL) e um sensor adicional foi utilizado para corrigir a rotação da cabeça. Sete sujeitos falantes do PB de São Paulo gravaram a frase-veículo "D**i**go _ para t**i**"[6]. As vogais selecionadas foram: (i) o /i/ de "Digo" (doravante /i/^{1}) e (ii) o /i/ do final na sentença em "ti" (doravante /i/^{2}), de forma que o primeiro era sempre vozeado e o segundo era mais susceptível ao desvozeamento.

Os dados foram segmentados manualmente no software Praat. Nos casos em que não houve vozeamento, todo o ruído fricativo foi rotulado como sílaba, pois o esperado era que os traços articulatórios de /i/ pudessem ser mapeados pelo EMA. Os valores foram medidos dentro dos limites do segmento e os dados foram normalizados usando um script automatizado no Matlab (Wong e Hermes, 2015). A posição e a velocidade média dos sensores foram as unidades mensuradas.

5.2.2 Resultados

UL, LL, TT, TM e TB foram submetidos à análise de componentes principais (PCA) e à análise linear discriminante (LDA) (Jombart, 2013). Dois PCs representaram de forma ideal a variação nos dados. O primeiro componente (PC1) explicou 40% da variância e apresentou forte correlação negativa com a protrusão labial e a elevação da língua. O segundo componente (PC2) explicou 23% da variância e se mostrou fortemente correlacionado com a anteriorização da língua. O LDA baseado nos PCs resultantes mostrou que ambos /i/^{1} e /i/^{2} foram claramente diferenciados ao longo dessas duas dimensões, como pode ser visto na Tabela 1.

[5] As siglas dos sensores se referem aos nomes em inglês: TT = *Tongue Tip*, TM = *Tongue Mid*, TB = *Tongue Body*, J/C = *Jaw/Chin*, UL = *Upper Lip* e LL = *Lower Lip*.

[6] As palavras-alvo (X) do experimento original (monossílabos nasais) foram desconsideradas em nossa análise.

Tabela 1 – Classificação dos resultados do LDA, mostrando a separação de /i/[1] e /i/[2] em duas classes

LD-Class	i[1]	i[2]
1	***57***	17
2	14	***58***

Levando em consideração a primeira derivada do sinal do corpo da língua, a velocidade vertical do TM (dTMy) para /i/[2] (em "para ti") foi consistentemente maior do que a velocidade para /i/[1] (em "digo"), como na comparação dos gráficos de *boxplot* da Figura 24. Esse resultado, mesmo que preliminar, vai ao encontro da hipótese desta tese, de que a posição fraca da palavra ou da frase favorece o desvozeamento, uma vez que não há tempo suficiente para o vozeamento.

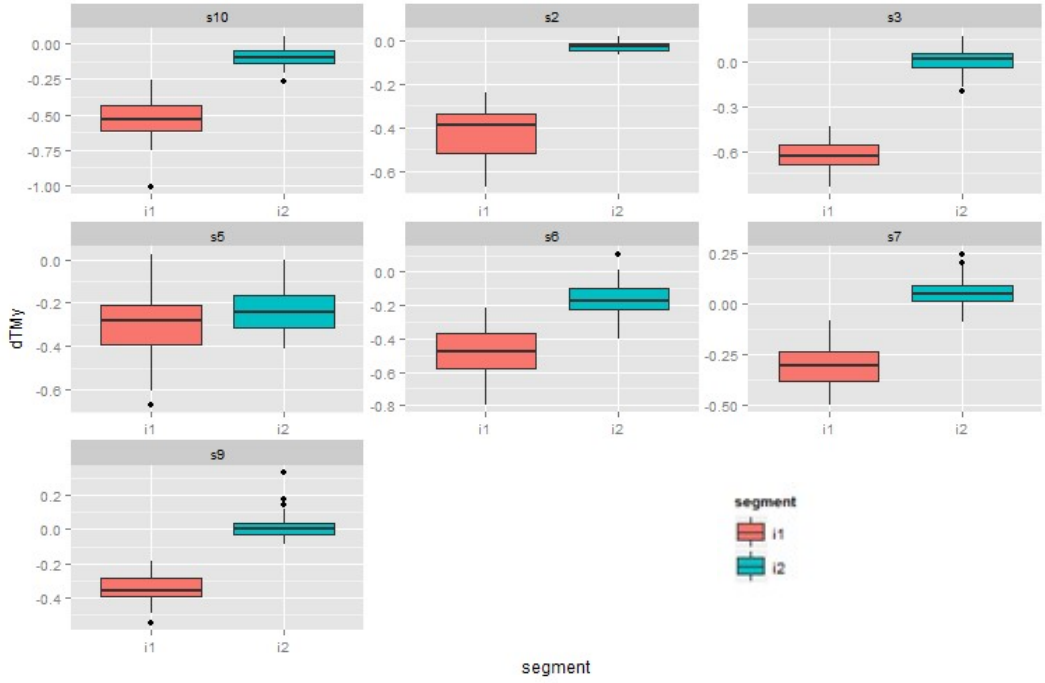


Figura 24 – *Boxplots* para a velocidade do sensor da posição medial da língua para sete sujeitos do PB (em mm/s) para /i/[1] (vermelho) e /i/[2] (azul)
Fonte: Meneses *et al*. (2016)

A Figura 25 a seguir mostra as vogais /i/[2] (em azul) e /i/[1] (em vermelho). De forma geral, podemos dizer que as duas vogais são articulatoriamente distintas para o trato vocal anterior: /i/[2] parece ser produzido com mais protrusão labial e levantamento da língua.

O dado acima é relevante, mas é importante lembrar que as vogais são produzidas em ambientes diferentes, uma seguida por /g/ se outra seguida por pausa.

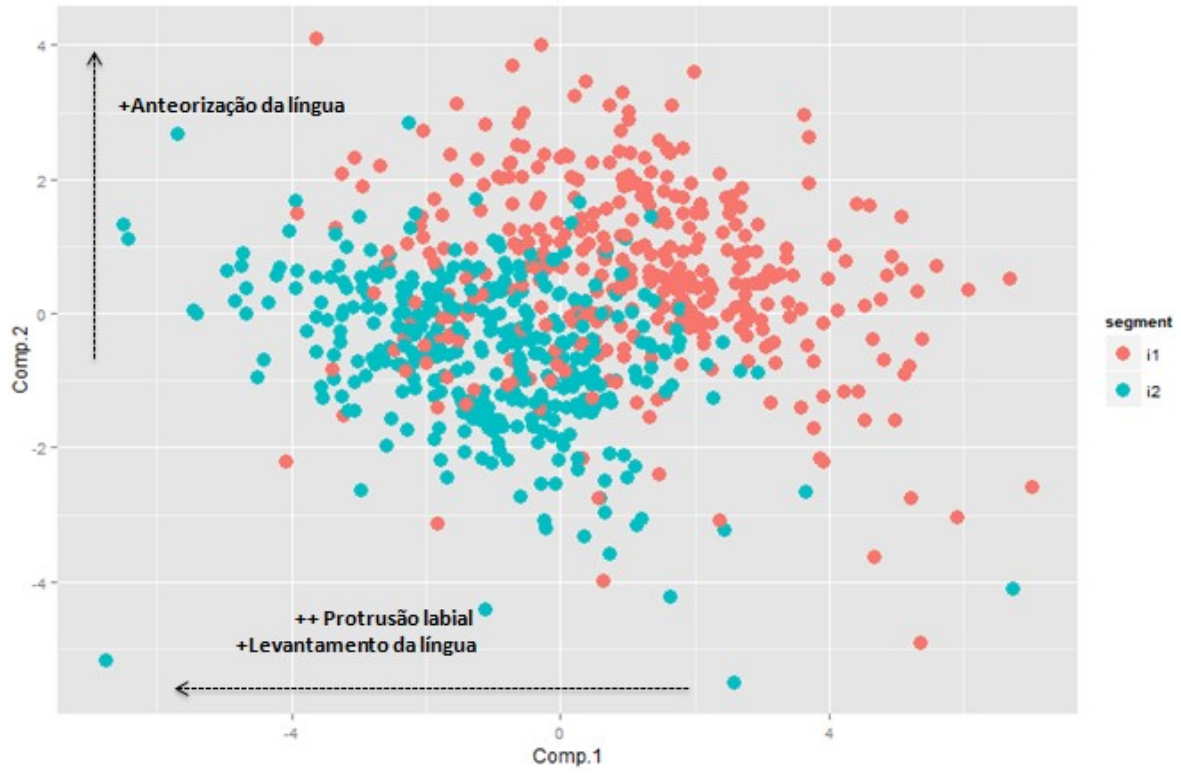


Figura 25 – Plano cartesiano com os escores de /i[1]/ (vermelho) e /i[2]/ (azul) em PC1 (x) e PC2 (y)
Fonte: Meneses *et al*. (2016)

Como o *design* do experimento não foi concebido tendo em vista o desvozeamento, poucos sujeitos desvozearam as vogais com suficiente frequência.[7] Por isso, optou-se por observar a produção de i[2] por um sujeito que consistentemente desvozeava tais vogais e aquele que não o fez. Foi feita, então, a comparação entre S9 e S10: S9 produz a vogal final desvozeada na maior parte das ocorrências e S10 apresenta isso com muito menor frequência. Para S9, a posição do corpo da língua foi maior em comparação aos demais sujeitos (incluindo S10), como podemos ver na Figura 26 a seguir. A velocidade absoluta de sensores de S9 também foi maior durante a vogal desvozeada.

[7] Apesar de ser a palavra final na sentença, a palavra alvo "ti" carrega mais força prosódica se comparada com uma sílaba pós-tônica final em ambiente prosódico átono e seguida de consoante surda, por exemplo. Por isso, é totalmente ju stificável a baixa frequência de desvozeamento total.

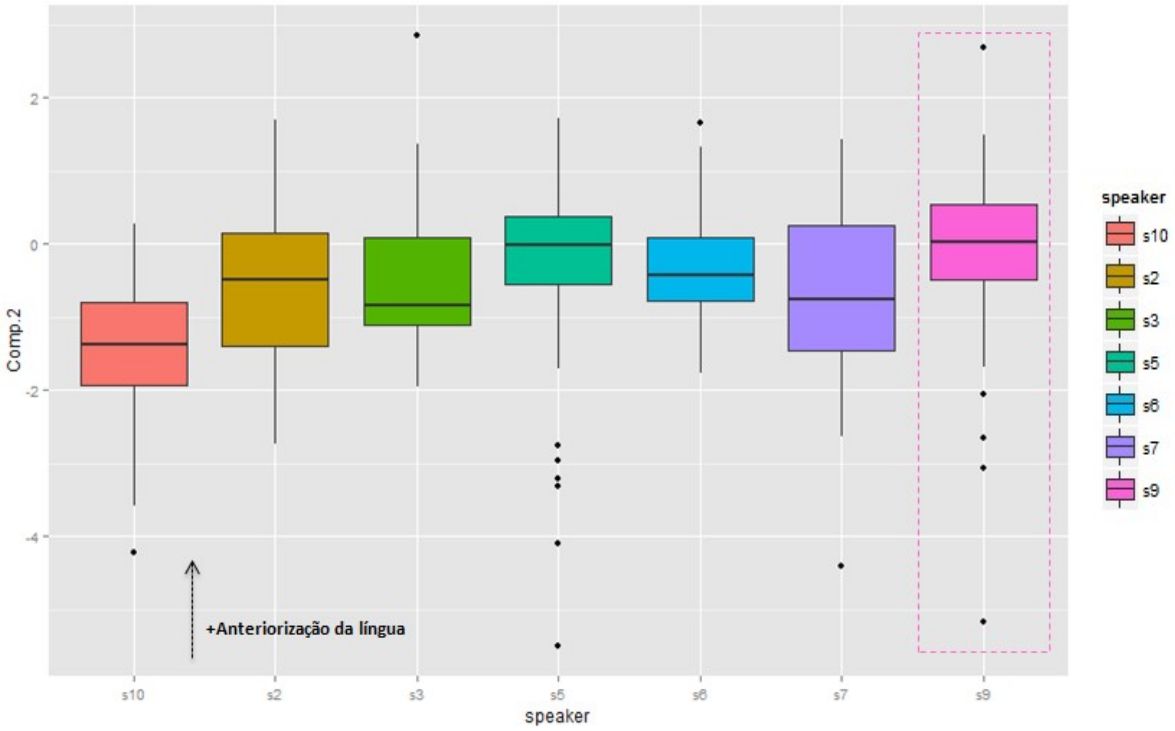


Figura 26 – *Boxplots* dos escores de sete sujeitos PC2 ordenados pelo grau de anteorização. S9 é destacado por o ser o que mais apresenta vogais desvozeadas.
Fonte: Meneses *et al*. (2016)

S9 e S10 parecem diferir significativamente considerando o PC2 (que pode ser geralmente descrito como a lâmina da língua, com base na análise de PC), sugerindo que existe uma diferença de articulação entre os /i/s vozeados e desvozeados, mesmo quando ambos estão na posição fraca: /i/ desvozeado é uma produção mais anterior.

A Figura 27 a seguir mostra que /i/^{2} foi produzido com mais protrusão do lábio e elevação da língua do que /i/^{1}. Da mesma forma, /i/^{2} também apresentou a posição da língua mais anterior quando foi desvozeado. Note na figura abaixo que os triângulos azuis geralmente têm a posição da língua mais anterior (maior PC2) do que círculos. Uma vez que S9 produz vogais desvozeadas de forma mais consistente, consideramos esta diferença associada ao desvozeamento.

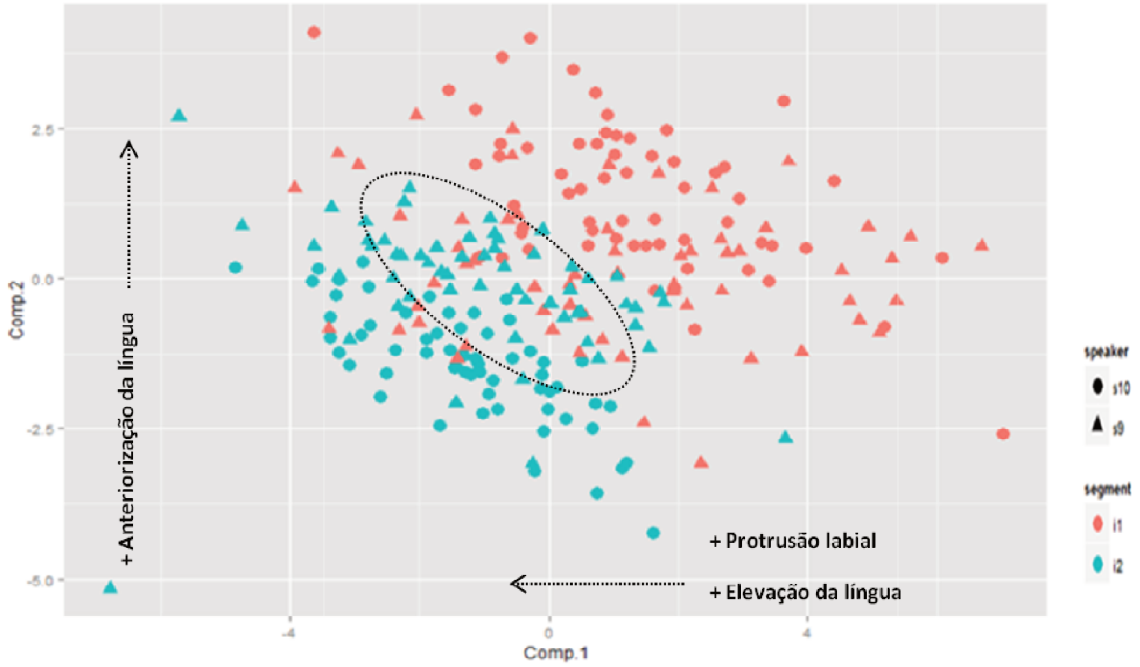


Figura 27 – Plano cartesiano de /i/[1] e /i/[2] para PC1 (x) e PC2 (y) dos sujeitos S9 e S10. Triângulos = S9; Círculo preto = S10; Círculo vermelho = /i/[1]; Círculo vermelho = /i/[2].
Fonte: Meneses *et al*. (2016)

Os resultados de Meneses *et al*. (2015) sugerem que a vogal desvozeada tem uma diferença na coordenação gestual em comparação com uma vogal totalmente vozeada. O aumento da velocidade e a elevação do corpo da língua observados também corroboram os dados acústicos/aerodinâmicos analisados descritos nesta tese e a hipótese descrita no Capítulo 3. As mudanças na configuração da língua são um apoio consistente à hipótese de que a vogal desvozeada no PB tem seu próprio ajuste do programa motor supralaríngeo.

Os dados acima ainda não permitem afirmar categoricamente que há diferença na coordenação gestual, pois pode haver uma diferença nos alvos das vogais, uma vez que /i/[1] e /i/[2] estão em diferentes posições prosódicas. Em contrapartida, esse tema carece de investigação, pois ainda temos pouca compreensão dos aspectos supralaríngeos da diferença entre as vogais desvozeadas e vozeadas. Ainda é possível trabalhar mais na tentativa de interpretar os resultados levantando questões sobre o que poderia estar acontecendo globalmente no trato vocal durante a produção de vogais surdas. Em trabalhos futuros, é possível observar diferentes porções da fricativa, antecipando quais os traços articulatórios da vogal podem existir no final da fricativa. Além disso, ainda é necessário controlar as condições prosódicas e o contexto fonético (anterior e posterior aos alvos).

5.3 Ultrassom

5.3.1 Visão geral do método

Os dados analisados em Meneses *et al*. (em preparação) serão apresentados aqui de forma breve. Em suma, o trabalho investiga o desvozeamento sob o viés articulatório através de imagens de ultrassom. O ultrassom é capaz de capturar a dinâmica da forma da língua, permitindo a aquisição de medidas relacionadas às posições de ponta, meio e raiz da língua, bem como as interações articulatórias entre vogais e consoantes (Gick, 2002). Recentemente, uma ampla variedade de tecnologias de imagem tem sido aplicada para medir e rastrear a articulação durante a fala.

No ultrassom, usa-se um transdutor em contato com a superfície corporal para coletar sinais ecográficos do trato vocal, acoplado a um computador com *software* de análise de imagem a fim de rastrear o movimento da língua através das imagens adquiridas. Ele permite a imagem do tecido mole, mas não de osso ou ar. Assim, quando mantido contra o tecido mole abaixo do queixo e virada para cima e para trás na direção da superfície da lingueta, duas secções dimensionais da superfície superior da língua podem ser mostradas a partir de raiz de lâmina (exceto para "sombras" lançadas pelo osso hioide, mandíbula e na cavidade sublingual).

A posição do *probe* foi mantida estável através do uso de um capacete que estabilizava sua posição. O sujeito foi mantido sentado em uma cadeira sólida de frente para a tela na qual as palavras alvos eram projetadas. Além disso, simultaneamente, é realizada a gravação da voz em tempo real.

5.3.2 Resultados iniciais

Dada a brevidade dos dados de Meneses *et al*. (em preparação), os resultados apresentados nesta subseção servem, até o momento, como um suporte ao nosso ponto de vista sobre a articulação das vogais reduzidas e desvozeadas no PB. Uma hipótese mais embasada só será possível depois da análise completa e publicação dos dados.

A Figura 28 a seguir mostra os *splines* (curvas) para o /s/ de coda. Observada a escala de cores que representa a movimentação da língua, é possível perceber que o movimento da língua durante a produção do /s/ é pequeno e estável.

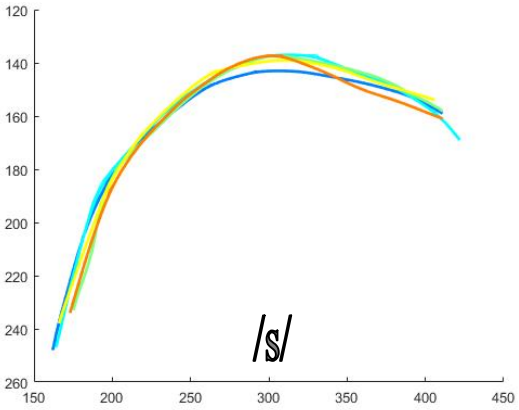


Figura 28 – Imagem sagital do quadro correspondente à fricativa /s/ no sinal acústico para "paz". A ponta da língua está à direita e a raiz da língua está do lado esquerdo.
Fonte: Meneses *et al.* (em preparação)

Há uma clara e esperada diferença entre o /s/ de coda e o /si/ vozeado, que pode ser visto na Figura 29 abaixo. O /si/ desvozeado, por outro lado, como indicavam os dados acústicos e aerodinâmicos, mantém o gesto de elevação da região medial da língua. Como afirmamos nos capítulos anteriores, a vogal desvozeada parece se diferenciar de sua contraparte sonora apenas pela falta de vozeamento e pela diminuição da duração e da magnitude do gesto (que também pode ser vista na forma dos *splines* na região de contato da língua com a região palato-alveolar).

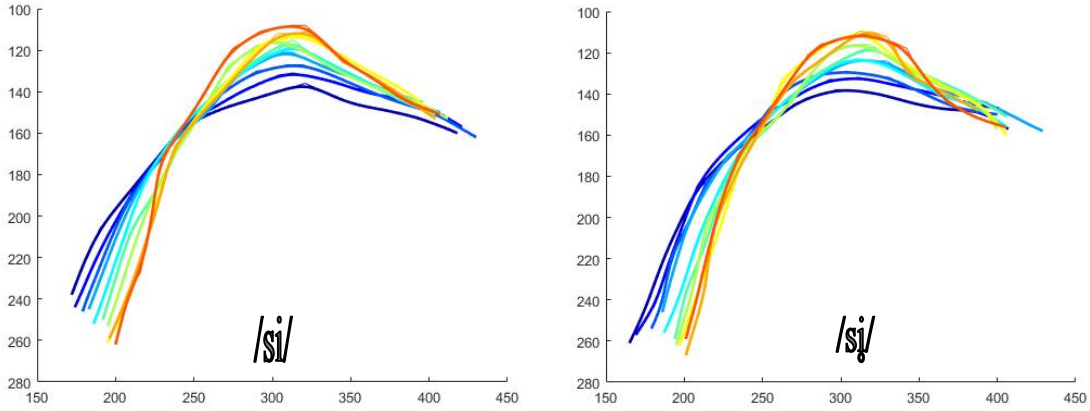


Figura 29 – Imagem sagital do quadro correspondente à sílaba /si/ com vogal vozeada (à esquerda) e desvozeada (à direita) no sinal acústico para a palavra "passe". A ponta da língua está à direita e a raiz da língua está à esquerda.
Fonte: Meneses *et al.* (em preparação)

Para /su/ vozeado e desvozeado, diferentemente de /si/, a diferença parece ser mais gritante. Nota-se na Figura 30 abaixo que, enquanto no /su/ vozeado é possível perceber a elevação da região mais posterior da língua, o /su/ desvozeado não apresenta a elevação da mesma grandeza.

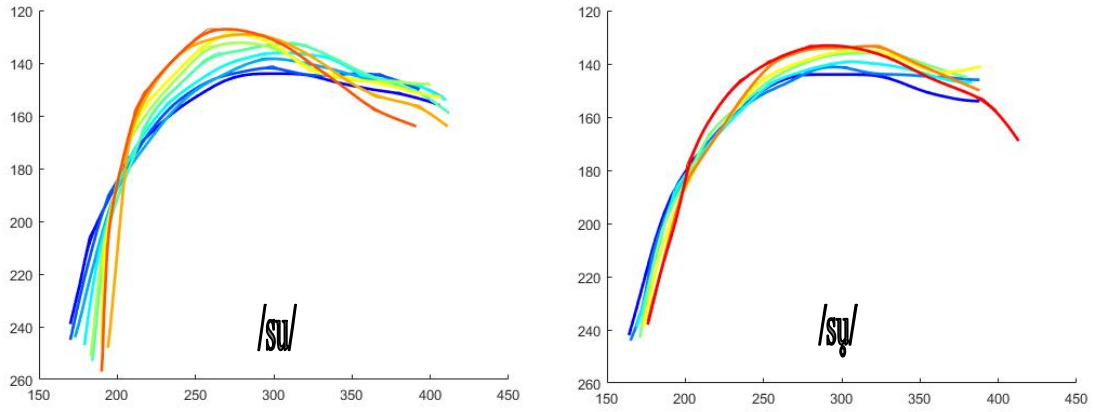


Figura 30 – Imagem sagital do quadro correspondente à sílaba /su/ com vogal vozeada (à esquerda) e desvozeada (à direita) no sinal acústico para a palavra "passo"
Fonte: Meneses *et al.* (em preparação)

Diante do exposto nos capítulos anteriores, os resultados das duas figuras acima permitem duas observações: em primeiro lugar, uma diminuição da magnitude do gesto de /u/ em /su/ poderia aproximar ainda mais /si/ desvozeado de /su/ desvozeado, fundamentando o resultado de nosso experimento de percepção que tinha, como mostrado, um viés de escolha claramente voltado para /i/; em segundo lugar, o mesmo relaxamento na constrição de /u/ pode ser efeito de um controle articulatório, uma vez que a articulação do /u/ ainda envolve o gesto de arredondamento dos lábios. Esse último, então, compensaria a afrouxamento daquele gesto mais anterior.

De forma geral, os dados de Meneses *et al*. (em preparação) exemplificados aqui parecem amparar (mesmo preliminarmente) a maioria dos resultados acústicos apresentados nos capítulos anteriores, umas vez que mostram que os gestos supraglóticos da vogal resistem ao ambiente pós-tônico, ainda que o vozeamento não resista. O trabalho em preparação pretende analisar a posição e o deslocamento da língua e pode iluminar o problema em questão bem como atestar (ou não) as hipóteses levantadas aqui.

5.4 Conclusão

Os nossos resultados indicam que a redução de vogais átonas finais no PB é difundida, mas progressiva, variando de encurtamento com desvozeamento parcial para o desvozeamento completo. Essa variabilidade parece resultar de diferenças na coordenação gestual e diferentes graus de sobreposição entre gestos de consoantes e vogais. Tal gradiente está mais relacionado a uma sobreposição extrema do que ao apagamento de vogal, como apontam os dados acústicos, aerodinâmicos e também articulatórios. No entanto, como a vogal é extremamente curta, os ouvintes podem "se enganar" por conta das fracas pistas vocálicas no sinal acústico, de modo que eles não consigam recuperar tal sobreposição.

O desenvolvimento de nossa hipótese nos leva a crer que a variação sincrônica existente (vogal reduzida/desvozeamento parcial e desvozeamento total) pode ser a matéria-prima para uma mudança fônica que caminha progressivamente ao apagamento completo da vogal.

Se, no final, como mostra a Figura 31 a seguir, a sobreposição com desvozeamento total for aceita como variante, a apócope pode ser gradualmente assumida por mais e mais ouvintes, uma vez que estes têm graus variáveis de sensibilidade aos sinais da vogal. Então, vogais desvozeadas, isto é, vogais fricatizadas completamente sobrepostas podem ser a matéria-prima e desencadear o processo de apócope baseado na percepção do ouvinte em um longo prazo, como defendido por Ohala (1981).

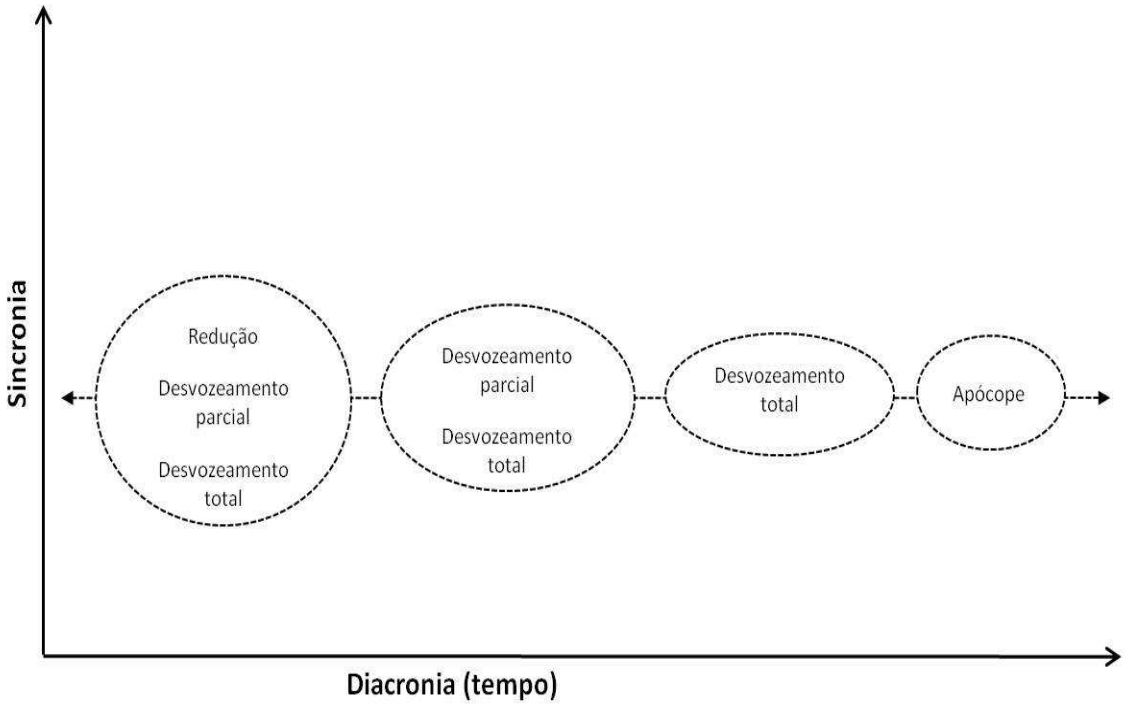


Figura 31 – Linha da variação sincrônica e da variação diacrônica hipotética. O eixo x representa a variação diacrônica e o eixo y o momento sincrônico.

PARTE 2

A ELISÃO NO SÂNDI
OU A REDUÇÃO DA VOGAL BAIXA

Capítulo 6

"Revisitando o processo de sândi vocálico no PB"

6.0 Introdução

Já foi visto que a redução dos segmentos é uma das variações mais comuns na fala espontânea. Nessa perspectiva, as vogais do PB constituem um terreno fértil para o estudo de fenômenos que estão no limiar da relação entre processos categóricos e gradientes (Albano, 1999, 2001; Ohala, 2005)[1], tais como a redução/vozeamento e a redução/apagamento (ver Capítulo 3). As vogais, conforme a posição relativa ao acento, têm produção variável, sendo que essa variabilidade resulta dos efeitos somados de processos aparentemente categóricos sensíveis à posição do acento e de processos gradientes sensíveis à taxa de elocução, entre outros fatores. Esse tipo de variabilidade surgiu na análise do desvozeamento na Parte 1 da tese e, a partir daqui, tomará a forma da elisão no sândi vocálico externo.

Se olharmos os fenômenos estudados aqui do ponto de vista categórico, há processos de neutralização das distinções de altura das vogais não baixas nas posições pré-tônica e pós-tônica, que provocam uma suposta queda da vogal; do lado gradiente, há processos de redução de todas as vogais nas posições átonas, especialmente a pós-tônica, que reduz drasticamente a magnitude dos gestos vocálicos.

A redução extrema de vogais é uma manifestação de uma tendência das línguas de reduzir gestos nas posições átonas. Essa tendência não parece ser igual no desvozeamento e no sândi vocálico, que são os casos analisados nesta tese. No caso específico do sândi, diferentemente do ocorrido no desvozeamento vocálico, o fato de a redução afetar a vogal baixa /a/, que não está sujeita à neutralização das átonas nem ao desvozeamento, revela sua particularidade mais intrigante.

O objetivo dos experimentos que apresentaremos a seguir é (re)investigar[2] o processo de sândi vocálico, também sob um ponto de vista dinâmico da produção da fala, na linha de Albano *et al*. (1998) e Albano (1999). Além disso, pretende-se comparar dois

[1] Albano (1999) aponta quatro fenômenos vocálicos que estão no limite entre o gradiente e o categórico, a saber: a nasalização da vogal, a palatalização das oclusivas coronais diante de /i/, a elisão de /a/ no sândi e a elevação do [ɪ] pós-tônico.

[2] A maioria dos trabalhos sobre o sândi é de cunho gerativo, como se verá a seguir.

processos de redução da vogal, levando-se em consideração a natureza particular de cada um deles. No final de nossa análise, poderemos integrar os processos em um mesmo modelo de redução da vogal no PB.

Nos próximos tópicos, revisaremos uma série de investigações que já tiveram o sândi vocálico como foco de interesse tanto na literatura internacional quanto na literatura do PB. Depois, apresentaremos os dados de produção do Experimento de Produção I e II, depois o Experimento Aerodinâmico. No fim, as considerações finais versarão sobre os resultados e as possíveis relações entre os dados analisados na tese.

6.1 O sândi nas línguas naturais

Sândi é o termo usado pelos antigos gramáticos indianos, a partir de Panini, para se referir a processos fonológicos ocasionados pela sequência direta de morfemas na formação/junção de palavras.[3] Uma diferenciação básica foi atribuída ao processo pelos gramáticos indianos: de um lado estava o sândi interno, processo que ocorre dentro de palavras; e de outro o sândi externo, processo que ocorre entre palavras na fala encadeada.

Em muitas escolas da fonologia, os processos que afetam vogais no sândi foram tratados através de regras fonológicas categóricas. A literatura internacional tem dividido o processo de sândi externo em três fenômenos gerais:

(i) As regras de modificação do morfema estão relacionadas com a "estrutura de sílaba", que altera sequências de consoantes e vogais. Em resumo, essa forma de sândi gera sílabas CV regularmente alternadas em detrimento de sequências "estranhas" de apenas consoantes ou vogais através da inserção ou elisão de consoantes como em (1a) para o holandês e em (1b) para o espanhol (Bergman, 1971; Hutchinson, 1974; Jongenburger & van Heuven, 1993).

(1) a. *moe oefent* /mu.u.fənt/ => [mu.ʔu.fənt] *'mamãe está praticando*'

b. *tu uniforme* /tu#uniforme/ => [tu.ni.for.me] '*seu uniforme*'

[3] O primeiro uso da palavra em inglês ocorreu em 1806, na qual o termo "sândi" referia-se ao processo de "permutação de letras ocasionada pela junção de sílabas" em sânscrito. Os estruturalistas americanos Edward Sapir e Leonard Bloomfield utilizaram a palavra sândi como sinônimo de "junção", e eles são os maiores responsáveis pela difusão do termo em uso moderno (Kaisse, 2006).

(ii) Também concerne ao sândi a assimilação de uma característica fonológica, ou seja, a alteração do segmento sob a influência de seu contexto fonético no encontro entre palavras. Nesse caso está a mudança de uma consoante dependendo do contexto fonético, como por exemplo, o ponto de articulação das plosivas e fricativas do holandês que assimilam o ponto do [j] (Jongenburger & van Heuven, 1993; Mitterer & Ernestus, 2006).

(2) a. /t/ alveolar => palatal [c]: *weet je* /we-t#ja/ => [we:ca] ‘você sabe’

b. /s/ alveolar => pós-alveolar [ʃ]: *was je* /was#jə/ => [waʃə] ‘foi você’

(iii) Por último, aquele que é objeto de nosso interesse: processos de sândi que se referem às vogais e são considerados formas de “ajustamentos vocálicos”. Para essa categoria de ajustes, o acento é um fator importante. Nesse caso, as vogais átonas, obviamente mais fracas, perdem parte de sua articulação (encurtamento ou redução da vogal) e podem até mesmo desaparecer completamente na fala normal e rápida, como nos exemplos em (3a) do igede (língua falada na região central da Nigéria, na qual o sândi é recorrente) e em (3b) do PB (Bergman, 1971; Jongenburger & van Heuven, 1993; Russel, 2008; Collischonn, 2011).

(3) a. yë àkpa => [yˈakpà] ‘ver uma caixa’

b. faixa etária => [faʃɛˈtariɐ]

O espanhol, assim como o português, é um bom exemplo de língua que conta com uma extensa descrição do processo de sândi externo. Nessa língua, estudos estruturalistas e gerativistas caracterizam o sândi vocálico externo como um processo de queda, como apontam os exemplos abaixo em (4).

(4) a. *esta hija => [esˈtiχa]* ‘esta filha’

b. *vendra Inez => [βenˈdrines]* ‘venha Inez’

Estudos mais recentes mostraram que essas análises são inadequadas, uma vez que elas se concentram exclusivamente em fenômenos segmentais, ignorando a interação entre fenômenos suprassegmentais e gradientes. Segundo Hutchinson (1974), o espanhol tende a ter sequências de sílabas com poucos encontros consonantais e alguns conjuntos de vogais, ou seja, o espanhol tende a ter sequências de sílabas que se aproximam de uma sequência /CV.CV.CV/ – o sândi, então, contribui para essa tendência do espanhol. O sândi da vogal pode, portanto, ser desencadeado pela existência de vogais contíguas e, segundo a autora, é controlado pelo ritmo da língua.

Como se sabe, o espanhol é uma língua de ritmo silábico (*syllable-timed*) e tende a manter constante a duração relativa da sílaba. Em geral, isso significa que, além das vogais justapostas, que se desfazem no sândi, as consoantes em torno delas também podem cair. Assim, sílabas inteiras podem se enfraquecer ou cair no interior da unidade atribuída a uma sílaba. No fim, a sílaba resultante é uma CV (C) que tem no núcleo uma vogal (é importante ressaltar que esse não parece ser um o caso do português). De acordo com Hutchinson, o enfraquecimento da primeira das duas vogais no sândi para o espanhol é consistente com a tendência específica da língua de manter um determinado tipo de sequência silábica, como nos exemplos dados acima em (4).

Parece natural que, no caso de vogais idênticas em uma sequência CV#V(C), uma delas seja apagada, uma vez que, articulatoriamente, nenhuma transição é necessária para ir de uma para a outra. Conforme Hutchinson (1974) e Perkell (1969), em sequências de vogais diferentes, como /ei/ e /ou/ e /ai/, a transição é um caso *default*, mas há uma grande variação da duração da sílaba resultante do sândi. No entanto, não fica claro nos trabalhos citados se a variação da duração é resultado de uma articulação diferente dependendo da natureza vocálica. Outro fato importante relatado por Hutchinson (1974) é a relação do processo de sândi com a qualidade vocálica. De maneira geral no processo de sândi, quando uma sequência CV*V(C) é investigada na fala rápida, o /a/ parece sumir do sinal acústico. Os dados de Hutchinson mostram que sequências /Ca#V(C)/ se comportam de uma maneira diferente em relação às sequências de /ei/, por exemplo, nas quais a queda é menos recorrente.

Do ponto de vista da articulação, Perkell (1969) sugere que o movimento da mandíbula e o corpo da língua na produção de vogais altas são diferentes em contraste com a produção das vogais baixas. Isso, segundo o autor, poderia explicar a queda da vogal baixa e a formação de um glide para as vogais médias em contexto de sândi vocálico no espanhol. O

trabalho de Perkell foi um dos primeiros, se não o primeiro, a relacionar a redução da vogal /a/ a uma articulação singular imposta pela espontaneidade da fala no sândi.

Segundo o autor, a altura da mandíbula para consoantes está nivelada com a de vogais altas, e dada a lentidão do movimento mandibular e a sua importância na articulação do /a/, é provável que haja uma perda de amplitude articulatória na produção de uma sequência /Ca#V(C)/ na fala espontânea. Além disso, esse movimento pode encontrar resistência dentro das limitações de tempo da fala corrente. Essas relações articulatórias podem explicar por que o /a/ é apagado e por que as vogais altas e médias se sustentam no sândi vocálico, como se verá adiante. Nesse ponto, cabe pensar na diferença entre o desvozeamento e o sândi, já que naquele processo são as vogais altas que se preservam mesmo diante do curto tempo em sílabas CV. O sândi, nesse contexto, parece ser consequência de efeitos de *timing* e do sequenciamento silábico na produção de sequências V-V.

A partir da década 1990, estudos fonéticos mais cuidadosos mostraram, caso após caso, que os fenômenos de sândi são, em sua maioria, gradientes e altamente variáveis. Enquanto Ladd & Scobbie (2003) verificaram alguns casos de sândi externo categórico, no qual há uma geminação de consoantes para três sujeitos da Sardenha[4], Nolan, Holst & Kühnert (1996), utilizando eletropalatografia, e Zsiga (1994, 1995), através da análise espectral, mostraram que a mudança no inglês de /s/ para um aparente [ʃ] em sequências como "*miss you*" era um processo gradiente e o gesto resultante era distinto do /ʃ/ de "*issue*". Para eles, essa produção intermediária foi mais bem explicada pela sobreposição gestual.

No que concerne à variabilidade, Ellis & Hardcastle (2002), usando eletropalatografia para estudar a assimilação decorrente do sândi em nasais do inglês em sequências /n#k/, encontraram uma intensa variabilidade entre falantes na fala rápida: alguns sujeitos nunca assimilam a nasal, alguns sempre fazem a assimilação categoricamente, alguns alternam entre nenhuma assimilação e assimilação completa[5]. Além desses, Russell (2008) mostra que o sândi é um processo diversificado que pode variar de um tipo categórico ao gradiente, resultado de diferentes coordenações e da sobreposição gestual. A discussão da

[4] Segundo os autores, há muitas variações sincrônicas em fronteiras de palavra na Sardenha e elas envolvem a lenição intervocálica de obstruintes [V#rV], o vozeamento de fricativas na posição inicial de palavras ([ˈfradɛ] 'irmão' => [suˈvɾadɛ] 'o irmão'), o vozeamento de oclusivas iniciais fracas ([ˈkanɛ] 'cachorro' => [suˈganɛ] 'o cachorro') e o desaparecimento de oclusivas iniciais sonoras, deixando vogais em hiato ([ˈbak:a] 'vaca' => [sa.ˈak:a] 'a vaca').

[5] Na nossa visão, essa variação dita como surpreendente é característica mais que comum em processos de redução, como já vimos nos casos de desvozeamento.

variação, tratada desde as primeiras linhas desta tese como matéria-prima da mudança sonora, é a grande controvérsia dos trabalhos citados acima e, no que toca à diacronia, é também uma grande lacuna. No geral, ela é a característica mais comum em processos de redução e os trabalhos citados divergem justamente quanto à natureza e à variabilidade do sândi, mostrando que ainda há muito a ser investigado sobre esse processo.

6.2 A influência do acento no PB e o processo de sândi vocálico

Os processos de sândi vocálico no PB foram amplamente estudados por Bisol (1992, 1996, 2003). Além dela, Abaurre, Galves & Scarpa (1999), Tenani (2002, 2003), Collischonn (2011, 2012) *inter alia* também investigaram o processo sob o viés da Teoria Gerativa. Todos os trabalhos são unânimes em reportar três processos fonológicos recorrentes em juntura vocálica na fronteira de palavras no PB: elisão, degeminação e ditongação.

Segundo Bisol (1992), o português é uma língua sensível ao peso da sílaba final, ao encontro de acentos e à juntura de núcleos silábicos em fronteira vocabular. Caso a última sílaba não esteja acompanhada por acento ou pausa, a perda de um dos núcleos silábicos é inevitável e, consequentemente, a sílaba que o domina é ressilabificada e isso dá origem aos três processos exemplificados abaixo em (5).

(5) a. *como uvas* (komʊ#uvas) – ko.ˈmu.vas

b. *comi alho* (komi#aʎu) – ko.mj.ˈaʎu

c. *compra ervas* (kompɾɐ#ɛrvɐs) – kom.ˈpɾɛ.rvɐs

De acordo com Bisol (1992, 1996, 2003, 2004), em casos como em (5a), as vogais idênticas degeminam e a consoante se torna ataque de uma sílaba inicial; no fim, perde-se uma sílaba e a vogal é "encurtada", pois não há vogais longas no PB. No exemplo em (5b), cria-se um ditongo, entretanto não se perde nenhum segmento. Na elisão, em (5c), a consoante da sílaba final passa a ser o ataque da sílaba inicial da palavra seguinte; aqui, segundo a autora, perde-se uma sílaba, uma vez que o /a/ foi apagado.

Nos casos gerais da elisão, é importante salientar que há um encontro entre uma pós-tônica final com uma pré-tônica ou tônica. Aqui, cabe lembrar rapidamente a escala

acentual de Câmara Jr. (1969) para a tonicidade no PB, já citada na seção 1.3 da Parte 1 deste trabalho. Como se sabe, o inventário das vogais do PB é maior em posição tônica. Por outro lado, nas demais posições, as vogais estão sujeitas ao enfraquecimento e até a queda. Câmara Jr. baseia-se em diferenças de intensidade de força prosódica de sílabas acentuadas e não acentuadas para propor uma pauta prosódica escalar. Ele designa 3 e 2 para o acento (forte) primário e secundário, 1 para a pretônica e 0 para átonas após o acento. Aqui, no sândi externo, a mais fraca tende a cair.

Em trabalho recente sobre a aquisição do sândi no PB, Komatsu & Santos (2007), baseadas nas análises de Bisol (1992, 1996), mostram de forma geral a aplicação das regras de sândi no PB:

(i) A degeminação, por exemplo, não ocorre se a segunda sílaba carregar o acento, mas pode ocorrer entre duas sílabas fracas ou entre uma sílaba forte e uma fraca;

(ii) A ditongação, por outro lado, ocorre nas sequências de duas vogais em que uma das vogais é alta e não acentuada;

(iii) A elisão é bloqueada se a segunda sílaba é acentuada; a primeira vogal é, geralmente, o /a/ que é átona e apagada diante de vogais altas e médias.

Segundo Bisol (1992, 1996, 2003), a regra de elisão da vogal é aplicada em 94% das vezes, ou seja, a elisão das vogais é a mais recorrente. Segundo a maioria dos trabalhos citados acima, a elisão fica restrita ao apagamento da vogal /a/ em posição não acentuada de final de palavra, quando a palavra seguinte começa por vogal de qualidade diferente, e ocorre quando a presença de duas vogais em sequência fica sob o domínio de uma mesma sílaba. Do ponto de vista formal, um dos segmentos vocálicos é sempre "apagado" em função do choque dos dois núcleos silábicos. Assim, a vogal átona final da primeira palavra é sempre "apagada", e uma nova sílaba é formada a partir da junção do núcleo da sílaba átona final da primeira palavra com a vogal inicial da segunda palavra.[6]

[6] Não é o foco desde trabalho, mas vale salientar que Vigário (1998), para o PE, defende a hipótese de que a elisão corresponde a um processo de redução típico de palavras não acentuadas. No PE, devido à pauta prosódica das vogais, a elisão pode ocorrer dentro da palavra. Isso, aparentemente, não ocorre no PB.

6.3 O sândi sob o viés da FonGest: elisão de /a/ átono em juntura com vogal átona na visão de Albano *et al*. (1998) e Albano (1999)

Tanto quanto o desvozeamento, a elisão de /a/ átono em juntura com outra vogal átona[7], como visto, é outro processo considerado categórico por diversos estudos do PB. Basicamente, em termos de representação tradicional feita sobre a visão dos traços distintivos, temos a seguinte formalização em (6) abaixo.

(6) /a/ [-acento] → 0 / _# V [-acento]

A representação em (6) simplifica o processo do apagamento em contexto de fronteira e junção vocálica no caso da elisão. Albano *et al*. (1998) demonstraram que, em alguns casos em que a transcrição de oitiva trata como elisão, a vogal supostamente elidida deixa um rastro claro no sinal acústico. Assim, a pronúncia de "cara idoso" tem, pelo menos, três versões atestadas acusticamente pelos autores: uma com encontro vocálico nítido, uma muito próxima de "caridoso" e outra intermediária, em que os valores de F1 e F2 logo após a liberação do [R] evidenciam um breve resquício de [a]. Abaixo, a Figura 32 apresenta os espectros que mostram essa gradação.

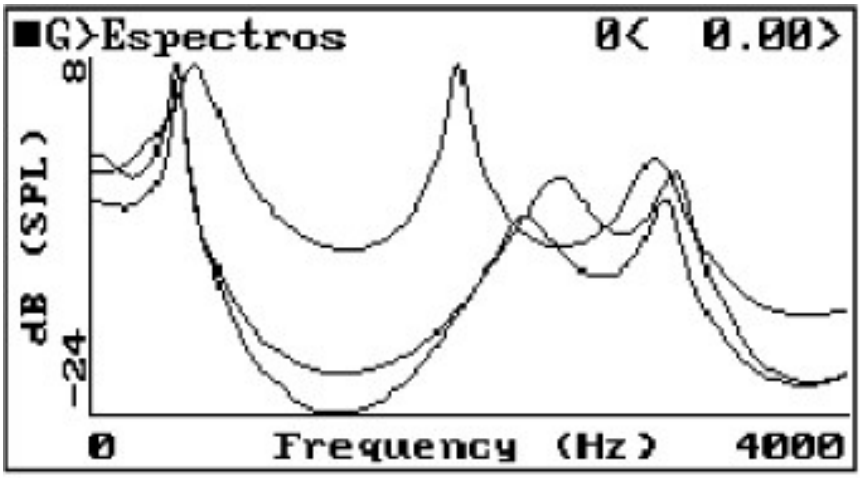


Figura 32 – Espectros de LPC após a liberação do [R] em três pronúncias (lenta, moderada e rápida)
Fonte: Albano (1999, p. 38)

[7] As palavras testadas por Albano *et al*. (1998) mostravam uma junção de uma átona final com uma pré-tônica, como por exemplo, "cara idoso".

Segundos Albano *et al*. (1998) e Albano (1999), na fala mais lenta, portanto, é possível depreender um encontro vocálico nítido e, na fala normal e rápida, essa distinção mais clara se perde, mas a vogal pode permanecer sobreposta no sinal da produção acústica.

A interpretação gradiente da elisão ganha ainda mais força quando se considera que ela interage com o processo de palatalização. A Figura 33 mostra o espectrograma da expressão "muita#idade" [mult'idad͡ʒi] pronunciada sem africação por um falante do dialeto carioca, no qual esse processo é padrão.

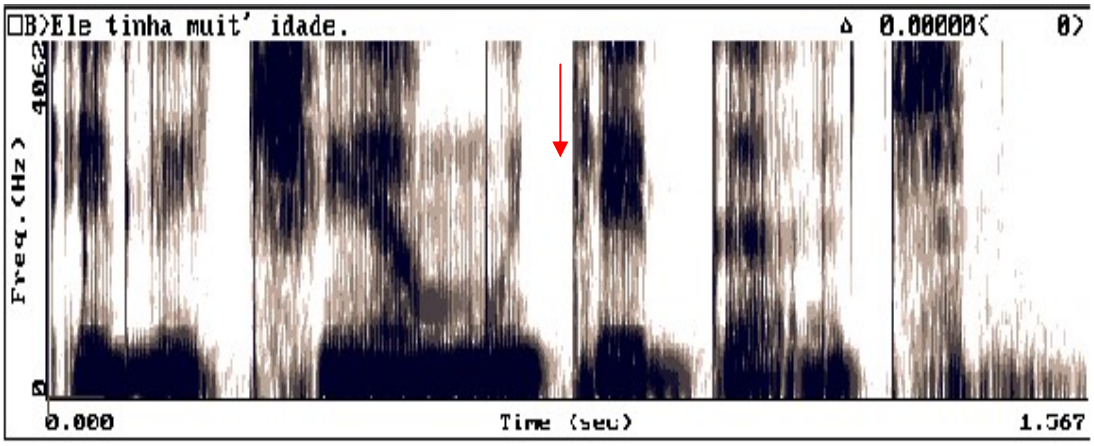


Figura 33 – Ausência de africação em /muit ˈidade/ pronunciada por um dos informantes do experimento de Albano *et al*. (1998). Nota-se a ausência de ruído fricativo na região indicada na seta, como reportado pelos autores.
Fonte: Albano (1999, p. 39)

Segundo a autora, enquanto a inibição da africação exigiria uma "ginástica teórica" de modelos gerativos, na análise da FonGest, proposta por Albano (1999), é o gesto oculto do /a/ que impede o contato do /t/ com o /i/, como mostra a Figura 34 a seguir. Na pauta gestual de "muita'idade", o gesto de /a/ é representado pela linha barra pontilhada.

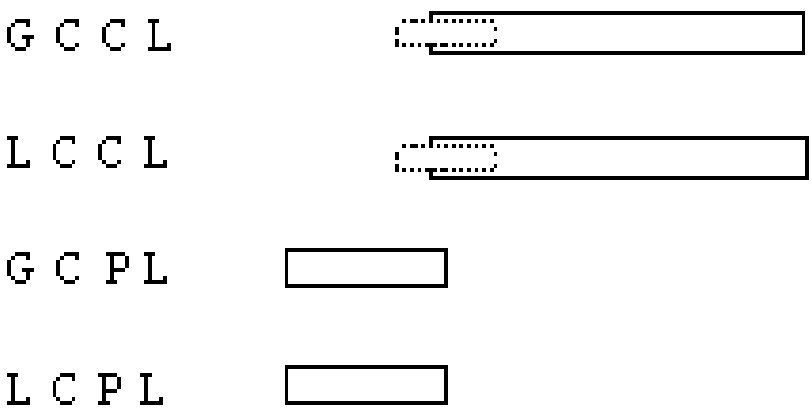


Figura 34 – Pauta gestual proposta por Albano (1999) para [mult'idad͡ʒi]
Fonte: Albano (1999, p. 39)

Depois de anos de pesquisas no PB e em outras línguas naturais, sabe-se que a gradiência reportada pelos autores supracitados é impossível de se representar em modelos fonológicos tradicionais, pois o único recurso disponível é associar /i/ à casa segmental já associada a /a/, o que não ocorre (Albano, 1999). Em contraste, a FonGest prevê que o gesto do /a/ pode diminuir e ocultar-se sob o gesto do /i/, o que explica os dados observados por Albano *et al*. (1998) e Albano (1999). No entanto, falta elucidar a diferença entre os casos com e sem rastros do /a/ no nível acústico.

Faz sentido, então, diante de dados como os expostos acima, investigar o processo de sândi como fenômeno gradiente de redução no PB, levantando hipóteses sobre o funcionamento geral da redução e destacando a particularidade desse processo. A discussão dos próximos capítulos segue nessa direção.

Capítulo 7

"Experimento de produção"

7.0 Introdução: experimentos de produção

Os dados que seguem descrevem e analisam os resultados de dois experimentos de produção controlados. O primeiro experimento, cujo procedimento é descrito na subseção 7.1.1.2, é uma tarefa de repetição contínua e de velocidade controlada que visava investigar a influência do aumento da velocidade na ocorrência ou não do sândi. O segundo experimento, cuja descrição encontra-se na subseção 7.2.1.1, reúne mais palavras e tem um número de repetições mais restrito. O objetivo era controlar melhor os contextos e permitir uma análise estatística mais robusta.

7.1 Experimento 1: tarefa de repetição contínua

7.1.1 Metodologia

7.1.1.1 Materiais e participantes

Dois participantes do sexo masculino foram selecionados para o Experimento-Piloto. A faixa etária (25-40 anos), a escolaridade (estudantes de graduação, neste caso) e o nível socioeconômico (classe média) foram os mesmos. Para esse experimento, o dialeto dos informantes não foi restringido, uma vez que o sândi é um processo que tem abrangência dialetal.

Para esse teste propriamente dito, a tarefa consistia na repetição de apenas uma palavra-alvo – "sincera idade" –, que é uma das palavras que compõe o *corpus* de palavras com sândi nesta tese.

7.1.1.2 Procedimento

O teste de repetição foi conduzido por meio de um metrônomo (aplicativo *online*), no qual era possível aumentar o ritmo das batidas gradativamente. O participante foi orientado a aumentar a velocidade no mesmo ritmo de batida do metrônomo e recebia um sinal sonoro e visual a cada aumento do passo das batidas. Com um total de 27 pontos de ocorrência por

sujeito, foram aplicadas seis fases de aumentos gradativos de velocidade, que partiram do passo normal, passando pelo moderado até chegar ao rápido.

7.1.1.3 Segmentação e medidas

Para a duração, as medidas obtidas nesse experimento foram as seguintes: a duração da palavra, a duração do encontro vocálico /a+i/ (rotulada como a#i), a duração do /i/ e da vogal /a/. Para essa última, as medidas foram feitas até onde era possível perceber transição formântica. A transição foi observada durante a segmentação dos dados e as medidas foram feitas automaticamente por meio de *script* adaptado para esse fim. Quando não havia mais transição entre /a/ e /i/, a duração do /a/ foi zero e só a duração do /i/ foi contabilizada. Além dessas, a medida dos formantes de /a/ e /i/ também foi tomada. As medidas também foram automáticas e seguiram o mesmo padrão de análise de medida de duração.

7.1.1.4 Análise estatística

Para o piloto, somente a estatística descritiva básica foi utilizada. A duração foi apresentada em dados brutos e por meio da linha de projeção logarítmica. Uma linha de tendência logarítmica é uma linha curva muito útil quando a taxa de alteração nos dados aumenta ou diminui rapidamente e depois se nivela. A fórmula da suavização é $Y = b0 + b1 * ln(X) + e$, na qual X representa a variável explicativa, Y indica a variável de resposta e (e) representa um erro aleatório.

7.1.2 Resultados

7.1.2.1 Duração

A Figura 35 a seguir apresenta a curva de tendência logarítmica da duração das três unidades analisadas, a saber: /a#i/ (porção a+i), /i/ e /a/, quando fosse possível identificar essa última vogal. Na mesma figura, é possível observar as linhas de oitiva e de transição (linhas verticais), que representam o limiar no qual era possível perceber a vogal auditivamente e o limite que marca o fim da transição vocálica.

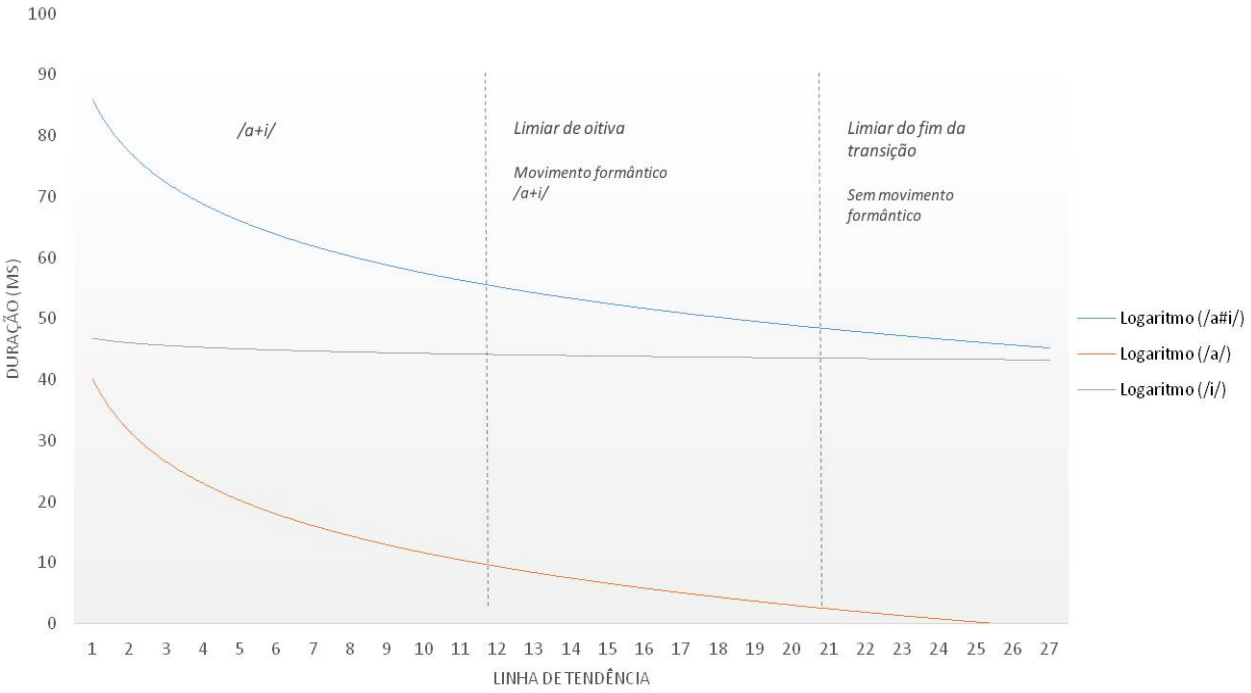


Figura 35 – Linha de tendência da duração de /a#i/, /i/ e /a/ durante a realização da tarefa de repetição. Note nas linhas verticais os limiares de oitiva e o limiar que marca o fim da transição entre as vogais.

No início da tarefa, como esperado, as fronteiras entre /a/ e /i/ estão bem definidas e é possível estabelecer bem a duração e cada um. A duração de /i/ é maior do que a do /a/ durante toda a fala considerada normal. Diferentemente do que predizem as regras de sândi, na fala normal não há nenhum apagamento da vogal. Mais do que isso, durante boa parte da tarefa, na qual a velocidade aumentava gradativamente, é possível estabelecer limites de oitiva e limites acústicos (transição formântica) entre /a/ e /i/.

A partir do limiar da oitiva, vista na Figura 35 acima, a porção /a+i/ só é identificada pela transição formântica. A oitiva é altamente prejudicada pela coarticulação entre /a+i/ e qualquer julgamento de transição é determinada pelo movimento de subida do F2. No fim, a partir do ponto onde os movimentos formânticos desapareceram, não é possível perceber mais nenhum resquício acústico da produção do /a/. Aparentemente, a partir desse momento a vogal /a/ foi apagada, restando, então, somente a vogal /i/ na juntura entre "sincer'idade". O aumento da velocidade parece ser um componente muito importante na redução e no apagamento do /a/.

É possível perceber por meio da linha de tendência apresentada acima, como também pelos resultados brutos (não suavizados pela linha de tendência) da Figura 36 abaixo,

que a tendência da duração das vogais e também da duração total da palavra é de queda gradativa e diretamente proporcional ao aumento da velocidade determinada na tarefa de produção.

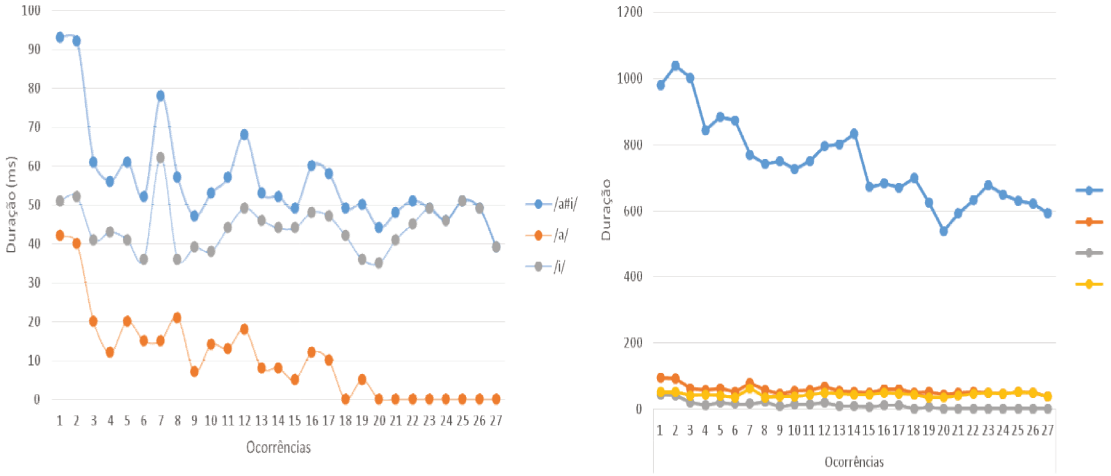


Figura 36 – Linha não suavizada da duração de /a#i/, /i/ e /a/ (à esquerda) e linha de duração da palavra (à direita)[1]

7.1.2.2 Formantes

Os formantes, analisados por meio da extração automática, mostram que, na maioria dos momentos, /a/ e /i/ apresentam diferença mínima, como podemos visualizar pela área de espaço vocálico da Figura 37 a seguir, na qual a grande maioria dos pontos (invisível justamente pela sobreposição) está concentrada na mesma região.

Na Figura 37, as únicas repetições que caem nas áreas esperadas do espaço vocálico de /a/ e /i/ são as quatro primeiras (círculo azul e vermelho, respectivamente), que parecem ter limites mais bem definidos entre as vogais. As demais caem em posições intermediárias que se sobrepõem parcialmente, aproximando-se cada vez mais de /i/.

[1] Como a duração da palavra é muito maior do que os segmentos da figura à esquerda, não foi possível colocá-las em uma mesma figura.

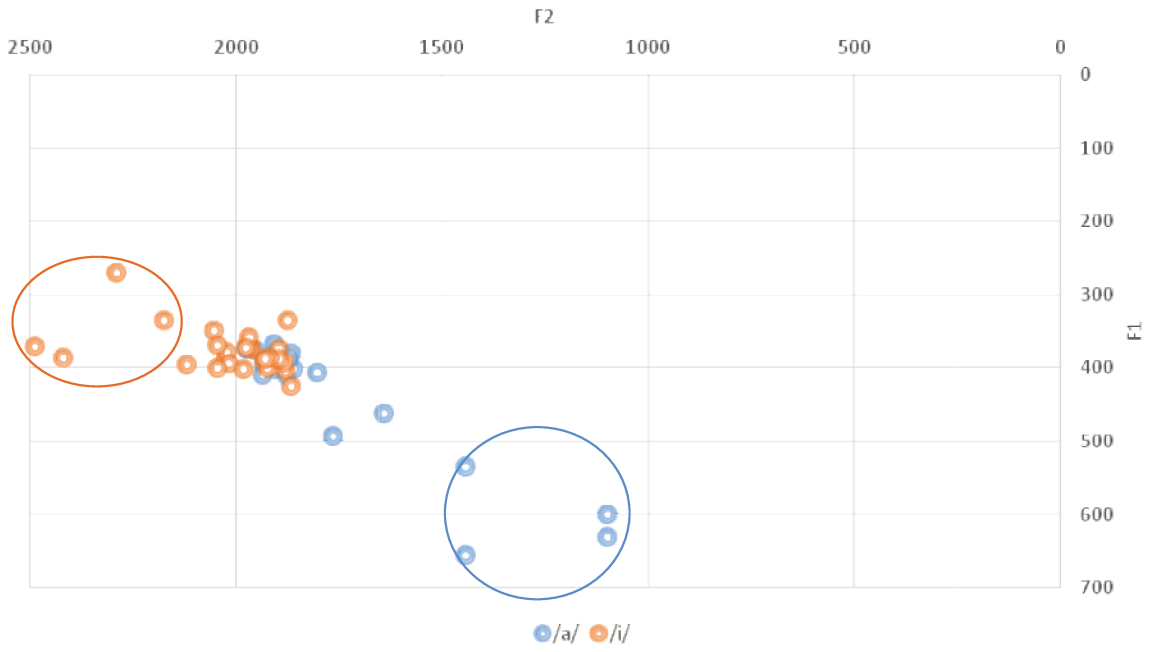


Figura 37 – Área de espaço vocálico das vogais /a/ e /i/. As áreas marcadas pelos círculos mostram os casos nos quais os valores médios dos formantes estavam dentro do esperado para /a/ (círculo azul) e /i/ (círculo laranja).

A tendência de F1 e F2 é de se estabilizarem, respectivamente, em torno de 400Hz e 2000Hz, o que torna a vogal resultante relativamente próxima de /i/. Essa tendência de estabilização fica mais clara na Figura 38 abaixo, que mostra, em linha, a evolução de F1 e F2 com o aumento da velocidade.

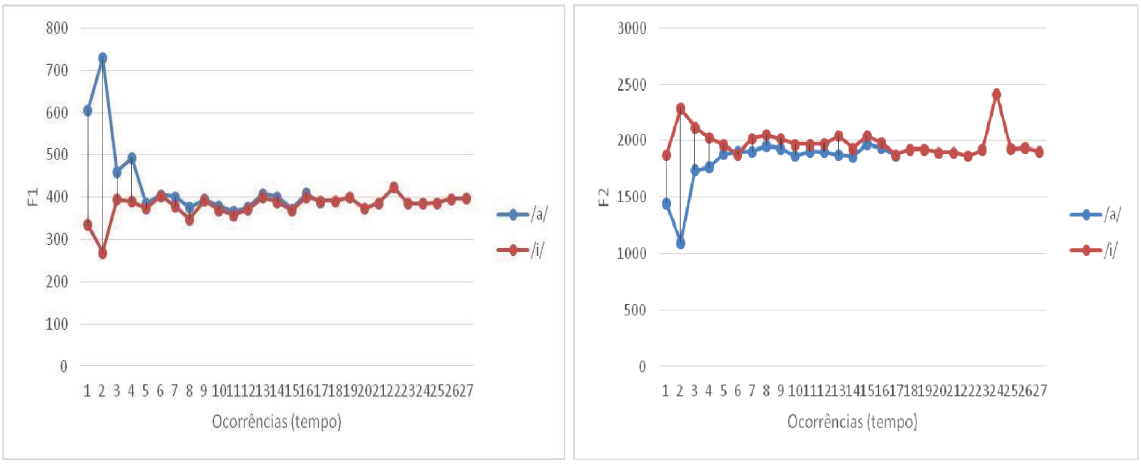


Figura 38 – Tendência das frequências de F1 (à esquerda) e F2 (à direita) durante a tarefa de repetição

O detalhe importante que as figuras acima revelam é que, enquanto as médias de F1 de /a/ e /i/ se estabilizam rapidamente na região de 400Hz, o F2 se mantém ligeiramente diferente até pouco mais da metade da tarefa de repetição.

A Figura 39 abaixo mostra dois exemplos retirados do Experimento-Piloto. Na primeira figura, é possível ver, claramente, a transição entre /a/ e /i/ na região de F2; para F1, por outro lado, a região de F1 fica estável. Na segunda, além de uma mudança radical da forma de onda, nenhuma transição formântica é vista.

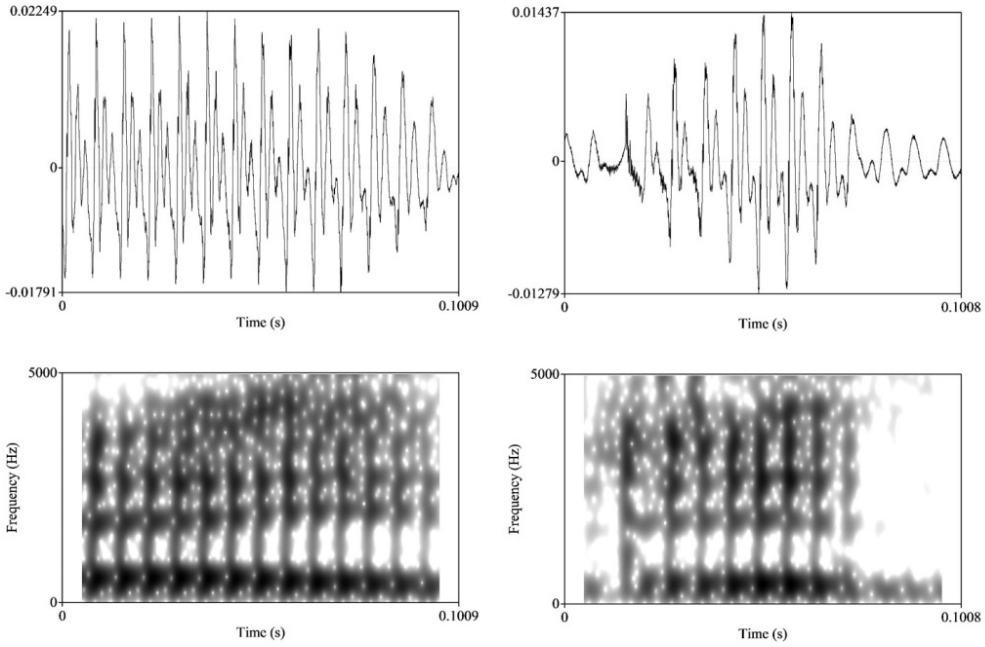


Figura 39 – Forma de onda e espectrograma de palavras retiradas de dois momentos diferentes da tarefa, normal e rápida

7.1.3 Discussão

Os resultados do Experimento 1 mostram que a variável velocidade é realmente relevante no processo de sândi vocálico e revelam que o aumento gradativo da velocidade é fator determinante na redução do /a/ e no possível apagamento da vogal na fala mais rápida. A redução da duração das vogais, além da redução da palavra como um todo, mostra que o aumento da velocidade de fala impõe um novo modo de coordenação sobre os gestos vocálicos que culmina da queda do /a/. Do ponto de vista metodológico, parece claro que essa variável, a velocidade, precisa ser investigada em um experimento maior e mais controlado.

Além da duração, as medidas dos formantes mostram resultados consideráveis. O F1 e o F2 das vogais, de maneira diferente, resultam na estabilização mais próxima dos

valores análogos ao /i/. Antes disso, porém, os valores de F1 de /a/, por exemplo, já são consideravelmente mais baixos, mesmo quando a vogal é realizada com duração relevante. Enquanto F1 diferencia-se menos, F2 parece se diferenciar por mais tempo.

De forma geral, sabe-se que os valores de F1 estão altamente relacionados com a elevação da língua na área na direção vertical do trato vocal e com o deslocamento da mandíbula. O F1 estabelece, por exemplo, a diferença entre vogais altas e baixas. Enquanto /i/ e /u/ têm os menores valores de F1, pois são as vogais mais altas e fechadas, /a/ tem os maiores valores de F1, pois é a vogal mais baixa e aberta. O F2, por sua vez, está relacionado com o deslocamento da língua na área horizontal do trato vocal e estabelece a diferença entre as vogais anteriores e as vogais posteriores. O /i/, por exemplo, é a vogal mais anterior e tem, portanto, o maior valor de F2.

Os valores de F1 parecem estar de acordo com o que sugere Perkell (1969), a saber, que a movimentação da mandíbula na sequência de /a/ para /i/ é mínima. A mandíbula, como sabemos, oscila para cima e para baixo durante a fala. Cada oscilação da mandíbula corresponde a uma sílaba e o núcleo corresponde à fase descendente do ciclo da mandíbula. No discurso normal, o comportamento da mandíbula é estabelecido. No entanto, para a fala mais rápida, mais casual, os movimentos da mandíbula são minimizados. Os valores de F2, por sua vez, sugerem que há uma pequena movimentação da língua na fala normal e em alguns casos da fala rápida, mas que essa movimentação é contida com o aumento da velocidade.

A constituição silábica também é importante aqui, uma vez que também influencia propriedades de magnitude gestual, sobreposição e variabilidade na produção da fala (Dell, Juliano & Govindjee, 1993; Byrd, 1996). Nas sequências CV, que era o caso do desvozeamento, C e V estão em fase[2] e o acoplamento é não só mais natural, mas também mais estável. Por outro lado, sequências que não estão em fase, como o encontro V-V heterossilábico do sândi, estão mais sujeitas ao refaseamento.

Somando os fatores citados acima, percebe-se que a tendência aparente da redução no sândi é de redução do /a/ na fala normal, sobreposição no início da fala rápida e queda total da vogal com o aumento mais radical da velocidade.

[2] Fase e antifase são os dois modelos básicos disponíveis para coordenação de ações motoras específicas.

O próximo experimento mantém a distinção da velocidade (normal *x* rápida), mas muda a tarefa a fim de controlar o ambiente experimental. A metodologia e os resultados estão expostos nas subseções seguintes.

7.2 Experimento de produção II

7.2.1 Metodologia

7.2.1.1 Materiais, participantes e procedimento

Quatro participantes do sexo feminino foram selecionados para o Experimento de Produção 2. A faixa etária (25-40 anos), a escolaridade (estudantes de graduação, neste caso) e o nível socioeconômico (classe média) foram novamente mantidos no mesmo padrão. Para esse experimento, assim como no anterior, o dialeto dos informantes não foi restringido.

No total, 16 palavras-alvo foram escolhidas para o experimento, de modo que oito são palavras nas quais o sândi poderia ocorrer e as oito restantes são seus pares mínimos sem ocorrência de sândi, como podemos observar no Quadro 6 abaixo. A frase-veículo foi produzida com a mesma estrutura e número de sílabas para cada palavra-alvo. Além disso, todas as frases foram construídas com algum sentido contextual, como por exemplo, "Ele chegou à paternidade agora". A lista completa das frases-veículo pode ser vista no Anexo 2.

Quadro 6 – Palavras-alvo usadas no Experimento de Produção 2

Vogais-alvo	**Palavras-alvo**			
	Palavra	*Transcrição*	Palavra	*Transcrição*
/a#i/	sincera idade	*[sĩˈsɛɾa#iˈdadʒ͡ɪ]*	sinceridade	*[sĩsɛɾiˈdadʒ͡ɪ]*
	pura idade	*[ˈpuɾa#iˈdadʒ͡ɪ]*	puridade	*[puɾiˈdadʒ͡ɪ]*
	materna idade	*[maˈteɣna#iˈdadʒ͡ɪ]*	maternidade	*[mateɣniˈdadʒ͡ɪ]*
	cara idoso	*[ˈkaɾa#idosʊ]*	caridoso	*[kaɾiˈdosʊ]*
	fraterna idade	*[fɾaˈtɛɣna#iˈdadʒ͡ɪ]*	fraternidade	*[fɾatɛɣniˈdadʒ͡ɪ]*
	paternidade	*[paˈtɛɣna#iˈdadʒ͡ɪ]*	paternidade	*[patɛɣniˈdadʒ͡ɪ]*
	rara idade	*[ˈɣaɾa#iˈdadʒ͡ɪ]*	raridade	*[ɣaɾiˈdadʒ͡ɪ]*
	para idade	*[ˈpaɾa#iˈdadʒ͡ɪ]*	paridade	*[paɾiˈdadʒ͡ɪ]*

Cada sujeito leu as frases cinco vezes em velocidade normal e rápida em seções diferenciadas, ou seja, uma repetição só para a fala normal e outra para a gravação de palavras em taxa rápida. Para diferenciar as velocidades, uma gravação modelo prévia foi feita com o auxílio de um metrônomo. Essa gravação, realizada por uma musicista bem familiarizada com o uso do metrônomo, foi utilizada como um modelo de fala normal e rápida. O sujeito, então, ouvia os modelos no início de cada repetição. O objetivo da gravação era fazer o sujeito perceber a diferença entre a taxa normal e rápida.

7.2.1.2 Segmentação e medidas

Para o Experimento 2, as medidas obtidas foram a duração da porção da união vocálica /a+i/ (rotulada como a#i) e a medida dos formantes de /a/ e /i/. Para a duração, já que o alvo era /a+i/, as medidas foram feitas por meio de medidas automáticas.

Para os formantes, as medidas foram feitas levando-se em consideração a existência ou não de transição formântica. Se não houvesse nenhum sinal de transição, o sinal era rotulado apenas como /i/. A medida dos formantes exigiu, então, uma análise qualitativa e quantitativa, a qual denominamos "*quase*-automática". A transição foi observada e segmentada manualmente e as medidas foram feitas automaticamente por meio de *script* (Anexo 3) adaptado para esse fim, isto é, buscar os rótulos específicos e medir os formantes no ponto médio do rótulo em questão. Essa opção foi necessária, pois a duração do /a/ era extremamente variável e o *script* automático estava passível de erros.

7.2.1.3 Análise estatística

Além da estatística descritiva e da contagem de ocorrências, o teste ANOVA *One-Way* foi utilizado para comparar valores de duração e, também, os valores dos formantes para os casos investigados.

7.2.2 Resultados

7.2.2.1 Ocorrências

Os resultados a seguir dizem respeito à aparente ocorrência de sândi nos dados analisados, ou seja, casos nos quais a vogal não deixa rastros aparentes na inspeção

qualitativa. Como esperado e como foi visto no experimento anterior, na fala normal os casos em que é possível identificar a existência da vogal /a/ superam 70% das ocorrências, como pode ser observado na Figura 40a a seguir.

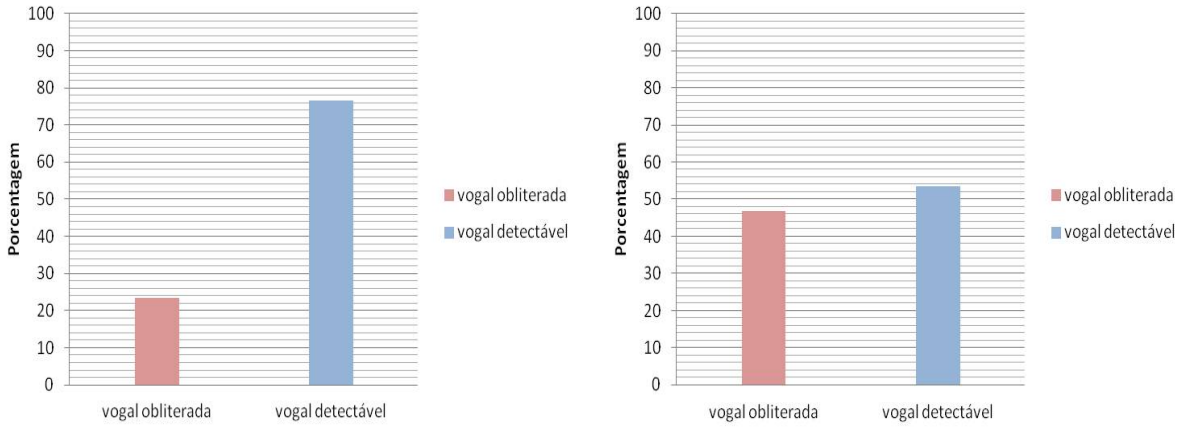


Figura 40 – Porcentagem de ocorrências de possíveis apagamentos e realização em casos de encontro de vogais na fala normal (à esquerda) e rápida (à direita)

Na fala rápida (Figura 40b), por outro lado, a ocorrência de possíveis apagamentos (possíveis diante dos parâmetros analisados) é maior. A taxa de ocorrência de sândi e de não sândi é quase a mesma.

7.2.2.2 Duração

Abaixo, apresentamos os resultados da duração para os contextos investigados. As Figuras 41 e 42 mostram os resultados para /a+i/ *x* /i/ ("cara idoso" *x* "caridoso", por exemplo) para a fala normal e rápida.

A Figura 41 mostra que a diferença entre /a+i/ e /i/ na fala normal é relevante e significativa ($F = 153.4$, d.f. = 2, $p < 0.001$). Na média, os valores de /a+i/ foram de aproximadamente 105 ms, enquanto /i/ teve a duração média de aproximadamente 65 ms. A diferença nos valores – alta, considerando o efeito do sândi na vogal /a/ –, era esperada, uma vez que na fala normal a coordenação entre /a/ e /i/ pode ser feita com mais controle sobre a articulação das vogais.

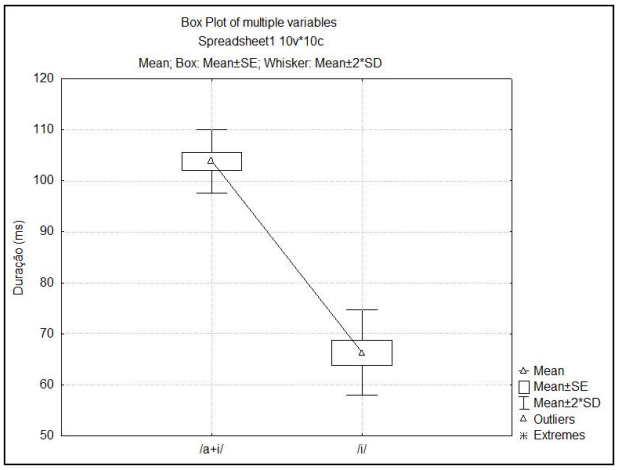


Figura 41 – Duração da sequência /a+i/ e /i/ em condição de taxa normal de fala

Enquanto a diferença para a fala normal já era esperada, os dados referentes à fala mais rápida foram surpreendentes: há também uma diferença significativa entre /a+i/ e /i/ na fala rápida ($F = 23.8$, d.f. = 2, $p < 0.01$). Se considerarmos que quase 50% das ocorrências reportadas aqui não apresentavam a mesma transição entre /a/ e /i/ da fala normal, o resultado da Figura 42 mostra que, mesmo nos casos nos quais o sândi reduz radicalmente o /a/, a duração permanece maior do que /i/.

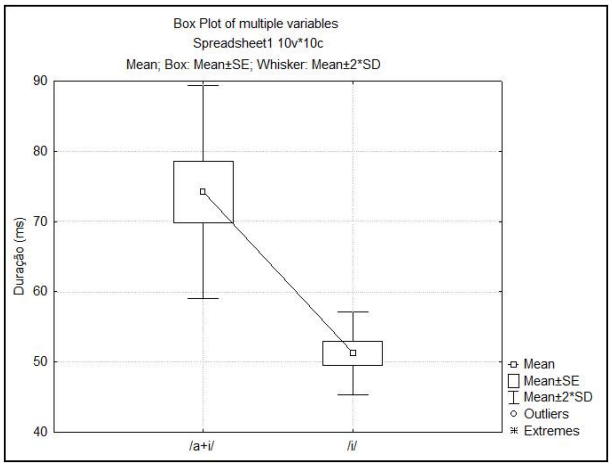


Figura 42 – Duração da sequência /a+i/ e /i/ em condição de taxa rápida de fala

7.2.2.3 Formantes

Os valores médios dos formantes são apresentados nas Figuras 43 e 44 abaixo.

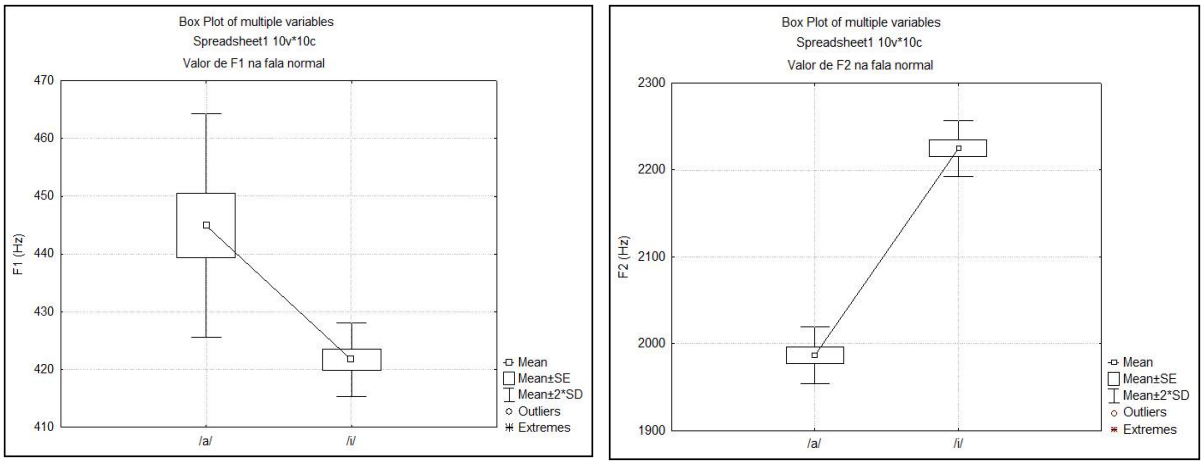


Figura 43 – Valores de F1 (à esquerda) e F2 (à direita) para /a/ e /i/ em sequências /a+i/ em condição de taxa normal de fala

Para a fala normal, os valores médios de F1 e F2 mostram-se estatisticamente significativos. Para F1, que está situado à esquerda na Figura 43, a diferença entre as médias de /a/ e /i/ é mínima (aproximadamente, é de 30 Hz), mesmo que a diferença seja significativa ($F = 15.5$, d.f. = 2, $p = 0.0179$). Esses resultados vão ao encontro dos resultados apresentados no Experimento 1, que já apontavam uma diferença pequena no F1 da transição /a+i/.

Para F2, os resultados também seguem a tendência já apontada no experimento anterior, ou seja, os valores de F2 também são significativos ($F = 325.3$, d.f. = 2, $p = 0.0005$) e mostram que há, em contexto de sândi, uma subida de 300 Hz da realização do /a/ até o *release* do /i/.

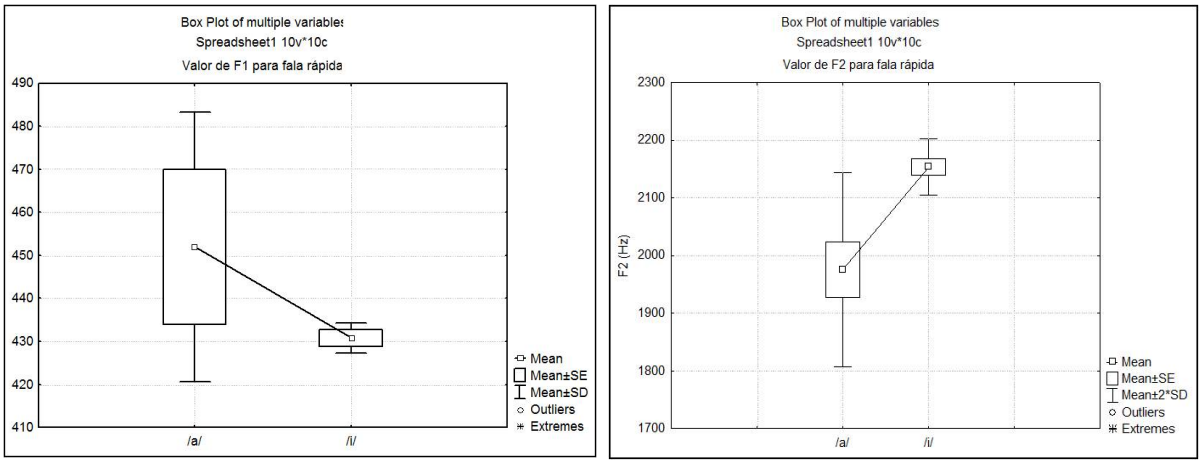


Figura 44 – Valores de F1 (à esquerda) e F2 (à direita) para /a/ e /i/ em sequências /a+i/ em condição de taxa rápida de fala

Quando nos voltamos para a fala rápida, os resultados tomam uma forma diferente dos resultados da fala normal. Para F1, parece não haver diferença entre /a/ e /i/ na sequência /a+i/, isto é, não há transição dos formantes na região de F1. Para F2, por outro lado, mesmo com as condições impostas pelo aumento da velocidade, a diferença constatada na fala normal se mantém ($F = 12.5$, d.f. = 2, $p = 0.025$). Há uma subida de F2 como esperado nas transições entre /a/ e /i/, mas a subida é reduzida.

7.2.3 Discussão

O Experimento 2, diferentemente do Experimento 1, não impôs aos sujeitos analisados condições extremas de taxa de velocidade. Sendo assim, nessas condições, os resultados mostrados são aqueles que mais se aproximam das condições reais de fala e merecem nosso destaque.

O primeiro experimento mostrou que, da fala de taxa normal à fala de taxa extremamente rápida (condições finais da tarefa), a duração e os valores de F1 e F2 de /a+i/ caminham gradativamente para a estabilização na direção de /i/. Nas condições da tarefa atual, na qual o falante determinava a taxa a partir de um modelo, os resultados mostram-se diferentes.

Para a duração, os resultados mostram que, na taxa de fala normal e rápida, há uma diferença significativa entre /a+i/ e /i/. Tais números causam surpresa, pois, diante dos resultados do primeiro teste, o esperado seria que não houvesse diferenças estatísticas na fala mais rápida. Na fala normal, como já havia sido mostrada por Albano (1999, 2001), a ocorrência do apagamento de /a/ no sândi é menos provável e as diferenças encontradas aqui corroboram essa ideia.

Os valores de F1 e F2 seguiram exatamente o mesmo padrão descrito na tarefa anterior: ambos apresentaram diferenças significativas entre /a/ e /i/ na fala normal e não houve diferenças para F1 na fala de taxa rápida. É importante, então, recuperarmos os resultados de Perkell (1969), que mostram que os valores de F1 referem-se à movimentação da mandíbula, que, na sequência de /a/ para /i/ na fala rápida, é mínima. Como também sugere Perkell (1969), os valores de F2, que foram diferentes e significantes também na fala mais rápida, parecem indicar que há uma pequena movimentação da língua na fala normal e em alguns casos da fala rápida, mas que essa movimentação é refreada com aumento da velocidade.

7.2.3.1 Hipóteses

Diante dos resultados acústicos reportados acima, já é possível levantar hipóteses sobre o encontro de vogais em fronteira e a realização do sândi no PB. Três dos resultados apresentados acima são os apoios da nossa hipótese, a saber: (i) a transição de F2 em fala normal e rápida; (ii) a duração maior de /a+i/ na fala normal; (iii) a duração maior de /a+i/ na fala rápida mostrada no Experimento de Produção 2.

Em primeiro lugar, é possível afirmar que a ocorrência categórica de sândi na fala normal é improvável, como já argumentava Albano (1999, 2001). A sobreposição existente no encontro de /a+i/ na fala normal não parece ser capaz de eliminar a vogal já mais centralizada, como mostra a representação da Figura 45 a seguir. Se isso se mostrar verossímil, os dados aerodinâmicos e de percepção mostraram pistas de existência da vogal mais claramente do que na fala rápida. Esse, por sua vez, parece ser o grande desafio.

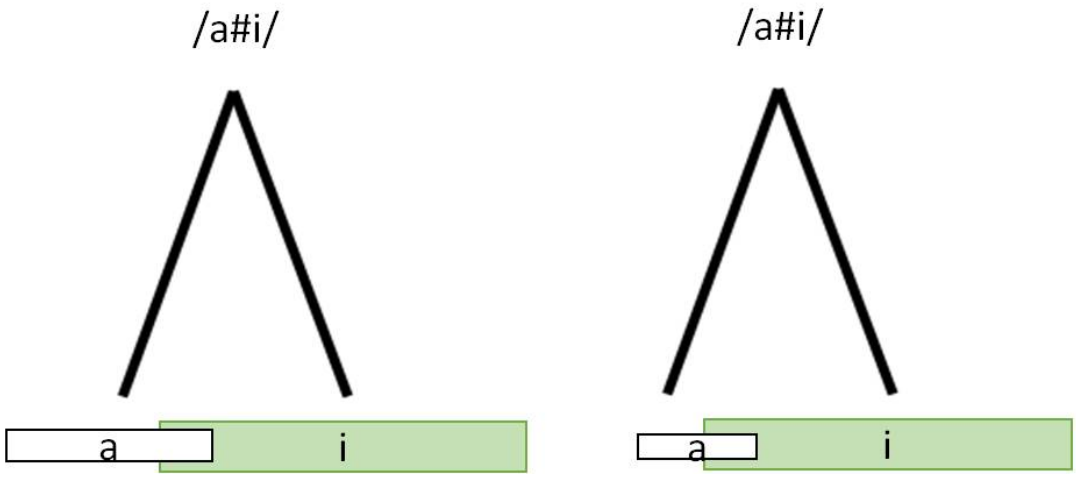


Figura 45 – Representação de hipótese do encontro de vogais em fronteira para a fala normal. As vogais em questão estão em antifase e não se trata de representação silábica.

Como mostramos no Experimento 1, o aumento da velocidade parece realmente levar à queda da vogal em muitos casos. Por outro lado, no segundo experimento, a duração de /a+i/ também foi significativa nesse contexto. Logo, para essa condição, temos duas possibilidades: (i) a vogal /a/ é extremamente reduzida, mas mantém-se no movimento mínimo dos articuladores; (ii) a vogal /a/ é completamente apagada da sequência /a+i/, resultando no sândi da vogal (ver Figura 46 abaixo).

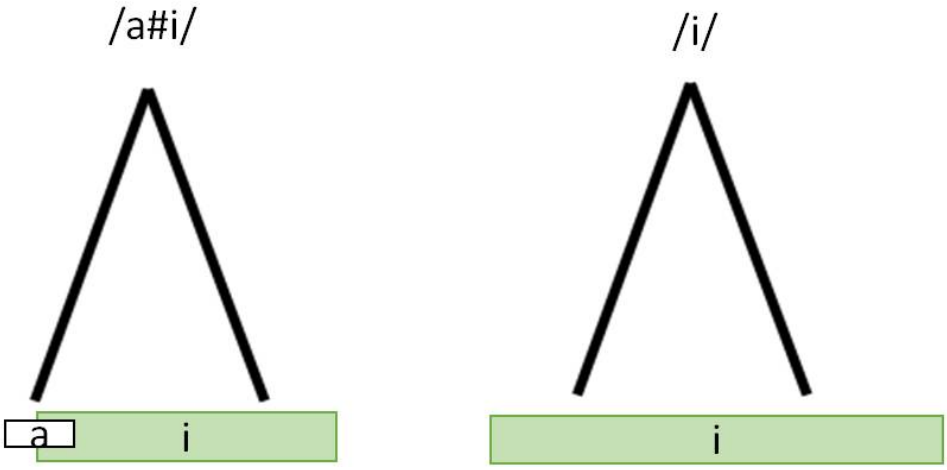


Figura 46 – Representação de hipótese do encontro de vogais em fronteira para a fala rápida

Dados complementares de análise aerodinâmica e de percepção vão nos ajudar a entender melhor o mecanismo envolvido na transição /a+i/ e o que exatamente diferencia /a+i/ de /i/ na fala rápida. No próximo capítulo, apresentamos esses dados.

Capítulo 8

"Teste Aerodinâmico"

8.0 Teste aerodinâmico: fluxo de ar oral

8.1 Teste aerodinâmico

8.1.1 Metodologia

8.1.1.1 Materiais, participantes e procedimento

O teste aerodinâmico idealizado para o sândi foi realizado com três participantes ($S1_a$...$S3_a$). Além das palavras gravadas no Experimento de Produção 2, mais dois contextos foram adicionados à investigação. Assim, além de /a+i/, foram gravadas sequências de /a+u/ e /i+a+e/, de acordo com as palavras listadas no Quadro 7 a seguir.

Novamente, apenas participantes do sexo feminino foram gravadas para manter a comparabilidade com o experimento anterior e os demais experimentos. O procedimento também seguiu o padrão do experimento anterior. Cada sujeito leu as frases-alvo três vezes em taxa rápida e normal. Todas as palavras-alvo foram gravadas em frases-veículo contextualizadas e estão listadas no Anexo 2.

Quadro 7 – Palavras-alvo usadas no Experimento Aerodinâmico

Vogais-alvo	**Palavras-alvo**			
	Palavra	Transcrição	Palavra	Transcrição
/a#i/	sincera idade	[sĩˈsɛɾa#iˈdadʒ͡ɪ]	sinceridade	[sĩsɛɾiˈdadʒ͡ɪ]
	pura idade	[ˈpuɾa#iˈdadʒ͡ɪ]	puridade	[puɾiˈdadʒ͡ɪ]
	materna idade	[maˈtɐɣna#iˈdadʒ͡ɪ]	maternidade	[mateɣniˈdadʒ͡ɪ]
	cara idoso	[ˈkaɾa#idosʊ]	caridoso	[kaɾiˈdosʊ]
	fraterna idade	[fɾaˈtɐɣna#iˈdadʒ͡ɪ]	fraternidade	[fɾatɐɣniˈdadʒ͡ɪ]
	paternidade	[paˈtɐɣna#iˈdadʒ͡ɪ]	paternidade	[patɐɣniˈdadʒ͡ɪ]
	rara idade	[ˈɣaɾa#iˈdadʒ͡ɪ]	raridade	[ɣaɾiˈdadʒ͡ɪ]
	para idade	[ˈpaɾa#iˈdadʒ͡ɪ]	paridade	[paɾiˈdadʒ͡ɪ]
/a#u/	ata urdida	[ˈata#ˈuɣdidɐ]	aturdida	[atˈuɣdidɐ]
/i+a+u/	pia elite	[ˈpia#eˈlitʃ͡ɪ]	pielite	[pieˈlitʃ͡ɪ]

8.1.2 Materiais

O EVA2 foi novamente a ferramenta utilizada para captar o fluxo de ar oral nesta tarefa (Ghio & Teston, 2004). O EVA2 permite a gravação simultânea do som e do fluxo de ar e exige a utilização de um "bocal" ligado ao suporte mecânico do microfone. Mais uma vez, os sensores de pressão foram manualmente calibrados e foram verificados para cada sujeito e entre todas as repetições. O Phonedit foi o *software* usado para a gravação e também para a análise dos dados.

8.1.3 Resultados

As figuras a seguir mostram os sinais sincronizados da forma de onda e do fluxo de ar oral. No exemplos abaixo, mostraremos os dados que foram recorrentes nas repetições e podem iluminar o processo de sândi no PB.

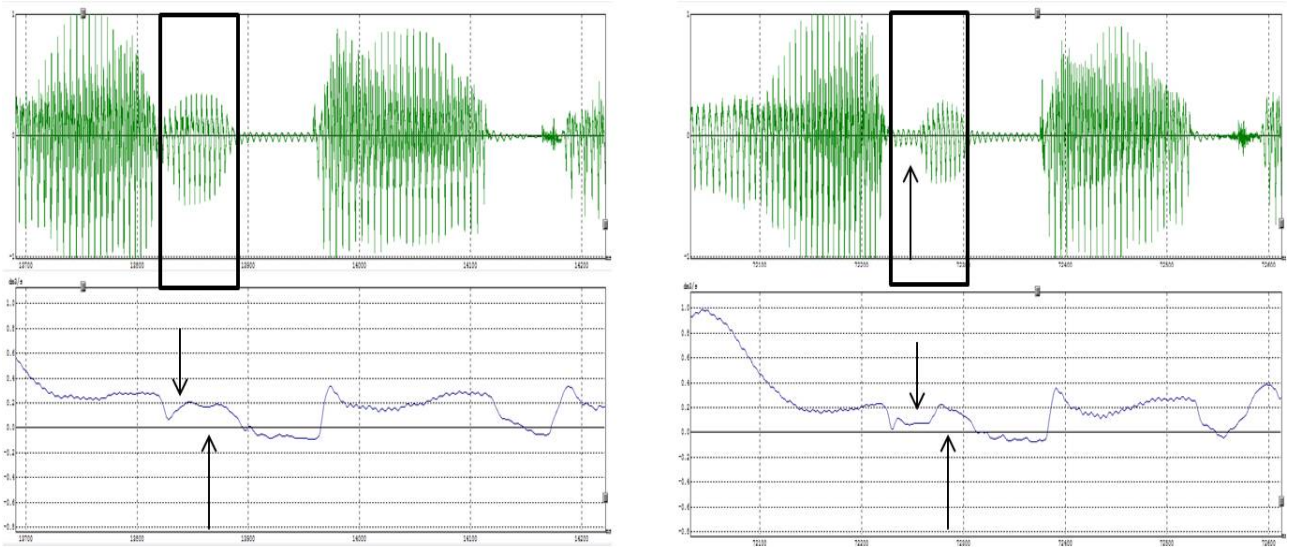


Figura 47 – Forma de onda (acima) e fluxo de ar oral (abaixo) de “raridade” (à esquerda) e “rara idade” (à direita) na fala normal. O destaque da forma de onda indica os alvos /i/ e /a+i/, respectivamente. As setas indicam as diferenças encontradas no fluxo de ar.

Na Figura 47, as setas à esquerda mostram a sequência de fluxo de ar para a sílaba /ri/ de “raridade”. É possível perceber um queda de fluxo de ar apontada pela primeira seta, que é caracteristica do tap, e depois uma soltura do ar que está diretamente relacionada à produção da vogal /i/. Os limites entre /r/ e /i/ não estão bem definidos e a vogal tem duração e amplitude bem evidentes. Na figura à direita, o quadro destaca a sequência /ra#i/ em “rara idade”. Prontamente, é possível perceber um limite bem estabelecido entre /r/ e a vogal seguinte. Assim como na figura à esquerda, também há uma queda pequena do fluxo de ar que se refere ao tap /r/. A diferença entre as figuras está na região apontada pela primeira seta à direita. Ali é possível perceber uma soltura pequena do fluxo e só depois uma soltura mais intensa que está relacionada à produção da vogal /i/. Esse tipo de soltura foi recorrente nos dados e parece estar relacionada a uma produção afrouxada da vogal /a/ em velocidade normal e rápida.

O padrão foi exatamente o mesmo para a fala rápida, considerando a diminuição do tempo de soltura do /a/ que aqui interpretamos como análoga à soltura de ar para um /a/. A seguir, na Figura 48, mostramos os casos de “raridade” e “rara idade” para a fala rápida. As seta indicam os pontos de semelhança entre as fala normal e rápida.

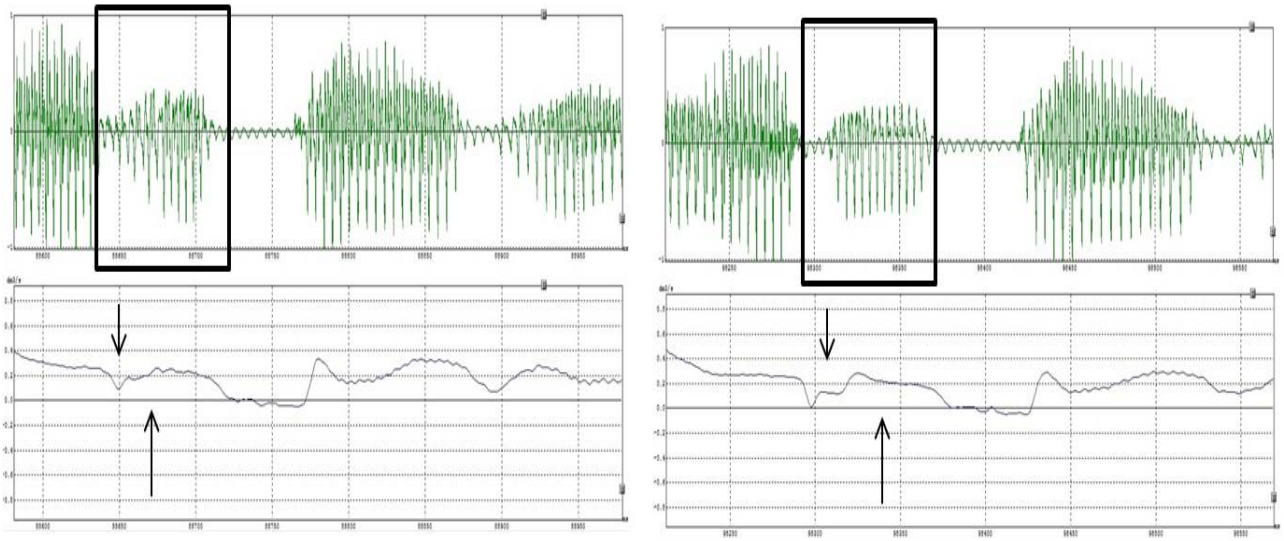


Figura 48 – Forma de onda (acima) e fluxo de ar oral (abaixo) de “raridade” (à esquerda) e “rara idade” (à direita) na fala rápida. O destaque da forma de onda indica os alvos /i/ e /a+i/, respectivamente. As setas indicam as diferenças encontradas no fluxo de ar.

Na fala rápida também houve casos no qual a vogal /a/ aparentemente não deixou pistas no sinal de fluxo de ar. Das palavras listadas no Experimento Aerodinâmico, “pia elite” foi aquela que não apresentou nenhuma diferença tanto na fala rápida quanto na fala normal. Abaixo, na Figura 49, mostramos o exemplo de “pielite” e “pia elite”. Aparentemente, não há qualquer diferença entre a sequência de /i+a+e/, esperada para “pia elite”, e /i+e/, provavelmente devido ao maior número de vogais em encontro.

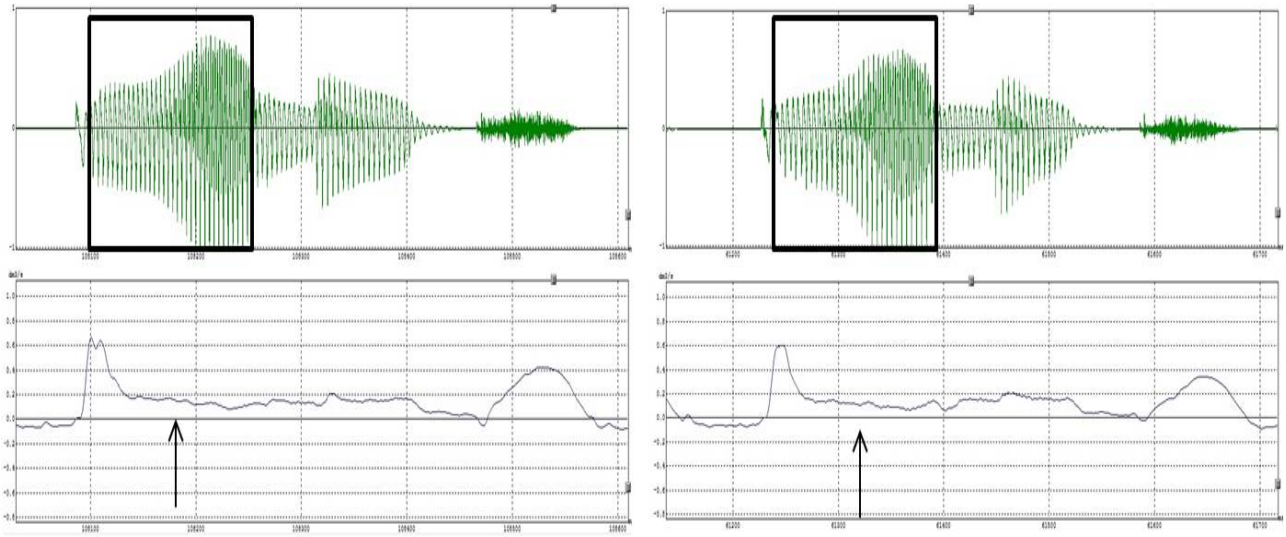


Figura 49 – Forma de onda (acima) e fluxo de ar oral (abaixo) de “pielite” (à esquerda) e “pia elite” (à direita) na fala rápida. O destaque da forma de onda indica os alvos /i+e/ e /i+a+e/, respectivamente. As setas indicam as diferenças encontradas no fluxo de ar.

8.1.4 Discussões

Os dados da subseção anterior visavam investigar quais os correlatos aerodinâmicos da produção da vogal /a/ em contexto de sândi. Os resultados dos testes aerodinâmicos, como esperado, vão ao encontro dos resultados acústicos.

Ao retomarmos as Figuras 47 e 48, tanto na fala normal quando na fala rápida, é possível identificar, imediatamente antes a produção do /i/, uma soltura breve do ar que é correlata à brevidade de produção da vogal /a/. Aparentemente, no sândi por elisão, o gesto do /a/ é reduzido e pode ocultar-se sob o gesto do /i/, o que explica os dados observados aqui e também corrobora a análise feita por Albano *et al*. (1998) e Albano (1999).

Muito possivelmente, a vogal /a/, em posição mais fraca na juntura entre as vogais, é reduzida ao máximo quando é imposta sobre ela o aumento da velocidade. Isso se dá porque a configuração do trato vocal para a vogal alta /i/ se sobrepõe à produção da vogal /a/. Em muitos contextos, a produção plena do /a/ só é possível se uma pausa é introduzida entre /a/ e /i/. Se isso não acontece, não há tempo suficiente para se alcançar a abertura da mandíbula necessária para a produção do /a/. Nesses casos, a vogal pode sim ser totalmente apagada, como parece ser o caso do exposto na Figura 49.

Assim, tal como defendido por Albano *et al*. (1998), a chamada elisão em contexto de sândi no PB é, na realidade, a realização de um gradiente de produção que vai da pronúncia evidente de /a+i/ (“cara idoso”) e passa por um leque de produções intermediárias dependendo do grau de sobreposição do /a/ até a produção de algo muito próxima de “caridoso”. Nesse último caso, a influência de fatores associados (posição prosódica fraca e aumento da velocidade) pode levar à elisão da vogal. O próximo passo, já iniciado no âmbito experimental, é a averiguação de dados articulatórios que podem fornecer pistas mais concretas da real influência dos fatores supracitados na redução ou elisão do /a/ em sândi no PB.

Capítulo 9

"Diretrizes para pesquisas futuras no sândi por elisão"

9.0 Breve visão dos dados

Tal como fizemos no Capítulo 5 da Parte 1 da tese, antes da conclusão, apresentamos dados de trabalhos ainda em preparação ou que ainda exigem uma coleta maior de dados e já são consequências dos dados mostrados na Parte 2 sobre a suposta elisão no sândi no PB. Aqui, em primeiro lugar, apresentaremos dados do teste de discriminação que foi indicação da banca de exame de qualificação[1]. Além disso, apresentamos dados iniciais de ultrassom para o sândi que também já é seguimento dos dados analisados na tese e também é fruto da parceria entre o LAFAPE-Unicamp e o Illinois Phonetics & Phonology Lab[2].

9.1 Percepção: teste de discriminação

9.1.1 Breve descrição da metodologia

A presente seção apresenta os resultados preliminares de um experimento de percepção, mais especificamente, um teste de discriminação que investiga se os sujeitos são capazes de diferenciar as possíveis ocorrências da vogal /a/ em contexto de sândi de seus pares de ocorrência apenas de /i/. De forma geral, a tarefa consiste na tentativa de diferenciar "cara idoso" de "caridoso", por exemplo. Esperava-se que os sujeitos fossem capazes de diferenciar as palavras-alvos na fala normal e que a fala rápida ocasionasse mais erros.

Quatro ouvintes do sexo feminino ($S1_c$...$S4_c$) foram selecionadas para esse experimento. Foram mantidos constantes a faixa etária (25-40 anos), a escolaridade (estudantes de graduação, neste caso) e o nível socioeconômico (classe média).

Para o teste propriamente dito, dois grupos de estímulos – palavras com sequência /a+i/ e palavra com /i/ – foram extraídos de nossas sentenças-alvo originais e tocados de forma isolada do contexto em uma tarefa de discriminação (cf. Quadro 7 no Capítulo 8). A escolha dos estímulos foi aleatória. Todos os estímulos tiveram a intensidade normalizada e um *script* do Praat foi utilizado para o teste (Anexo 4).

---

[1] Agradeço ao Prof. Dr. Wilmar D'Angelis pela indicação.
[2] O projeto em questão foi financiado pelo Lemann Institute for Brazilian Studies – UIUC.

9.1.2 Resultados

Os dados apresentados na Figura 50 abaixo mostram os resultados dos dois tipos de estímulos aos quais os sujeitos foram submetidos. Levando-se em conta os resultados acústicos, o esperado para o teste de discriminação era um número alto de acertos na fala normal e um aumento de erros na fala rápida.

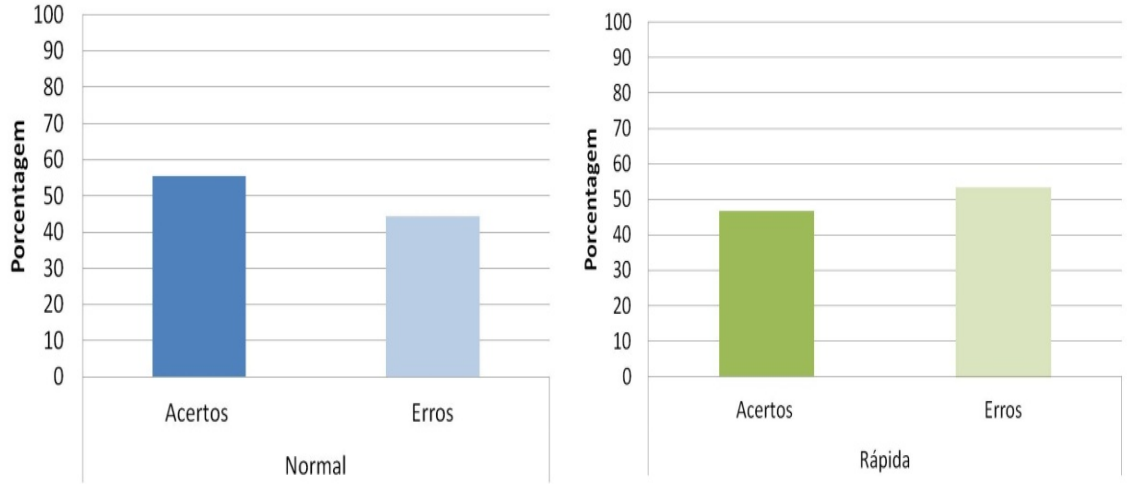


Figura 50 – Porcentagem de acerto e erro para os estímulos de taxa normal e rápida

Até então, ainda diante de dados preliminares, os sujeitos analisados mostravamm-se indecisos na tarefa de discriminação. Na fala normal, a taxa de acerto era maior, mas não era uma diferença relevante. Na fala rápida, o número de erros aumentou, mas a diferença também era pequena.

O interessante dos dados é a diferença da taxa de acerto entre as palavras. Levando-se em consideração a fala normal, "fraterna idade *x* fraternidade" e "sincera idade *x* sinceridade" tiveram as maiores taxa de reconhecimento (94% e 88%, respectivamente).

Tabela 2 – Porcentagem de acerto por palavra com velocidade de fala normal

X	caridoso	fraternidade	maternidade	puridade	sinceridade
fraterna#idade	94%				
sincera#idade		88%			
cara#idoso			61%		
pura#idade				33%	
materna#idade					27%

Para os estímulos de taxa rápida, exatamente as mesmas palavras tiveram as maiores taxa de acerto com 66% e 72%, respectivamente. Por outro lado, "pura idade *x* puridade" teve a maior taxa de erro (95%).

Tabela 3 – Porcentagem de acerto por palavra com velocidade de fala rápida

X	caridoso	fraternidade	maternidade	puridade	sinceridade
materna#idade	88%				
sincera#idade		72%			
fraterna#idade			66%		
cara#idoso				22%	
pura#idade					5%

É necessário salientar que o efeito do número de ocorrências das palavras é quase inevitável no teste de percepção.

Além do teste de discriminação, um teste de identificação será necessário para entender melhor o julgamento dos falantes sobre os dados de sândi. Os dados acima mostram o padrão na direção esperada, mas as diferenças ainda não estão na proporção que os dados acústicos mostram.

9.2 Dados articulatórios (prévia de Meneses *et al.*, em preparação)

O mesmo aviso dado na seção 5.3.2 da Parte 1 serve para os dados apresentados a seguir: dado o momento inicial da análise, as imagens apresentadas nesta seção servem como uma primeira visão da articulação das vogais reduzidas em contexto de sândi no PB. Uma hipótese mais embasada só será possível depois da análise completa e publicação dos dados articulatórios.

A Figura 51 a seguir mostra os *splines* (curvas) para /i/ na fala de taxa normal e rápida. Observada a escala de cores que representa a movimentação da língua, é possível perceber que o movimento da língua durante a produção do /i/ é na fala normal é como o esperado em um movimento mais estável: o dorso da língua se eleva em direção ao palato duro. Na produção do /i/, a língua movimenta-se ligeiramente para a frente em diferentes proporções, sobe e anterioriza-se, diminuindo o volume da cavidade anterior e aumentando o da posterior.

Nas figures abaixo para fala normal (à esquerda) e rápida (à direita), o movimento da língua é exatamente como o esperado, mas na fala rápida os alvos não são alcançados com a mesma magnitude e provavelmente sem a mesma rigidez (*stiffeness*).

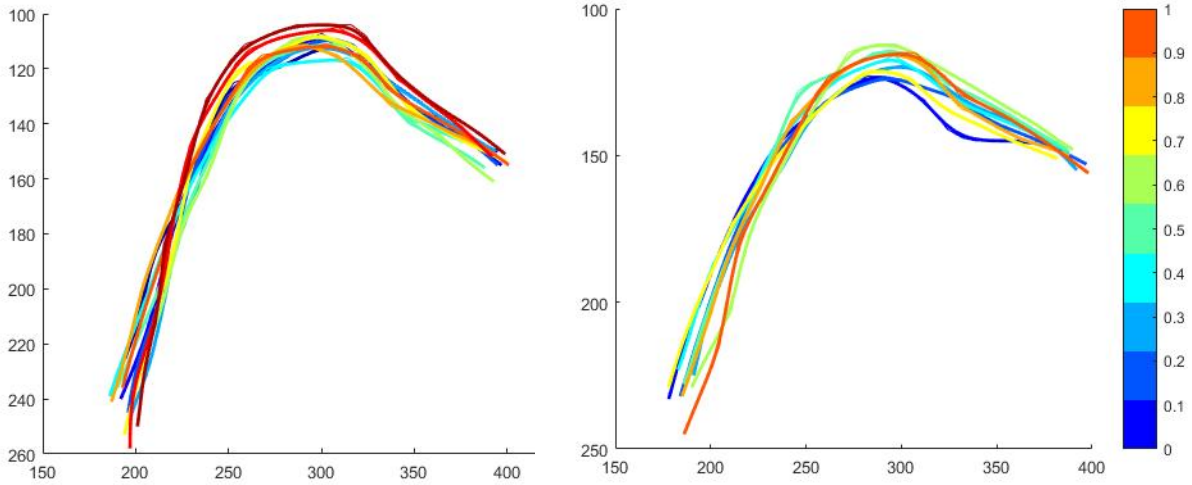


Figura 51 – Imagem sagital do quadro correspondente à /i/ na fala de velocidade normal (à esquerda) e fala de velocidade rápida (à direita)
Fonte: Meneses *et al.* (em preparação)

Diferentemente das imagens da Figura 51, é possível perceber, na Figura 52 à esquerda, a língua mais recuada e ligeiramente mais baixa durante a produção do /a/[3] em relação ao final do movimento, que está mais próxima do alvo da vogal /i/. Sobretudo, nessa imagem, fica clara o movimento esperado para a sequência /a+i/.

[3] No [a], a raiz da língua recua e o ponto de constrição ocorre exatamente na zona da faringe.

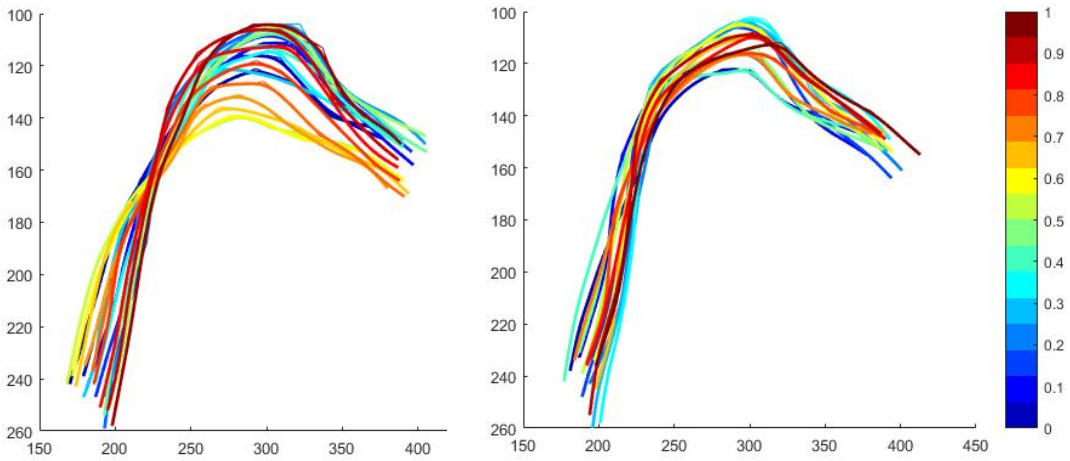


Figura 52 – Imagem sagital do quadro correspondente à sequência /a+i/ na fala de velocidade normal (à esquerda) e fala de velocidade rápida (à direita)
Fonte: Meneses *et al.* (em preparação)

No que concerne à fala rápida (figura à direita), se comparada com o /i/ rápido da Figura 51, é possível perceber mais movimento, mas o *release* do movimento é muito mais próximo do alvo esperado para /i/. O trabalho em preparação pretende analisar a posição e o deslocamento da língua. Somente a análise estatística das trajetórias pode esclarecer a questão, bem como atestar (ou não) as hipóteses levantadas aqui.

9.3 Conclusão

Em suma, os nossos resultados mostram que a chamada elisão vocálica em juntura de palavras mostra-se, assim como a apócope de vogais finais, um fenômeno gradiente que envolve dois aspectos principais: a redução do /a/ em fronteira prosódia átona e o aumento da velocidade de fala. Na medida em que a velocidade aumenta, mais difícil é para o falante coordenar os gestos de /a/ e /i/ com os alvos planejados. O resultado, então, é uma gama de possibilidades de alvo que vai do /a+i/ e pode, em condições mais extremas, resultar na elisão do /a/. Essa variabilidade e o gradiente mencionado parecem resultar de diferenças na coordenação gestual e diferentes graus de sobreposição entre gestos das vogais em questão.

Considerações Finais

A produção, percepção e disseminação dos processos fônicos de redução são tão complexas que mal conhecemos seus mecanismos gerais. Conforme discutido ao longo desta tese, os modelos abstratos podem explicar descobertas em torno da redução atribuindo à pronúncia reduzida um estatuto de variante regida por regras ou restrições. Por outro lado, modelos fonológicos mais comprometidos com as bases fonéticas precisam de pressupostos adicionais sobre a física da fala para explicar a redução.

Dois tipos diferentes de redução foram alvos de estudo no presente trabalho, a saber: o desvozeamento das vogais pós-tônicas altas e a elisão da vogal baixa em sândi. No primeiro caso, a visão inicial de que as vogais finais sofrem apócope em posições fracas no PB é desconstruída em favor de uma leitura mais reveladora: a vogal não é apagada, mas desvozeada. Dados acústicos, aerodinâmicos e de percepção sustentam essa ideia, a saber: a centralização, o deslocamento do centroide, o alongamento aparente do ruído, a semelhança entre o fluxo de ar de vozeadas e desvozeadas e a identificação da vogal pelo ruído alongado.

No segundo caso de nosso estudo, a interpretação da elisão da vogal /a/ em sândi dá lugar a uma visão mais gradiente da produção: na maioria dos casos, a vogal supostamente elidida deixa um rastro claro no sinal acústico; em casos extremos, restam poucas ou nenhuma pista da vogal reduzida no sinal acústico. Dados acústicos e aerodinâmicos atestam nossa ideia, a saber, maior duração de /a+i/ em relação a /i/ na fala rápida e normal, F1 e F2 significativos na fala normal e somente F2 significativo na fala rápida.

No desvozeamento, a vogal final que parece apagada está apenas radicalmente reduzida e quase inteiramente sobreposta à consoante precedente. Essa sobreposição de gestos supraglóticos inibe o vozeamento, pois os gestos glóticos correspondentes, sendo antagônicos, não podem se sobrepor. Nesse caso específico, a aerodinâmica obrigatória da consoante faz prevalecer a abdução das pregas vocais.

Em uma sílaba padrão, observa-se sempre o início síncrono de C e V seguido de uma porção assíncrona, resultante da maior duração da vogal, a qual costuma incluir um estado estacionário. No desvozeamento, por outro lado, o gesto vocálico “encolhe” no tempo e no espaço e a porção estacionária é truncada. O resultado é o *blending* dos gestos C e V. Se

uma porção da vogal "resiste" à sobreposição, torna-se curta e fraca demais para sustentar a produção da voz.

No caso da elisão, não estamos mais falando da coordenação CV, mas da coordenação mais complexa de um encontro de CV#V. Na fala mais lenta, é possível depreender o encontro das vogais, e na fala rápida, essa distinção pode se perder. Na maioria dos casos, esse encontro permanece no sinal acústico, uma vez que o gesto do /a/ pode diminuir e ocultar-se sob o gesto do /i/. Aqui, é preciso pensar em uma dinâmica mais ampla do movimento dos articuladores, dada a diferença entre movimentos da mandíbula e do corpo da língua na produção de vogais altas e baixas.

A altura da mandíbula para consoantes está nivelada com a de vogais altas, e dada a lentidão do movimento mandibular e a sua importância na articulação do /a/, é provável que haja uma perda de magnitude gestual na produção de uma sequência /Ca#V/ na fala mais rápida. A trajetória e os alvos da língua também são modificados pela velocidade da transição entre /a/ e /i/. É essa coordenação que explica por que o /a/ pode parecer apagado e por que as vogais altas e médias se sustentam no sândi vocálico.

À luz das informações acima, é possível afirmar que ambos os processos parecem ser resultado de questões específicas de *timing* e *target* articulatório, sendo que, no desvozeamento, há uma alteração radical da aerodinâmica supraglótica e glótica.

No desvozeamento, são a velocidade de ativação (*timing)* e a sobreposição dos gestos supraglóticos das vogais átonas (*target* articulatório) que inibem o vozeamento. A força do acoplamento da sílaba CV favorece o processo porque a sobreposição extrema dos gestos de C e V faz cair a pressão supraglótica, impedindo o vozeamento de recomeçar (cf. Ohala, 1983). Em casos como esse, C e V estão em fase e o acoplamento não só é mais natural, mas também é mais estável; por isso a queda da vogal é algo mais improvável. Em contrapartida, sequências que não estão em fase, como o encontro V#V heterossilábico do sândi, estão mais sujeitas ao refaseamento, por isso a vogal baixa pode ser significativamente ocultada no processo.[4] Nesse contexto, é a velocidade de transição entre /a#i/ (*timing*) que impede a movimentação regular da mandíbula, reduzindo a magnitude ou inibindo a produção do /a/ (*target* articulatório).

[4] Do ponto de vista do ouvinte, a coordenação antifase é até preferível, porque minimiza a sobreposição temporal entre articulações. Os gestos da vogal atingem os seus alvos mais lentamente e envolvem um menor grau de constrição. No entanto, no caso da elisão em sândi, no encontro de /CV_1#V_2/, o gesto de V_1 é iniciado simultaneamente com o gesto consonantal e truncado pelo gesto de V_2.

Outro ponto a ser considerado é a posição dos segmentos envolvidos na escala de sonoridade. No desvozeamento, as vogais altas são extremamente reduzidas porque o gesto da consoante prevalece sobre o da vogal. Na elisão em sândi, a vogal alta /i/ prevalece sobre a vogal baixa, tanto em duração quanto em magnitude. Em ambos os casos, é sempre o segmento de menor sonoridade que determina o grau de constrição do resultado do *blending*. Dessa forma, o enfraquecimento atinge o segmento de maior sonoridade em ambos os casos (CV e V#V). Esses fatos parecem indicar que a redução da vogal por sobreposição depende da sonoridade relativa dos segmentos envolvidos.

Os fenômenos descritos acima – redução, sobreposição gradiente e *blending* de gestos articulatórios – dão lugar a articulações "obscuras" que dificultam, por exemplo, a percepção de uma diferença de ponto de articulação entre consoantes adjacentes (Browman & Goldstein, 1992), ou de um /i/ em *blending* com uma fricativa, ou até mesmo de um encontro vocálico.[5]

Muitos dos efeitos linguísticos da coordenação dos gestos dos falantes são susceptíveis e motivados pela necessidade de se garantir e concluir com sucesso o ato da comunicação. Isso quer dizer que a variabilidade temporal mostrada na coordenação dos movimentos articulatórios parece ser sensível à necessidade de se recuperar as unidades linguísticas a partir do sinal acústico (cf. Chitoran, Goldstein & Byrd, 2002; Son, Kochetov & Pouplier, 2007; Tilsen, 2013).

A questão da recuperação dos gestos sobrepostos e redução foi abordada em Chen (2003), que utilizou logatomas com /db/ e /bd/. Os gestos variaram não só em termos de sobreposição, mas também em termos de redução do gesto da consoante. Um modelo de recuperação perceptual foi utilizado para detectar o grau de sobreposição e redução dos estímulos do sinal acústico. A autora concluiu que um gesto sobreposto produzido por um sujeito poderia ser recuperado por um ouvinte como um segmento gradientemente reduzido e imitado de forma incorreta. Tais "erros" poderiam ser fonologizados e, portanto, resultar em um processo de assimilação, seja gradiente ou categórico (cf. Beddor *et al*., 1986; Ohala 1990).

O fato de haver, respectivamente, 25% e 40% de erros no teste de identificação para as sílabas desvozeadas alongadas e de duração média (cf. subseção 4.2.5) e de haver 40%

[5] A questão da percepção de processos de redução tem sido um tema de crescente interesse e controvérsia, especialmente no contexto dos fenômenos de fala corrente (cf. Browman & Goldstein, 1990, 1992; Ohala, 1990; Byrd, 1996; Steriade, 2001; Chen, 2003).

e 50% de erros no teste de discriminação para fala normal e rápida no sândi (cf. seção 9.1.2), mostra que, em ambos os casos estudados, a recuperabilidade das sequências já não é completa. Nossos achados também vão ao encontro das ideias de Beddor *et al*. (1986) e Ohala (1990): a sobreposição de CV no desvozeamento e de V#V no sândi geram "erros" de percepção que podem, eventualmente, desencadear uma mudança sonora.

Referências

ABAURRE, M.; GALVES, C.; SCARPA, E. (1999) A interface fonologia-sintaxe. Evidências do português brasileiro para uma hipótese top-down na aquisição da linguagem. In: SCARPA, E. (Org). *Estudos de prosódia*. Campinas: Editora da UNICAMP, p. 285-323.

ALBANO, E. (1999) O português brasileiro e as controvérsias da fonética atual: pelo aperfeiçoamento da Fonologia Articulatória. *D.E.L.T.A.* [online] 15: 23-50.

_____. (2001) *O gesto e suas bordas*. Campinas: Mercado de Letras/São Paulo: FAPESP.

ALBANO, E. *et al*. (1998) A interface fonética-fonologia e a interação prosódia-segmentos. Relatório de grupo de trabalho. *Estudos Linguísticos XXVII*: Anais do XLV Seminário do Grupo de Estudos Linguísticos do Estado de São Paulo. São José do Rio Preto: UNESP-IBILCE, p. 135-143.

ALBANO, E.; MENESES, F. (2015) Novas luzes sobre a Dinâmica Sincrônica e Diacrônica do Desvozeamento Vocálico (apresentação oral). *Cadernos de Resumos do X Congresso Internacional da Abralin*. Belém, p. 59-60.

ANDREEVA, B.; KOREMAN, J. (2008) The status of vowel devoicing in Bulgarian: phonetic or phonological? In: ZYBATOW, T. *et al*. (Eds.). *Formal description of Slavic languages*: The Fifth Conference (FDSL5), Leipzig 2003. Frankfurt am Main: Peter Lang, p. 81-91.

AQUINO, P. (1997) *O papel das vogais reduzidas pós-tônicas na construção de um sistema de síntese concatenativa para o português do Brasil*. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem/Universidade Estadual de Campinas, Campinas.

BAKEN, R.; ORLIKOFF, R. (2000) *Clinical measurement of speech and voice*. 2.ed. San Diego: Singular.

BATTISTI, E.; HERMANS, B. (2007) Fixed and variable properties of Brazilian Portuguese palatalization. In: *Phonetics and Phonology in Iberia 2007 (PaPI'07)*. Braga: Ed.UMinho, p. 21-22.

BECKMAN, M. *et al*. (1992) The interaction of coarticulation and prosody in sound change. *Language and Speech* 35(1-2): 45-58.

van BERGEM, D. (1995) *Acoustic and lexical vowel reduction*. Ph.D. Dissertation, Universiteit van Amsterdam, Amsterdam.

BEDDOR, Patrice; KRAKOW, Rena; GOLDSTEIN, Louis. (1986) Perceptual constraints and phonological change. Phonology Yearbook 3: p. 197–217.

BERGMAN, R. (1971) Vowel sandi and word division in Igede. *Journal of West African Languages* 8(1): 13-25.

BERTI, L. (2006) *Aquisição incompleta do contraste entre /s/ e /ʃ/ em crianças falantes do português brasileiro.* Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem/Universidade Estadual de Campinas, Campinas.

BISOL, L. (1992) Sândi vocálico externo: degeminação e elisão. *Cadernos de Estudos Linguísticos* 23: 83-101.

_____. (1996) Sândi externo: o processo e a variação. In: ILARI, R. (Ed.). *Gramática do Português Falado IV*. Campinas: Editora da Unicamp.

_____. (2003) Sandhi in Brazilian Portuguese. *Probus* 15(2): 177-200.

_____. (2004) Mattoso Câmara Jr. e a Palavra Prosódica. *D.E.L.T.A.* 20(Número Especial): 59-70.

BROWMAN, C.; GOLDSTEIN, L. (1987) Tiers in articulatory phonology, with some implications for casual speech: Haskins Laboratories Status Report on Speech Research. In: KINGSTON, J.; BECKMAN, M. (Eds.). *Papers in Laboratory Phonology I*: between the grammar and the physics of speech. Cambridge: Cambridge University Press, p. 1-30.

_____. (1989) Articulatory gestures as phonological units. *Phonology* 6(2): 201-251.

_____. (1990) Gestural specification using dynamically-defined articulatory structures. Journal of Phonetics, 18, p. 299-320.

_____. (1992) Articulatory phonology: an overview. *Phonetica* 49(3-4): 155-180.

BYRD, D. (1996) Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24(2): 209-244.

BYRD, D.; SALTZMAN, E. (1998) Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics* 26(2): 173-199.

CÂMARA JR, J. M. (1953) *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Organização Simões.

_____. (1969) *Problemas de linguística descritiva*. Petrópolis: Vozes.

_____. (1970) *Estrutura da língua portuguesa*. 15.ed. Petrópolis: Vozes.

CAMPBELL, A. (1959) *Old English grammar*. Oxford: Oxford University Press.

CHEN, L (2003) The origins in overlap of place assimilation. In WCCFL 22. Proceedings of the XXIIth West Coast Conference on Formal Linguistics, Gina Garding and Mimu Tsujimura (eds.), 137–150. Somerville, MA: Cascadilla Press.

CHITORAN, I.; GOLDSTEIN, L.; BYRD, D. (2002) Gestural overlap and recoverability: articulatory evidence from Georgian. In: GUSSENHOVEN, C.; WARNER, N. (Eds.). *Papers in Laboratory Phonology VII*. Berlin: Mouton de Gruyter, p. 419-448.

CHITORAN, I; ISKAROUS, K. (2008) Acoustic evidence for high vowel devoicing in Lezgi. *Proceedings of ISSP 8*. Strasbourg, p. 93-96.

CHITORAN, I.; MARSICO, E. (2010) *Vowel devoicing*: an updated phonetic typology. Paper presented at BLS 36.

COLLISCHONN, G. (2011) A interação entre acento e processos de (re)estruturação silábica: um desafio para a Teoria da Otimidade. *Revista Linguística/Revista do Programa de PósGraduação em Linguística da Universidade Federal do Rio de Janeiro* 7(1): 88-98.

_____. (2012) Sândi vocálico no português brasileiro: como o acento determina sua realização. *Letras&Letras* 28(1): 13-27.

CUNHA, C. (2015) Portuguese lexical clusters and CVC sequences in speech perception and production. *Phonetica* 72(2-3): 138-161.

DAVIDSON, L. (2006) Schwa elision in fast speech: segmental deletion or gestural overlap? *Phonetica* 63(2-3): 79-112.

DELL, G.; JULIANO, C.; GOVINDJEE, A. (1993) Structure and content in language production: a theory of frame constraints in phonological speech errors. *Cognitive Science* 17(2): 149-195.

DEMOLIN, D.; STORTO, L. (2015) Temporal coordination of glottalic gestures in Karitiana. In: AVELINO, H.; COLER, M.; WETZELS, L. (Orgs.). *The phonetics and phonology of laryngeal features in native American languages*. Leiden: Brill, p. 308-321.

DUBIELA, M. (2013) *A vogal /e/ átona final em falantes curitibanos*. Monografia (Graduação em Letras-Inglês) – Universidade Federal do Paraná, Curitiba.

ELLIS, L.; HARDCASTLE, W. (2002) Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics* 30(3): 373-396.

ERNESTUS, M. (2000) *Voice assimilation and segment reduction in Casual Dutch*: a corpus-based study of the phonology-phonetics interface. Utrecht: LOT.

FAGYAL, Z.; MOISSET, C. (1999) Sound change and articulatory release: where and why are high vowels devoiced in Parisian French? *Proceedings of the XIVth International Congress of Phonetic Sciences*. Berkeley, University of California, p. 309-312.

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

FENG, L. (2007) The articulatory and acoustic study of fricative vowels in Suzhou Chinese. *16th ICPhS XVI*. Saarbrücken.

FERNANDES, A. (2007) *Apagamento de vogais átonas em trissílabos proparoxítonos*: um contributo para a compreensão da supressão vocálica em português europeu. Dissertação (Mestrado em Linguística) – Universidade do Porto, Porto.

FERREIRA, A. P. (2007) *Pet or petty? Diferenças entre palavras CVC e CVCV do inglês por aprendizes brasileiros: uma análise acústica*. Dissertação (Mestrado em Letras) – Universidade Federal do Paraná, Curitiba.

GHIO, A.; TESTON, B. (2004) Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proceedings of International*

Conference on Voice Physiology and Biomechanics. Marseille, Université de la Méditerranée, p. 55-58.

GOLDSTEIN, L.; CHITORAN, I.; SELKIRK, E. (2007) Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. *Proceedings of the XVI International Congress of Phonetic Sciences*. Saarbrücken, p. 241-244.

GOLDSTEIN, L. *et al*. (2009) Coupled oscillator planning model of speech timing and syllable structure. In: FANT, G.; FUJISAKI, H.; SHEN, J. (Eds.). *Frontiers in phonetics and speech science*. Beijing: The Commercial Press, p. 239-250.

GORDON, Matthew. (1998) The phonetics and phonology of non-modal vowels: a cross linguistic perspective. *Berkeley Linguistics Society* 24: 93-105.

HAN, M. (1962) Unvoicing of vowels in Japanese. *Nihon Onsei Gakkai* 10: 81-100.

HARRINGTON, J. (2012) The relationship between synchronic variation and diachronic change. In: COHN, A.; FOUGERON, C.; HUFFMAN, M. (Eds.). *Handbook of laboratory phonology*. Oxford: Oxford University Press, p. 321-332.

HARRINGTON, J.; FLETCHER, J.; ROBERTS, C. (1995) Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics* 23(3): 305-322.

HAYES, B. (1989) Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2): 253-306.

HOGG, R. (1992) Phonology and morphology. In: _____. (Ed.) *The Cambridge history of the English language* – vol. 1. Cambridge: Cambridge University Press, p. 67-123.

HUTCHINSON, S. (1974) *Spanish vowel sandhi*. Paper presented in the Parasession on Natural Phonology of the Regional fleeting of the Chicago Linguistic Society.

ISKAROUS, K.; MCDONOUGH, J.; WHALEN, D. (2012) A gestural account of the velar fricative in Navajo. *Laboratory Phonology* 3(1): 195-210.

JANNEDY, S. (1995) Gestural phasing as an explanation for vowel devoicing in Turkish. *Ohio State University Work Papers in Linguistics* 45: 56-84.

JOHNSON, K. (1997) Speech perception without speaker normalization: an exemplar model. In: JOHNSON, K.; MULLENNIX, J. (Eds.). *Talker variability in speech processing*. San Diego: Academic Press.

_____. (2004) Massive reduction in conversational American English. In: YONEYAMA, K., MAEKAWA, K. (Eds.) *Spontaneous speech*: data and analysis. (Proceedings of the 1st Session of the 10th International Symposium. The National International Institute for Japanese Language, Tokyo, p. 29-54.

JOMBART, T. (2013) A tutorial for Discriminant Analysis of Principal Components (DAPC) using *adegenet* 1.4-0. Disponível em: < http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.407.8769&rep=rep1&type=pdf>. Acesso em: 20 jul. 2016.

JONGENBURGER, W.; van HEUVEN, V. (1993) Sandhi processes in natural and synthetic speech. In: van HEUVEN, V.; POLS, L. (Eds.). *Analysis and synthesis of speech, strategic research towards high-quality text-to-speech generation*. Berlin: Mouton de Gruyter, p. 261-276.

KAGER, R. (1997) Rhytmic vowel deletion in Optimality Theory. In: ROCA, I. (Ed.). *Derivations and constraints in phonology*. Oxford: Oxford University Press, p. 463-499.

KAISSE, E. (2006) Sandhi. In: ASHER, R.; SIMPSON, J. (Eds.). The encyclopedia of language and linguistics. Oxford: Pergamon, p. 740-741.

KENT, R.; KIM, Y. (2003) Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics* 17(6): 427-445.

KOHLER, K. (2000) Investigating unscripted speech: implications for phonetics and phonology. *Phonetica* 57(2-4): 85-94

KOMATSU, M.; SANTOS, R. (2007) A variação na aquisição de regras de sândi externo em Português Brasileiro. *D.E.L.T.A.* 23(2): 223-244.

KONDO, M. (1997) Syllable constraints on Japanese vowel duration and intensity. In: Groupe Acoustique, Apprentissage et Interprétation (Eds.). *La voyelle dans tous ses états*. Nantes, Université de Nantes, p. 32-37.

_____. (2005) Syllable structure and its acoustic effects on vowels in devoicing environments. In: van de WEIJER, J.; NANJO, K.; NISHIHARA, T. (Eds.). *Voicing in Japanese*. Berlin: Mouton de Gruyter, p. 229-245.

KOOPMANS-VAN BEINUM, F. (1985) *Vowel contrast reduction, an acoustic and perceptual study of Dutch vowels in various speech conditions*. Ph.D. Dissertation, Universiteit van Amsterdam, Amsterdam.

LABOV, W. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

LADD, D.; SCOBBIE, J. (2003) External sandhi as gestural overlap? Counter-evidence from Sardinian. In: LOCAL, J.; OGDEN, R.; TEMPLE, R. (Eds.). *Papers in Laboratory Phonology VI*: phonetic interpretation. Cambridge: Cambridge University Press, p. 164-182.

LEMLE, M. (1966) *Phonemic system of the Portuguese of Rio de Janeiro*. Master Thesis, University of Pennsylvania, Philadelphia.

LEWIS, A. (2001) *Weakening of intervocalic /p, t, k/ in two Spanish dialects*: toward the quantification of lenition processes. Ph.D. Dissertation, University of Illinois, Urbana-Champaign.

LINDBLOM, B. (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35(11): 1773-1781.

_____. (1990) Explaining phonetic variation: a sketch of the H and H theory. In: HARDCASTLE, W.; MARCHAL, A. (Eds.). *Speech production and speech modeling*. Dordrecht: Kluwer, p. 403-439.

MENESES, F. (2012) *As vogais desvozeadas no português brasileiro*: investigação acústico-articulatória. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem/Universidade Estadual de Campinas, Campinas.

MENESES, F.; ALBANO, E. (2013) From devoicing to apocope: an acoustic study of poststressed high vowel lenition in Brazilian Portuguese. *Phonetics and Phonology in Iberia*, Universidade de Lisboa.

_____. (2015) From reduction to apocope: final poststressed vowel devoicing in Brazilian Portuguese. *Phonetica* 72(2-3): 121-137.

MENESES, F. *et al*. (2016) Gesture and velocity in voiced and devoiced vowels: EMA evidence from Brazilian Portuguese. *5th EICEFALA 9*.

MENESES, F.; POZZANI, D. (Em preparação) *Efeitos das características de africadas e fricativas no desvozeamento das vogais do PB*.

MITTERER, H.; ERNESTUS, M. (2006) Listeners recover /t/s that speakers reduce: evidence from /t/-lenition in Dutch. *Journal of Phonetics* 34(1): 73-103.

MOON, S.; LINDBLOM, B. (1994) Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96(1): 40-55.

NAM, H.; GOLDSTEIN, L.; SALTZMAN, E. (2009) Self-organization of syllable structure: a coupled oscillator model. In: PELLEGRINO, F.; MARSICO, E.; CHITORAN, I. (Eds.). *Approaches to phonological complexity*. Berlin/New York: Mouton de Gruyter, p. 299-328.

NAPOLEÃO, R. (2012) *A redução de vogais altas pretônicas no português de Belo Horizonte*: uma abordagem baseada na gradiência. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte.

NOLAN, F.; HOLST, T.; KÜHNERT, B. (1996) Modelling [s] to [ʃ] accommodation in English. *Journal of Phonetics* 24(1): 113-137.

OHALA, J. J. (1981) The listener as a source of sound change. In: MASEK, C.; HENDRICK, R.; MILLER, M. (Eds.). *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Linguistics Society, p. 178-203.

_____. (1983) The origin of sound patterns in vocal tract constraints. In: MACNEILAGE, P. (Ed.) *The Production of Speech* (pp. 189-216). New York: Springer Verlag.

_____ (1990) There is no interface between phonetics and phonology. A personal view. J. Phonetics, 18, p. 153-171.

_____. (1993a) The phonetics of sound change. In: JONES, C. (Ed.). *Historical linguistics*: problems and perspectives. London: Longman, p. 237-278.

_____. (1993b) Sound change as nature's speech perception experiment. *Speech communication* 13(1-2): 155-161.

_____. (2005) The marriage of phonetics and phonology. *Acoustical Science & Technology* 26(5): 418-422.

PAGEL, D. (1993) Contribuição para o estudo das vogais orais inacentuadas no português brasileiro. *Cadernos de Estudos Linguísticos* 25: 85-99.

PERKELL, J. (1969) Physiology of speech production: results and implications of a quantitative cineradiographic study. Cambridge: The MIT Press.

POZZANI, D. (2011) *Gradientes alofônicos de oclusivas alveolares do português brasileiro em uma situação de contato dialetal*. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem/Universidade Estadual de Campinas, Campinas.

RAMUS, F.; DUPOUX, E; MEHLER, J. (2003) The psychological reality of rhythm classes: perceptual studies. *15th ICPhS*. Barcelona, p. 337-342.

RECASENS, D. (2014) Gradient phonetic implementation of regressive voicing assimilation in Catalan heterosyllabic two- and three-consonant clusters. *Phonetica* 71(2): 128-156.

ROLO, M. (2010) *Apócope das vogais átonas finais [i] e [u] em duas localidades do Centro Sul Baiano*: Beco e Seabra. Dissertação (Mestrado em Letras e Linguística) – Universidade Federal da Bahia, Salvador.

ROLO, M.; MOTA, J. (2012) Um estudo sociolinguístico sobre o apagamento de vogais finais em uma localidade rural da Bahia. *SIGNUM: Estudos Linguísticos* 15(1): 311-334.

ROMERO, J. (1996) Articulatory blending of lingual gestures. *Journal of Phonetics* 24(1): 99-111.

RUSSEL, K. (2008) Sandhi in Plains Cree. *Journal of Phonetics* 36(3): 450-464.

SALTZMAN, E; MUNHALL, K. (1989) A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1(4): 333-382.

SALTZMAN, E. *et al*. (2008) A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proceedings of the 4th International Conference on Speech Prosody*. Campinas, p. 175-184.

SCOBBIE, J.; POUPLIER, M.; WRENCH, A. (2007) Conditioning factors in external sandhi: an EPG study of English /l/ vocalization. *Proceedings of the XVIth ICPhS*. Saarbrucken, p. 441-444.

SHOSTED, R. K. (2006) The Aeroacoustics of Nasalized Fricatives. PHD dissertation, University Of California: Berkeley.

_____. (2008) The Acoustics of Voiceless Nasal Vowels. Laboratory Phonology 11. Wellington, New Zealand, p. 121-122.

SILVA, A. (2002) *As fronteiras entre fonética e fonologia e a alofonia dos róticos iniciais em PB*: dados de dois informantes do sul do país. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem/Universidade Estadual de Campinas, Campinas.

SMITH, C. (2003) Vowel devoicing in contemporary French. *Journal of French Language Studies* 13(2): 177-194.

SON, M.; KOCHETOV, A.; POUPLIER, M. (2007) The role of gestural overlap in perceptual place assimilation: evidence from Korean. In: COLE, J.; HUALDE, J. (Eds.). *Papers in Laboratory Phonology IX*. Berlin: Mouton de Gruyter, p. 507-534.

STERIADE, D. (2001) Directional asymmetries in place assimilation: A perceptual account. In The Role of Speech Perception in Phonology, Elizabeth Hume and Keith Johnson (eds.), 219–250. San Diego: Academic Press

TAYLOR, M. (1994) *The interaction of vowel deletion and syllable structure constraints*. Ph.D. Dissertation, Simon Fraser University, Burnaby.

TENANI, L. (2002) Sândi vocálico e estrutura prosódica. *Estudos Linguísticos* 31: 1-4.

_____. (2003) Domínios prosódicos e processos de reestruturação silábica. *Estudos Linguísticos* 32(edição eletrônica).

TILSEN, S. (2013) A dynamical model of hierarchical selection and coordination in speech planning. *PloS One* 8(4).

TORREIRA, F.; ERNESTUS, M. (2010) Phrase-medial vowel devoicing in spontaneous French. *INTERSPEECH 2010*. Chiba, p. 26-30.

TSUCHIDA, A. (1994) Fricative-vowel coarticulation in Japanese devoiced syllables: acoustic and perceptual evidence. *Working Papers of the Cornell Phonetics Laboratory* 9: 183-222.

_____. (1997) *Phonetics and phonology of Japanese vowel devoicing*. Ph.D. Dissertation, Cornell University, Ithaca.

VAISSIÈRE, J. (2009) Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal* 2: 22-33.

VIARO, M. (2005) Algumas considerações acerca do português falado quatrocentista e quinhentista. *Papia* 15: 80-101.

VIEGAS, M.; OLIVEIRA, A. (2008) Apagamento da vogal átona final em Itaúna/MG e atuação lexical. *Revista da Abralin* 7: 119-138.

VIGÁRIO, M. (1998) Elisão da vogal não-recuada final e a palavra prosódica no Português Europeu. *Atas do XIII Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: Associação Portuguesa de Linguística.

WARNER, N. (2011) Reduction. In: van OOSTENDORP, M.; HUME, E.; RICE, K. (Eds.). *The Blackwell companion to phonology*. Malden; Oxford: Wiley-Blackwell, p. 223-240.

WHANG, J. (2014) *Effects of predictability on vowel reduction*. Pôster apresentado no Labphon 2014, Tokyo.

WHEELER, M. (2007) La síncope i l'apòcope en la història del català: una aproximació des de l'Optimitat. In: _____. (Ed). *Morfologia i fonologia catalana i romànica*: estudis diacrònics. Valência: Universitat de València.

ZSIGA, E. (1994) Gestural overlap in consonant sequences. *Journal of Phonetics* 22(1): 121-140.

_____. (1995) An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. In: CONNELL, B.; ARVANITI, B. (Orgs.). *Papers in Laboratory Phonology IV*: phonology and phonetic evidence. Cambridge: Cambridge University Press, p. 282-302.

_____. (2000) Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28(1): 69-102.

Anexos

Anexo 1: Resumo Labphon 2016

Gesture and velocity in Brazilian Portuguese devoiced vowels: a preliminary EMA study

Francisco Meneses, Denise Pozzani*, Nicole Wong+, Zainab Hermes+, Torrey Loucks+, Ryan Shosted+*
*University of Campinas, +University of Illinois at Urbana-Champaign
chico.meneses@gmail.com

Vowel reduction with devoicing in certain contexts occurs in synchronic variation in Brazilian Portuguese (BP). It is well known that devoicing not only eliminates f0 but also weakens the regular formant structure of vowels. Ohala (1983) stated that a smaller supralaryngeal constriction reduces the transglottal pressure differential, which yields devoicing. Among devoiced poststressed vowels in BP, reduction of duration is caused by extreme overlap between consonant and vowel gestures. This overlap implies that there is insufficient time to achieve the oral opening necessary for a significant drop in supraglottal pressure (Meneses & Albano, 2015). These conditions result in a narrow oral aperture during devoiced vowels. Thus, timing of the onset and nucleus gestures of the syllable containing the devoiced vowel seems to determine the amount of voicing. We argue that supralaryngeal articulatory differences in voiceless vowels arise from online adjustments of their motor program, with specific targets, suggesting that vowel devoicing in BP is not just epiphenomenally tied to specific phonological contexts.

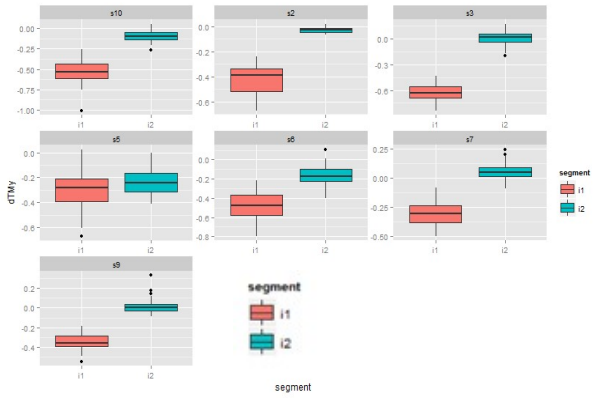
We present preliminary electromagnetic articulography (EMA) data to help clarify the articulation of devoiced final poststressed vowels in BP and to explain how their supralaryngeal characteristics differ with respect to those of their voiced counterparts. Although the data were gathered in the context of another experiment, this analysis can clear the ground for future research.

Lingual position data were collected using the NDI Wave system. Three sensors were glued at 1 cm intervals on the anterior portion of the tongue (tip: TT, mid: TM, and back: TB). Other sensors tracked jaw/chin (JC), upper and lower lip (UL, LL) movements and another one was used for head correction and rotation to a common occlusal plane. Seven female speakers of BP from São Paulo produced utterances in a carrier sentence including the utterance-initial verb *Digo* ('I say') and the utterance-final prepositional phrase *para ti* ('to you')—in this case /t/ is most commonly produced as an alveopalatal affricate.

Cross-linguistically, there is a preference for high vowel devoicing (Gordon, 1998), and /i/ devoices more frequently than /u/ (Rodgers, 1997). Considering this, we decided to compare the production of the two /i/s in the same carrier sentence. The first /i/ (*i1*) was expected to be voiced and the second /i/ (*i2*) was expected to devoice often due to their respective phonological contexts.

Data were manually segmented in Praat with reference to audio which was recorded simultaneously with the EMA data. When no voicing was present during the vowel, we segmented the entire fricative noise, expecting that articulatory traces of /i/ would still be evident. Values were collected within segment boundaries and positional data were z-normalized within speaker. The average position and velocity of the sensors was measured. For the inferior–superior dimension (y), jaw/chin (JC) position was subtracted. Vertical velocity of TM (the first derivative of the corrected TM signal; dTMy) for *i1* (in 'digo') was consistently lower than the velocity for *i2* (in 'para ti'), as in Figure 1.

Figure 1: Tongue mid sensor velocity for seven BP speaker (in mm/s).



UL, LL, TT, TM and TB values were submitted to a combined principal components analysis (PCA) and linear discriminant analysis (LDA: Jombart, 2013). Two PCs optimally accounted for the variance in the data. The first component (PC1) explained 40% of the variance and was strongly, negatively correlated with lip protrusion and tongue elevation. The second component (PC2) explained 23% of the variance, and was strongly correlated with tongue fronting. LDA based on the resulting PCs showed that both *i1* and *i2* were clearly differentiated along these two dimensions, as seen in Table 1.

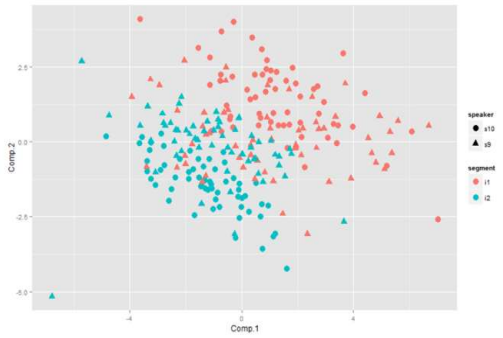
Table 1: Classification results of LDA, showing separation of *i1* and *i2* into two classes

LD-Class	i1	i2
1	***57***	17
2	14	***58***

Our materials do not permit us to observe the variation of vowel voicing for the same speaker in the same phonological environment. Since some speakers categorically voiced all vowels, we chose to observe *i2* as produced by one speaker who consistently devoiced it and by another who did not. We compared S9 and S10: S9 devoiced *i2* most of the time and S10 devoiced it much less frequently. For S9, the position of the two posterior sensors, corresponding to the tongue body, was higher than it was for the others (including S10). The absolute velocity of S9's sensors was also greater during the devoiced vowel.

With respect to aspects of supralaryngeal articulation other than tongue velocity, Figure 2 shows that *i2* was produced with more lip protrusion and tongue elevation than *i1*. In addition, *i2* manifested more anterior tongue position when it was devoiced.

Figure 2: Phrase-initial (*i1*) and phrase-final (*i2*) /i/ from S9 and S10; PC1 is on the x-axis and PC2 is on the y-axis. Blue markers represent the post-stressed vowels (*i2*). Note that blue triangles generally have a more anterior tongue position (greater PC2) than circles. Because S9 devoiced more consistently, we consider this difference to be associated with *i2* when it is devoiced.



We suggest that the devoiced vowel may involve a difference in gestural coordination as compared to a fully voiced vowel. The increased velocity and tongue body elevation observed are in agreement with the hypothesis and the acoustic/aerodynamic data analyzed by Meneses & Albano (2015), as mentioned above. Changes in lingual configuration are consistent with the hypothesis that voiceless /i/ in BP manifests its own online adjustments of the supralaryngeal motor program. To further investigate the tongue velocity and fronting issues, we are designing a new experiment, in which prosodic conditions and phonetic contexts are more controlled.

References

Gordon, M. (1998) The phonetics and phonology of non-modal vowels: A cross-linguistic perspective. *Proceedings of the Berkeley Linguistics Society,* 24, 93-105.
Meneses, F.O.; Albano, E.C. (2015) From reduction to apocope: Final post-stressed vowel devoicing in Brazilian Portuguese. *Phonetica,* 72, 121-137.
Jombart, T. (2013). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 1.4-0.
Ohala, J. J. (1983) The origin of sound patterns in vocal tract constraints. In: P. F. MacNeilage (ed.), *The production of speech*. New York: Springer-Verlag. 189-216.
Rodgers, J. E. J. (1997). Vowel devoicing/deletion in English and German. In A. P. Simpson, K. J. Kohler, and T. Rettstadt (Eds.), The *Kiel Corpus of Read/Spontaneous Speech — Acoustic data base, processing tools and analysis results*, AIPUK 32, 177–195.

Anexo 2: Lista completa de frases-veículo para Experimento de Produção II

1. Estão chamando isso de paridade agora
2. Estão chamando isso de para idade agora
3. Ele chegou a paternidade agora
4. Ele chegou a paterna idade agora
5. Estão chamando isso de maternidade agora
6. Estão chamando isso de materna idade agora
7. Eu não faria isso com o cara idoso jamais
8. Eu não faria isso com o caridoso jamais
9. A pia elite está incomodando
10. A pielite está incomodando
11. Acabaram encontrando a ata urdida na sala
12. Acabaram encontrando a aturdida na sala
13. Chegou-se a fraternidade agora
14. Chegou-se a fraterna idade agora
15. Estão chamando isso de raridade agora
16. Estão chamando isso de rara idade agora
17. Estão chamando isso de puridade agora
18. Estão chamando isso de pura idade agora
19. Estão chamando isso de sincera idade agora
20. Estão chamando isso de sinceridade agora

Anexo 3: Script de extração de formantes Kawahara (2010)

```
# This Praat script will get F1, F2, and F3 at the midpoints of all the intervals of the all files in the specified folder.

# Version: 3 Feb 2010

# Author: Shigeto Kawahara

# Input: TextGrid and wav in the same directly. They must have the same name.


form Get F1, F2, F3
	sentence Directory ./
	comment If you want to analyze all the files, leave this blank
	word Base_file_name
	comment The name of result file
	text textfile result.txt
endform


# Write-out the header


fileappend "'textfile$'" soundname'tab$'intervalname'tab$'F1'tab$'F2'tab$'F3'tab$'
fileappend "'textfile$'" 'newline$'


#Read all files in a folder
Create Strings as file list... wavlist 'directory$'/'base_file_name$'*.wav
Create Strings as file list... gridlist 'directory$'/'base_file_name$'*.TextGrid
n = Get number of strings


for i to n
clearinfo


#We first extract a formant tier
	select Strings wavlist
```

```
        filename$ = Get string... i
        Read from file... 'directory$'/'filename$'
        soundname$ = selected$ ("Sound")
        To Formant (burg)... 0 5 5500 0.025 50

# We now read grid files and extract all intervals in them
        select Strings gridlist
        gridname$ = Get string... i
        Read from file... 'directory$'/'gridname$'
        int=Get number of intervals... 1

# We then calculate F1, F2 and F3

for k from 1 to 'int'
        select TextGrid 'soundname$'
        label$ = Get label of interval... 1 'k'
        if label$ <> ""

                # calculates the mid point
                vowel_onset = Get starting point... 1 'k'
                vowel_offset = Get end point... 1 'k'
                midpoint = vowel_onset + ((vowel_offset - vowel_onset) / 2)

                # get the formant values at the midpoint
                select Formant 'soundname$'
                f_one = Get value at time... 1 'midpoint' Hertz Linear
                f_two = Get value at time... 2 'midpoint' Hertz Linear
                f_three = Get value at time... 3 'midpoint' Hertz Linear
                resultline$ = "'soundname$"tab$"label$"tab$"f_one"tab$"f_two"tab$"f_three"tab$"'
                fileappend "'textfile$'" 'resultline$'
        endif
```

```
endfor

fileappend "'textfile$'" 'newline$'



endfor

# clean up

select all
Remove
```

Anexo 4: script para experimento de percepção

"ooTextFile"

"ExperimentMFC 5"

stimuliAreSounds? <yes>

stimulusFileNameHead = "stimuli/"

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 0.5 seconds

stimulusMedialSilenceDuration = 0.5 seconds

numberOfDifferentStimuli = 40

"cara_idosoxcara_idoso.wav" ""

"cara_idosoxcara_idoso_fast.wav" ""

"cara_idosoxcaridoso.wav" ""

"cara_idosoxcaridoso_fast.wav" ""

"caridosoxcara_idoso.wav" ""

"caridosoxcara_idoso_fast.wav" ""

"caridosoxcaridoso.wav" ""

"caridosoxcaridoso_fast.wav" ""

"fraterna_idadexfraterna_idade.wav" ""

"fraterna_idadexfraterna_idade_fast.wav" ""

"fraterna_idadexfraternidade.wav" ""

"fraterna_idadexfraternidade_fast.wav" ""

"fraternidadexfraterna_idade.wav" ""

"fraternidadexfraterna_idade_fast.wav" ""

"fraternidadexfraternidade.wav" ""

"fraternidadexfraternidade_fast.wav" ""

"materna_idadexmaterna_idade.wav" ""

"materna_idadexmaterna_idade_fast.wav" ""

"materna_idadexmaternidade.wav" ""

"materna_idadexmaternidade_fast.wav" ""

"maternidadexmaterna_idade.wav" ""

"maternidadexmaterna_idade_fast.wav" ""

"maternidadexmaternidade.wav" ""

"maternidadexmaternidade_fast.wav" ""

"pura_idadexpura_idade.wav" ""

"pura_idadexpura_idade_fast.wav" ""

"pura_idadexpuridade.wav" ""

"pura_idadexpuridade_fast.wav" ""

"puridadexpura_idade.wav" ""

"puridadexpura_idade_fast.wav" ""

"puridadexpuridade.wav" ""

"puridadexpuridade_fast.wav" ""

"sincera_idadexsincera_idade.wav" ""

"sincera_idadexsincera_idade_fast.wav" ""

"sincera_idadexsinceridade.wav" ""

"sincera_idadexsinceridade_fast.wav" ""

"sinceridadexsincera_idade.wav" ""

"sinceridadexsincera_idade_fast.wav" ""

"sinceridadexsinceridade.wav" ""

"sinceridadexsinceridade_fast.wav" ""

numberOfReplicationsPerStimulus = 3

breakAfterEvery = 0

randomize = <PermuteBalancedNoDoublets>

startText = "Click to start."

runText = "Os sons são iguais ou diferentes?"

pauseText = ""

endText = "Acabou! Obrigado!"

maximumNumberOfReplays = 0

replayButton = 0 0 0 0 "" ""

okButton = 0 0 0 0 "" ""

oopsButton = 0 0 0 0 "" ""

responsesAreSounds? <no> "" "" "" "" 0 0

numberOfDifferentResponses = 2

0.3 0.4 0.5 0.6 "Igual: hit 1" 15 "1" "1"

0.5 0.6 0.5 0.6 "Diferente: hit 0" 15 "0" "2"

numberOfGoodnessCategories = 0