



Domínios de Lingu@gem

Fonética e Fonologia -
Homenagem a Gisela
Collischonn

Organização: José Magalhães (UFU)

2º Trimestre 2016
Volume 10, número 2

ISSN: 1980-5799

Expediente

Universidade Federal de Uberlândia

Reitor

Prof. Elmiro Santos Resende

Vice-Reitor

Prof. Eduardo Nunes Guimarães

Diretora da EDUFU

Profa. Belchiolina Beatriz Fonseca

Diretora do Instituto de Letras e Linguística

Profa. Maria Inês Vasconcelos Felice

EDUFU – Editora e Livraria da Universidade Federal de Uberlândia
Av. João Naves de Ávila, 2121 - Bloco 1S - Térreo - Campus Santa Mônica - CEP: 38.408-
144 - Uberlândia - MG
Telefax: (34) 3239-4293
Email : vendas@edufu.ufu.br | www.edufu.ufu.br

Editoração e Diagramação: Guilherme Fromm

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

Domínios de Linguagem, v. 10, n. 2, 2016, Uberlândia, Universidade Federal
de Uberlândia, Instituto de Letras e Linguística, 2007-

Trimestral.

Modo de acesso:

<http://www.seer.ufu.br/index.php/dominiosdelinguagem>.

Editoração: Guilherme Fromm.

Organização: José Magalhães.

ISSN: 1980-5799

1. Linguística - Periódicos. 2. Linguística aplicada - Periódicos.
I. Universidade Federal de Uberlândia. Instituto de Letras e Linguística.

CDU: 801(05)

Todos os artigos desta revista são de inteira responsabilidade de seus autores, não cabendo qualquer responsabilidade legal sobre seu conteúdo à Revista, ao Instituto de Letras e Linguística ou à Edufu.

Domínios de Linguagem

Diretor

Guilherme Fromm (UFU)

Conselho Editorial

Ariel Novodvorski (UFU)

Betina Ribeiro Rodrigues da Cunha (UFU)

Eliana Dias (UFU)

Fabio Izaltino Laura (UFU)

Cristiane Carvalho de Paula Brito (UFU)

Marleide Dias Esqueda (UFU)

Comissão Científica

Adriana Cristina Cristianini (UFU), Aldo Luiz Bizzocchi (FMU), Alice Cunha de Freitas (UFU), Ataliba T. de Castilho (USP/UNICAMP), Carla Nunes Vieira Tavares (UFU), Cecília Magalhães Mollica (UFRJ), Cintia Vianna (UFU), Cirineu Cecote Stein (UFPB), Claudia Maria Xatara (UNESP), Claudia Zavaglia (UNESP/SJ Rio Preto), Cláudio Márcio do Carmo (UFOP), Cleci Regina Bevilacqua (UFRGS), Clecio dos Santos Bunzen (UNIFESP), Cristiane Brito (UFU), Dánie Marcelo Jesus (UFMT), Deise Prina Dutra (UFMG), Dilys Karen Rees (UFG), Eduardo Batista da Silva (UEG), Elisa Battisti (UFRGS), Elisete Carvalho Mesquita (UFU), Ernesto Sérgio Bertoldo (UFU), Fabiana Vanessa Gonzalis (UFU), Fernanda Costa Ribas (UFU), Francine de Assis Silveira (UFU), Francis Henrik Aubert (USP), Gabriel Antunes Araujo (USP), Gabriel de Avila Othero (UFRGS), Giacomo Figueredo (UFOP), Hardarik Bluehdorn (Institut für Deutsche Sprache Mannheim – Alemanha), Heliana Mello (UFMG), Heloisa Mara Mendes (UFU), Igor Antônio Lourenço da Silva (UFU), Irenilde Pereira dos Santos (USP), Jacqueline de Fatima dos Santos Moraes (UERJ), Janice Helena Chaves Marinho (UFMG), João Bôscó Cabral dos Santos (UFU), Jose Luiz Fiorin (USP), José Ribamar Lopes Batista Júnior (CAF/UFPI), José Sueli de Magalhães (UFU), Karylleila Santos Andrade (UFT), Luiz Carlos Travaglia (UFU), Liliane Santos (Université Charles-de-Gaulle - Lille 3 - França), Manoel Mourivaldo Santiago-Almeida (USP), Marcelo Módolo (USP), Márcia Mendonça (UNICAMP), Maria Angélica Furtado da Cunha (UFRN), Maria Aparecida Resende Ottoni (UFU), Maria Cecília de Lima (UFU), Maria Célia Lima-Hernandes (USP), Maria de Fátima Fonseca Guilherme (UFU), Maria do Perpétuo Socorro Cardoso da Silva (UEPA), Maria Helena de Paula (UFG), Maria José Bocorny Finatto (UFRGS), Maria Luisa Ortiz Alvarez (UnB), Maria Luiza Braga (UFRJ), Maria Suzana Moreira do Carmo (UFU), Marlúcia Maria Alves (UFU), Maurício Viana Araújo (UFU), Michael J. Ferreira (Georgetown University – EUA), Montserrat Souto (Universidade Santiago de Compostela – Espanha), Nilza Barrozo Dias (UFF), Patricia de Jesus Carvalhinhos (USP), Paulo Osório (Universidade da Beira Interior – Portugal), Paulo Rogério Stella (UFAL), Pedro Malard Monteiro (UFU), Pedro Perini-Santos (PUC-Minas), Raquel Meister Ko. Freitag (UFS), Roberta Rego Rodrigues (CLC/UFPEL), Rolf Kemmler (Universidade de Trás-os-Montes e Alto Douro – Portugal), Sebastião Carlos Leite Gonçalves (UNESP/S.J. Rio Preto), Silvana Maria de Jesus, (UFU), Silvia Melo-Pfeifer (Universidade de Hamburgo – Alemanha), Simone Floripi (UFU), Simone Tiemi Hashiguti (UFU), Sinara de Oliveira Branco (UFCG), Stella Esther Ortweiler Tagnin (USP), Tommaso Raso (UFMG), Ubirajara Inácio Araújo (UFPR), Valeska Virgínia Soares Souza (IFTM), Vanessa Hagemeyer Burgo (UFMS), Vânia Cristina Casseb Galvão (UFG), Vera Lucia Menezes de Oliveira e Paiva (UFMG), Vitalina Maria Frosi (UCS), Waldenor Barros Moraes Filho (UFU), Zelina Márcia Pereira Beato (UESC).

Participaram dessa edição como pareceristas *ad hoc*

Antonio Carlos Silvano Pessotti (UNICAMP)

Carlos Alexandre Victorio Gonçalves (UFRJ)

Eleonora Cavalcante Albano (UNICAMP)

Ester Mirian Scarpa (UNICAMP)

Giovana Ferreira Gonçalves (UFPeI)

Juliene Lopes Ribeiro Pedrosa (UFPB)

Miguel Oliveira Jr. (UFAL)

Plínio Almeida Barbosa (UNICAMP)

Rogério V. Ferreira (UFMS/UNICAMP)

Ronaldo Mangeira Lima Jr. (UFC)

Rubens Marques de Lucena (UFPB)

Rui Rothe-Neves (UFMG)

Seung Hwa Lee (UFMG)

Susiele Machry da Silva (UTFPR)

Táise Simioni (UNIPAMPA)

Tatiana Keller (UFSM)

Waldemar Ferreira Netto (USP)

Domínios de Linguagem

Sumário

Expediente.....	435
Sumário	438
Apresentação	439
Fonética, Fonologia e o legado de Gisela Collischonn – José Magalhães (UFU)	439
Artigos	449
Sobre a preservação de expoentes morfológicos na fonologia variável do português brasileiro - Luiz Carlos Schwindt (UFRGS)	449
O mapeamento fonético-fonológico das vogais postônicas finais no português brasileiro - Carmen Lúcia Matzenauer (UCPEL).....	466
Uma abordagem fonológica para as postônicas médias não-finais - Arthur Pereira Santana (USP).....	494
A estrutura silábica em esperanto - Karina Gonçalves de Souza de Oliveira (USP).....	519
Variações de F0 e configurações de frase entoacional: análise de estruturas contrastivas - Geovana Soncin (UNESP/Rio Preto), Luciani Tenani (UNESP/Rio Preto).....	534
Análise perceptiva e acústica em fonética forense: uma pesquisa em disfarce de voz - Maria Lúcia de Castro Gomes (UTFPR), Denise de Oliveira Carneiro (IC-PR), Andrea Alves Guimarães Dresch (IC-PR)	559
Salvador, Vitória da Conquista e Teófilo Otoni: cidades e falares diferentes? Uma análise discriminante da F ₀ - Vera Pacheco (UESB), Marian Oliveira (UESB), Tássia da Silva Coelho	590
A aquisição do sistema vocálico do português por falantes nativos da variedade rio-platense de espanhol: uma discussão sobre a bidirecionalidade da transferência vocálica - Letícia Pereyron (UFRGS), Ubiratã Kickhöfel Alves (UFRGS).....	616
A Aquisição Fonológica Variável da Nasal Velar por Aprendizes de Inglês-L2: Análise pela Teoria da Otimidade Estocástica - Athany Gutierrez (UFRGS)	646
Nuclear stress placement by Brazilian users of English as an international language - Leonice Passarella dos Reis (UFSC/ EAMSC), Rosane Silveira (UFSC).....	673
Adaptações fonológicas na pronúncia de nomes comerciais com elementos do inglês no Brasil - Natália Cristine Prado (UNIR).....	703
Uma discussão acerca da aplicação do <i>Perceptual Assimilation Model-L2</i> à percepção fônica de língua estrangeira: questões de pesquisa e desafios teóricos - Reiner Vinicius Perozzo (UFRGS), Ubiratã Kickhöfel Alves (UFRGS)	733

Fonética, Fonologia e o legado de Gisela Collischonn Phonetics, Phonology and the legacy of Gisela Collischonn

José Magalhães*

Nascida em Lajeado-RS, em dois de abril de mil novecentos e sessenta e quatro, de família luterana, Gisela começou, ainda jovem, a cantar em coros. Sempre apaixonada pela música, foi nesse contexto que conheceu seu marido, Manuel, músico e regente de coro. Na tenra juventude, remava, o que lhe conferiu braços invejavelmente fortes. Com uma singular delicadeza de comportamento e caráter, ela adorava também a delicadeza dos passarinhos, os quais sempre contemplava, sabendo nomeá-los todos.

Na contramão dos discursos sórdidos e propagadores de ódio que o Brasil vem presenciando ultimamente, muitas vezes reverberados por políticos, religiosos e setores da mídia, Gisela abominava qualquer discurso que se pautasse nesses vieses; também desprezava atitudes machistas, mas não se autointitulava feminista. Era ciumenta e defensora radical dos amigos, sem ser possessiva. Embora fosse conhecida por ter um paladar muito amplo (diziam que saboreava até pedra), não comia, sob qualquer circunstância, nem peito de frango nem peixe anjo.

Gisela era um ser humano de superior generosidade. Não era, todavia, uma generosidade artificial ou que vislumbrasse retorno. Seu interesse era genuíno pelo ser humano em sua essência, por seu comportamento, pelo que as pessoas pensavam e sobre como viam o mundo. Isso fazia dela uma criatura naturalmente despida de preconceitos de qualquer tipo; conseqüentemente, rejeitava com rigor aqueles que detinham preconceitos de qualquer natureza.

Curiosa, compenetrada e pesquisadora incansável, Gisela concluiu, em 1987, seu curso de graduação em Letras/Licenciatura com habilitação em português e inglês, na Universidade

* Professor do Programa de Pós-Graduação em Estudos Linguísticos (Mestrado e Doutorado) e do Programa de Pós-Graduação Mestrado Profissional em Letras, do Instituto de Letras e Linguística da Universidade Federal de Uberlândia.

Federal do Rio Grande do Sul, para onde voltaria poucos anos depois, como professora efetiva. Isso após concluir o mestrado (1993) e o doutorado (1997), ambos na Pontifícia Universidade Católica do Rio Grande do Sul, sob orientação da Professora Leda Bisol. Em sua dissertação de mestrado, realizou o primeiro grande estudo (talvez o mais importante até hoje) sobre o acento secundário no português brasileiro com base no modelo de fonologia métrica de Halle e Vergnaud (1987). No doutorado, voltou-se para a descrição da estrutura da sílaba no português brasileiro, debruçando-se sobre os modelos não-lineares, especialmente, o proposto por Itô (1986, 1989), avaliando ser este mais econômico e mais restrito do que os modelos que operam com regras de silabação. Ainda em sua tese, Gisela fez um amplo estudo sobre a epêntese vocálica em português, como nos casos de [pnew] ‘pneu’; [pɛpsi] ‘Pepsi’ e [klubɪ] ‘clube’, respectivamente epêntese na sílaba inicial, na medial e na final. Adotando a proposta de Pigott (1995), concluiu que sílabas epentéticas não têm peso, o que as tornam invisíveis às regras de acento em português; outra conclusão importante acerca da epêntese é que este processo ocorre no léxico e, portanto, deve ser submetido às regras de natureza lexical. A importância de suas pesquisas de mestrado e doutorado ecoou por anos seguidos e, mesmo nos dias de hoje, qualquer investigação séria que se volte para estudos de acento e sílaba no português deve, necessariamente, retomar os trabalhos de Collischonn como suporte indispensável para novas análises. Por exemplo, Collischonn (1993) é referência fundamental para o recente estudo sobre o acento nas variedades brasileira e europeia do português, em Magalhães (2016).

Defendida a tese de doutorado em 1997, ela não se acomodou. Sua aptidão como investigadora contumaz rendeu-lhe inúmeros projetos de pesquisa. Coordenou e concluiu importantes projetos, tais como: i) Variação da epêntese no português do sul do Brasil, em que se propôs a contribuir para uma caracterização mais precisa da frequência e dos modos de ocorrência da epêntese vocálica no Português Brasileiro (variedades da região sul do Brasil abrangidas pelo Projeto VARSUL), o que permitiria uma avaliação da adequação descritiva das análises teóricas propostas para a fonologia do português brasileiro; ii) Realização variável da vogal em contextos SC iniciais, em que realizou um estudo quantitativo da epêntese vocálica como fenômeno variável no português falado na região sul do Brasil, considerando a realização variável da vogal em início de palavra, antecedendo sequências de /s/ + outra consoante, como em ‘spa’, ‘skol’, ‘estoque’; iii) Padrões de acento na poesia brasileira, em que investigou o ritmo de acento na poesia, partindo das categorias linguísticas propostas pelas teorias fonológicas do acento e dos constituintes prosódicos; iv) Regras fonológicas variáveis e

fonologia lexical do português, em que buscou reanalisar o comportamento de determinadas regras fonológicas variáveis do português brasileiro, almejando obter um enquadramento dessas regras dentro de uma perspectiva abrangente da organização do componente fonológico da língua; v) Processos vocálicos no PB: resolução de hiato na palavra e na frase, quando investigou os processos fonológicos que se aplicam a sequências de vogais na fronteira entre palavras e no interior destas, no português, focalizando a aplicação dos fenômenos em fronteira de palavras na língua falada e seus fatores condicionadores, a aplicação destes fenômenos e/ou seu bloqueio internamente às palavras e a compreensão dos constituintes prosódicos e/ou morfológicos que limitam ou determinam a aplicação dos fenômenos. Por último, estava à frente de dois projetos de pesquisa simultaneamente, a saber: juntamente com a Professora Sônia Frota, coordenava o projeto “Atlas Interactivo da Prosódia do Português”, e, em parceria com sua ex-orientanda de doutorado, Juliana Escalier Ludwig Gayer, coordenava o projeto “Fonologia do nível da frase: a proeminência acentual/tonal e processos de resolução de hiato”.

Em sua breve, porém profícua carreira, Gisela trabalhou com garra e dedicação, tendo publicado vinte e seis artigos em periódicos nacionais e internacionais e vinte e dois capítulos de livro; soma-se a esta vasta produção a edição/organização de muitos livros e revistas acadêmicas, a participação em dezenas de congressos, publicação de trabalhos em anais, organização de eventos importantes para o cenário linguístico nacional e internacional.

Gisela nos deixa ainda seu legado como formadora de novos pesquisadores em todos os níveis: orientou dezessete trabalhos de iniciação científica, dez trabalhos de conclusão de curso de graduação, quatro teses de doutorado e treze dissertações de mestrado. Infelizmente, quis o destino que fosse interrompida a orientação de mais um trabalho de iniciação científica, três teses de doutorado e uma dissertação de mestrado, quando, na manhã do dia 15 de junho de 2016, em Porto Alegre, deixou-nos e foi transferida para outro plano. Contudo, sua presença permanecerá entre aqueles que com ela conviveram, que simplesmente a conheceram ou que tiveram e terão contato com ela por meio de seu trabalho. As emocionantes palavras do Professor Luiz Carlos Schwindt, amigo-irmão, colega e parceiro inseparável de Gisela, traduzem de forma precisa o grande ser humano que não está mais fisicamente conosco: “Era generosa. Profundamente generosa. Sábia. Profundamente sábia. Sensível. Profundamente sensível. Gisela olhava igualmente para uma pessoa com muitos títulos e para um estudante de início de curso. Interessavam-lhe as ideias, e estabelecia relações entre as diferentes ideias com destreza invejável. No meio disso tudo, sentia. E expressava seu sentimento com a timidez

desajeitada dos artistas. Sim, a arte foi sempre parte essencial de sua vida e se casou lindamente com a fonologia que praticou”.

Somos cientes de que qualquer homenagem que se faça à Gisela é algo muito pequeno diante da grandeza do ser humano que ela foi. Mesmo assim, dedicamos-lhe este volume da revista Domínios de Lingu@gem, que reúne trabalhos de membros do Grupo de Trabalho de Fonética e Fonologia da Anpoll, cuja existência, produtividade e atuação na linguística nacional também é devida à nossa Gisela Collischonn, que esteve à frente do Grupo por quatro anos, em um trabalho conjunto com a Professora Thaís Cristófaró Silva (UFMG).

Este volume reúne doze artigos inéditos de autoria de pesquisadores brasileiros que têm se dedicado aos estudos de Fonética e de Fonologia, seja do ponto de vista descritivo, seja abordando diferentes dimensões teóricas dos mais recentes modelos e instrumentos de análise.

Para a abertura deste volume, como forma de deferência não apenas pelo seu trabalho, mas também por sua proximidade com Gisela Collischonn, apresentamos o trabalho de Luiz Carlos Schwindt, *Sobre a preservação de expoentes morfológicos da fonologia variável do português brasileiro*. Há muitos anos, o autor vem investigando fenômenos de natureza morfofonológica no português, com trabalhos caracterizados por uma criteriosa análise de dados e profundas reflexões teóricas. Também um olhar atento a questões relativas à variação no português brasileiro perpassa vários trabalhos do autor. Neste artigo, Schwindt aborda dois fenômenos variáveis conhecidos no português brasileiro: a desnasalização de ditongos átonos finais, caso de “homem ~ homi”; “pedem ~ pedi”, e o apagamento de ‘r’ em coda final tônica, caso de “amor ~ amoØ”; “amar ~ amaØ”, como evidências de que expoentes de morfemas monossegmentais são mais protegidos contra apagamentos do que porções fonológicas distribuídas em unidades morfológicas maiores. Para tanto, o autor parte do pressuposto de que processos fonológicos variáveis podem acessar informações morfológicas, apresentando três questões norteadoras de suas investigações: (i) expoentes de morfemas monossegmentais são protegidos de – ou são menos suscetíveis a – apagamentos variáveis? (ii) pode-se falar em efeitos compensatórios no caso de apagamento de morfemas? Nesse sentido, há, de fato, apagamento, ou se trata de processo gradiente que deixa resíduo fonético? (iii) quais as alternativas para lidar formalmente com a proteção ao apagamento de morfemas? Após discutir os fenômenos apresentados e refletir sobre estas questões, propõe-se uma formalização na perspectiva da Teoria da Otimidade, pelo viés da teoria da correspondência (McCARTHY; PRINCE, 1995).

O segundo artigo – *O mapeamento fonético-fonológico das vogais postônicas finais no português brasileiro* – de Carmen Lúcia Matzenauer, retoma uma reflexão de Camara Jr (1970), segundo a qual o funcionamento das vogais átonas é um dos problemas mais intrincados da fonêmica portuguesa no Brasil. Tão certo estava Camara Jr que, ainda hoje, existem inúmeras investigações com o objetivo de descrever a riqueza dos fenômenos, por vezes, variáveis que têm como alvo as vogais átonas do português brasileiro. Em seu trabalho, Matzenauer traz à tona o subsistema átono final /i, u, a/ de modo a ponderar sobre tipologias de línguas, sobre o processo de aquisição da linguagem, e também o mapeamento entre os níveis fonético e fonológico dessas vogais, a partir de dados de produção e de percepção dos segmentos vocálicos por falantes nativos de português do Brasil. Para a análise dos dados, a autora se vale do Modelo Bidirecional de Processamento de Língua Materna – BiPhon – proposto por Boersma (2007, 2011) e Boersma e Hamann (2009). Ao final, são apresentados quatro tipos de evidências que podem oferecer suporte para a escolha do subsistema de três vogais átonas finais.

Na sequência, o trabalho de Arthur Pereira Santana – *Uma abordagem fonológica para as postônicas médias não-finais* – continua a tratar das vogais átonas do português brasileiro. Diferentemente de Matzenauer no artigo anterior, que tratou do subsistema átono final, Santana volta-se para as vogais médias em sílaba átona postulando que, enquanto diversos estudos a respeito do vocalismo no português brasileiro tratam das vogais em posição pretônicas, as postônicas não-finais carecem de análises que consigam captar formalmente fatos como a alternância entre médias-altas e altas em todos os dialetos do País. O autor trabalha com dados de fala controlada, via experimentos, de vinte informantes de São Paulo (SP) e vinte de São Luís (MA), objetivando comparar os resultados na busca de similaridades e diferenças entre os dois dialetos. Ressalta-se, contudo, que a opção foi discutir apenas variáveis de natureza fonológica, quais sejam: ponto de articulação da vogal tônica, o ponto de articulação da vogal átona final, o ponto de articulação do contexto fonológico precedente, o ponto de articulação do contexto fonológico seguinte, a altura da vogal tônica e a altura da átona final.

No artigo *A estrutura silábica em esperanto*, Karina Gonçalves de Souza de Oliveira começa fazendo um preâmbulo acerca das razões por que existem as línguas planejadas; no caso do esperanto, para servir de língua auxiliar internacional. O objetivo primeiro do trabalho é retomar a literatura fonológica desta língua a fim de realizar uma discussão sobre sua estrutura silábica, ao que a autora conclui ser, no esperanto, maximamente (C)(C)V(C)(C). Para

caracterizar a organização interna do constituinte sílaba, buscaram-se construtos teóricos já canônicos na literatura, tais como Ataque e Coda.

Geovana Soncin e Luciani Tenani, em *Variações de F0 e configurações de frase entoacional: análise de estruturas contrastivas*, consideram questões relativas ao fraseamento prosódico de sentenças e ao modo como diferentes fraseamentos alteram os contornos da entonação, para descreverem as variações de F0 em sentenças de mesma sequência segmental, mas que se diferem estruturalmente da frase entoacional. Em especial, o trabalho busca verificar de que modo os parâmetros acústicos de F0 contribuem para a configuração de sentenças nas quais uma mesma cadeia segmental pode se organizar em diferentes frases entoacionais, considerando-se a alteração de suas fronteiras internas à mesma sequência sintática, em função dos componentes semânticos que as perpassam, como (i) a. não espere; b. [não espere]; c. [não] [espere] e (ii) a. isso só ele resolve; b. [isso] [só ele resolve]; c. [isso só] [ele resolve]. Para desenvolver seu trabalho, as autoras valem-se, metodologicamente, da fonologia de laboratório, enquanto que as pressuposições teóricas estão alicerçadas na Fonologia Prosódica, de Nespor e Vogel (1986).

No artigo *Análise perceptiva e acústica em fonética forense: uma pesquisa em disfarce de voz*, Maria Lúcia de Castro Gomes, Denise de Oliveira Carneiro e Andrea Alves Guimarães Dresch envolvem-se em métodos de análise perceptiva e acústica para introduzirem alunos de graduação em Letras no campo da fonética forense. Para tanto, comparam a voz de cinquenta falantes, distribuídos em gênero e em faixa etária, na simulação de um telefonema para pedido de resgate por sequestro, em vozes normais e disfarçadas, a partir de um texto com setenta e cinco palavras. Os resultados obtidos a partir dos experimentos realizados levam as autoras à conclusão de que é necessário que mais pesquisadores invistam em abordagens multidisciplinares na análise de fala em fonética forense.

Na sequência, Vera Pacheco, Marian Oliveira e Tássia da Silva Coelho, no artigo *Salvador, Vitória da Conquista e Teófilo Otoni: cidades e falares diferentes? Uma análise discriminante da F0*, voltam-se à investigação das vogais /i/, /a/ e /u/, agora a partir de parâmetros acústicos, observando sua realização em sílabas tônicas e pretônicas. As autoras buscam, assim, avaliar o papel da frequência fundamental F0 na delimitação dos falares de três cidades, a saber, Teófilo Otoni, no noroeste de Minas Gerais; Vitória da Conquista, no sudoeste da Bahia; e Salvador, a capital da Bahia, localizada mais ao norte deste estado. Justificando-se pelo fato de essas três cidades terem algum elo de natureza social, Pacheco, Oliveira e Coelho

questionam se esses três importantes centros populacionais podem ter outro tipo de elo revelado por F0, perseguindo a hipótese de que falantes de Vitória da Conquista teriam padrão de F0 intermediário entre o padrão dos falantes de Teófilo Otoni e o dos soteropolitanos, já que a primeira cidade recebe, costumeiramente, cidadãos das outras duas. Sobre isto, concluem que o falar conquistense pode ser caracterizado como baiano-mineiro. O quadro de informantes compõe-se de seis sujeitos de 25 a 40 anos, sendo um homem e uma mulher, naturais e sempre residentes, de cada uma das cidades e com nível superior de escolaridade, finalizado ou em andamento.

O trabalho *A aquisição do sistema vocálico do português por falantes nativos da variedade rio-platense de espanhol: uma discussão sobre a bidirecionalidade da transferência vocálica*, de Letícia Pereyron e Ubiratã Kickhöfel Alves, parte da perspectiva dinâmica que avalia a língua como um sistema adaptativo complexo, em que mudanças se dão em ciclos contínuos e ilimitados, para verificar esse dinamismo no que respeita a aquisição do português brasileiro porto alegreense (L2) por falantes do espanhol rio-platense (L1). Com isso, os autores pretendem investigar de que modo uma língua adquirida posteriormente – o português – influencia sobre o sistema materno – o espanhol, vislumbrando identificar a possibilidade de transferência bidirecional (L2-L1 e L1-L2) entre esses sistemas. Para tanto, os autores comparam os sistemas vocálicos (valores de F1, F2 e duração) de falantes monolíngues de espanhol rio-platense, residentes na Província de Buenos Aires (Argentina), com a produção, em espanhol (L1) e português (L2), por falantes deste mesmo dialeto de língua-materna, residentes na cidade de Porto Alegre.

No artigo *A Aquisição Fonológica Variável da Nasal Velar por Aprendizes de Inglês-L2: Análise pela Teoria da Otimidade Estocástica*, Athany Gutierrez busca, no algoritmo de aprendizagem da chamada Teoria da Otimidade Estocástica (BOERSMA; HAYES, 2001), elementos para realizar uma análise da aquisição da consoante nasal velar /ŋ/ do inglês (L2) por aprendizes falantes do português brasileiro (L1). Considerando que as nasais em posição de coda em inglês possuem estatuto de fonema, dada a sua função distintiva neste sistema, o que não acontece no português brasileiro, a investigação é norteadada pelo questionamento acerca de como acontece a aquisição da consoante velar na interlíngua português-inglês. O *corpus* adveio da fala gravada de aprendizes de inglês, divididos em um grupo com cinco informantes de nível básico e outro com cinco estudantes de nível pré-intermediário. No total, foram gravadas dez

horas de fala de cada grupo em encontros de conversação, sendo o *corpus* tratado conforme a análise variacionista laboviana.

Leonice Passarella e Rosane Silveira, no trabalho *Nuclear stress placement by Brazilian users of English as an international language*, trazem à tona importante discussão acerca da competência de falantes de inglês como língua internacional. O trabalho apresenta resultados de um estudo piloto destinado a investigar o modo como o acento nuclear fora produzido por falantes brasileiros em nível intermediário de proficiência de inglês em interação com outros falantes brasileiros de inglês, estes como ouvintes. Após a produção de cento e sessenta assertivas gravadas em áudio, as autoras recorreram ao *software* Praat de modo a efetuarem uma descrição acústica e auditiva dos dados, visando verificar se o acento nuclear fora corretamente alocado conforme os contextos discursivos; disso concluíram que os falantes demonstraram dificuldades nessa tarefa, o que pode comprometer a forma como esses mesmos falantes são interpretados quanto interagindo em inglês.

No artigo *Adaptações fonológicas na pronúncia de nomes comerciais com elementos do inglês no Brasil*, Natália Cristine Prado também traz a língua inglesa como central para sua investigação. A autora busca analisar a formação de nomes comerciais com elementos do inglês em português brasileiro, atentando para como os empréstimos se comportam fonologicamente na língua falada no Brasil, especialmente nos casos em que se percebe a ocorrência de epêntese e apagamento de algum elemento. Nomes como *Libertway Motel*, *Click Computadores*, *Fast Printer*, *Baby Shop*, *Big Mix* e *Clean Plus* são alguns exemplos que se encontram no *corpus* analisado. Prado observa que o falante do português brasileiro tende a realizar epênteses e apagamentos para “resolver” as sílabas inglesas que não são possíveis nessa língua; observou ainda que o apagamento ocorre bem menos que a epêntese, o que leva à conclusão de que, para salvaguardar a estrutura interna da sílaba em português, o falante brasileiro prefere utilizar com estratégia a inserção de elementos.

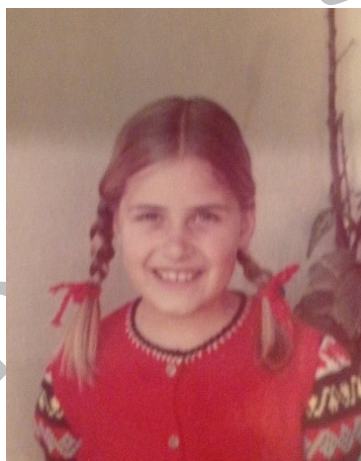
Fechando os doze artigos que compõem este volume da revista *Domínios de Lingu@gem*, edição especial de Fonética e Fonologia dedicada a Gisella Collischon, Reiner Vinicius Perozzo e Ubiratã Kickhöfel Alves, no artigo *Uma discussão acerca da aplicação do Perceptual Assimilation Model-L2 à percepção fônica de língua estrangeira: questões de pesquisa e desafios teóricos*, voltam-se à pesquisa com dados de percepção. Os autores iniciam seu trabalho com um breve histórico acerca dos estudos de percepção da fala, referindo, ao final desta introdução, a preponderância de três modelos perceptuais da fala não nativa no âmbito

das pesquisas nacionais, entre os quais é dada especial atenção ao PAM-L2, ou Modelo de Assimilação Perceptual da Aprendizagem da Fala em Segunda Língua (BEST; TYLER, 2007), cujo foco recai sobre aprendizes de uma segunda língua (L2) que estão adquirindo o sistema fônico alvo, com o postulado fundamental de que a aprendizagem perceptual é determinada por princípios diferentes daqueles do idioma materno. O objetivo central do artigo é apontar possíveis alterações ao PAM-L2, de modo que este dê conta da percepção fônica em contexto de línguas estrangeiras.

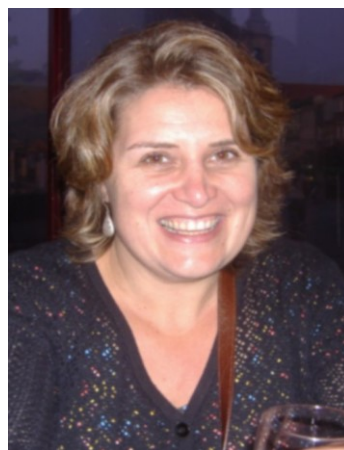
Todos os artigos acima apresentados revelam a solidez dos estudos em Fonética e Fonologia que têm sido desenvolvidos no Brasil, especialmente no âmbito do Grupo de Trabalho da Anpoll. Esperamos que os trabalhos aqui presentes sejam sempre consultados para que sirvam de alicerce a novas pesquisas, o que, certamente, dará continuidade ao legado de Gisela Collishonn, pesquisadora responsável pela refundação do GT, à frente do qual esteve por quatro anos.

Por fim, novamente saudamos e homenageamos nossa colega Gisela que tão cedo partiu. Contudo, apesar do pouco tempo que esteve entre nós, deixou-nos um imensurável legado de generosidade, sabedoria e competência que permanecerá sempre vivo em nossas lembranças e que este volume da Revista Domínios de Lingu@gem contribui para eternizar.

À Gisela Collishonn



☆ 02 de abril de 1964
(Lajeado, RS)



† 15 de junho de 2016
(Porto Alegre, RS)

Referências

- BEST, C.; TYLER, M. Nonnative and Second-Language Speech Perception: Commonalities and Complementarities. In: BOHN, O.; MUNRO, M. (Orgs.). **Language Experience in Second Language Speech Learning: In honor of James Emil Flege**. Filadélfia: John Benjamins Publishing Company, 2007.
- BOERSMA, P. Cue constraints and their interactions in phonological perception and production. Rutgers Optimality Archive 944, 2007.
- BOERSMA, P. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: BENZ, A. & MATTAUSCH, J. (eds.) **Bidirectional Optimality Theory**, 33-72. Amsterdam: John Benjamins, 2011
- BOERSMA, P.; HAMANN, S. In: CALABRESE, A.; WETZELS, W. L. (eds.) **Loanword phonology**. Amsterdam: John Benjamins, 2009.
- BOERSMA, P.; HAYES, B. Empirical Tests of the Gradual Learning Algorithm. **Linguistic Inquiry** 32, p. 45-86, 2001.
- CAMARA JR, J. M. **Estrutura da Língua Portuguesa**. Petrópolis: Vozes, 1970.
- COLISCHONN G. **Um estudo do acento secundário em português**. Dissertação (mestrado), Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1993.
- COLISCHONN, G. **Análise prosódica da sílaba em português**. Tese (Doutorado). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.
- HALLE, M.; VERGNAUD, J. R. **An essay on stress**. Cambridge, Mass.: Mit Press, 1987.
- ITÔ, J. **Syllable Theory in Prosodic Phonology**. Tese de doutorado. University of Massachusetts, 1986.
- ITÔ, J. A Prosodic Theory of Epenthesis. **Natural Language and Linguistic Theory**, 1989.
- MAGALHÃES, J. Main stress and secondary stress in Brazilian and European Portuguese. In WETZELS, W. L; MENUZZI, S.; COSTA, J. (eds) **The handbook of Portuguese linguistics**. Wiley-Blacwell, 2016, p. 107-124.
- NESPOR, M.; VOGEL, I. **Prosodic Phonology**. Dordrecht, Holanda: Foris, 1986.
- PIGOTT, G. Epenthesis and syllable weight. **Natural Language and Linguistic Theory**, v.13, p. 283-326, 1995.

Sobre a preservação de expoentes morfológicos na fonologia variável do português brasileiro

On the preservation of morphological exponents in the Brazilian Portuguese variable phonology

Luiz Carlos Schwindt*

RESUMO: A partir do pressuposto de que processos fonológicos variáveis podem acessar informações morfológicas, defendemos, neste texto, que expoentes de morfemas monosssegmentais são mais protegidos contra apagamentos do que porções fonológicas distribuídas em unidades morfológicas maiores. Dois fenômenos variáveis bastante debatidos em português brasileiro são analisados como evidências: a desnasalização de ditongos finais átonos (ex. homem ~ homi; pedem ~ pedi) e o apagamento de r em coda final tônica (ex. amor ~ amo∅; amar ~ ama∅). Propomos a formalização desses fenômenos na perspectiva da Teoria da Otimidade, numa abordagem em que restrições de natureza morfológica interagem com restrições fonológicas, a fim de assegurar consistência de exponência.

PALAVRAS-CHAVE: Consistência de Exponência. Desnasalização. Apagamento de r. Teoria da Otimidade.

ABSTRACT: From the assumption that phonological variation can access morphological information, we advocate in this text that monosegmental morphemes are more protected from deletion than morphemes distributed in greater phonological structures. Two variable phenomena of Brazilian Portuguese quite discussed in the literature are analyzed here as evidence: denasalization of final unstressed diphthongs (eg. homem ~ homi 'man'; pedem ~ pedi 'they ask') and r-deletion in stressed final codas (eg. amor ~ amo∅ 'love'; amar ~ ama∅ 'to love'). We propose a formal approach for these data in the Optimality Theory framework, assuming that morphological constraints interact with phonological ones in order to ensure Consistency of Exponence.

KEYWORDS: Consistency of Exponence. Denasalization. r-deletion. Optimality Theory.

1. Introdução

O modo como a fonologia interage com a especificação lexical de morfemas não é tema trivial. A ideia de que morfemas têm correspondentes necessários em nível fonológico está presente nas análises fonológicas e morfológicas formais desde o estruturalismo até a Teoria da Otimidade clássica (OT). Na abordagem estruturalista, estava em jogo o respeito à exigência de univocidade nas relações entre as unidades do plano da expressão e do plano do conteúdo;

* Professor do Departamento de Linguística, Filologia e Teoria Literária do Instituto de Letras da Universidade Federal do Rio Grande do Sul e pesquisador do Conselho Nacional de Desenvolvimento Científico e Tecnológico.

desencontros nesse mapeamento eram de tal modo rejeitados que justificaram a proposição de expedientes analíticos como morfemas \emptyset (quando a um morfema não correspondia som algum) ou de morfes vazios (quando a um som não correspondia morfema claramente identificável). Em trabalho inaugural em OT, McCarthy e Prince (1993), num modelo que se tornou conhecido como PARSE/FILL ou Teoria da Contenção, propuseram um princípio chamado Consistência de Exponência, que pretendia dar conta da exigência de pareamento entre morfologia e fonologia, definido nos termos de (1).

(1) Consistência de Exponência

Mudanças no expoente de um morfema fonologicamente especificado não são permitidas. (MCCARTHY; PRINCE, 1993:21)

De acordo com esse princípio, qualquer expoente fonológico de um morfema especificado no input (segmentos, moras etc.) deve ter correspondente idêntico no output.¹ Segmentos epentéticos, portanto, não têm afiliação morfológica, e são interpretados como implementações fonéticas de posições vazias. Porções fonológicas não pronunciadas, por outro lado, que permanecem presentes no candidato, apenas falham ao serem incorporadas à estrutura prosódica (ou seja: expoentes de morfemas não são apagados).

O modelo PARSE/FILL foi substituído por uma subteoria mais poderosa, envolvendo relações mais abstratas, conhecida como MAX/DEP ou Teoria da Correspondência (MCCARTHY; PRINCE, 1995). De acordo com essa nova abordagem, GEN pode produzir candidatos que contemplem mudanças na afiliação morfológica de um segmento ou mesmo apagamentos de segmentos associados a morfemas.

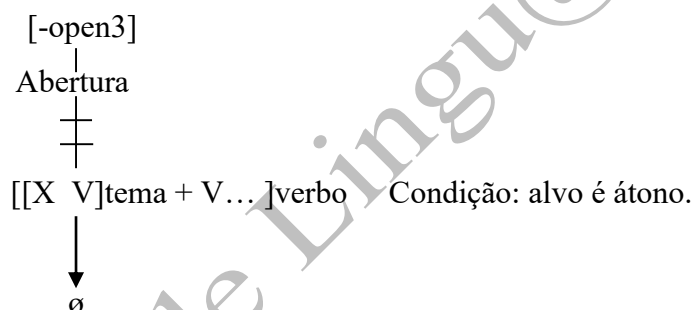
Apesar de não mais se assumir explicitamente consistência de exponência como um princípio, a ideia de que a afiliação morfológica de material fonológico deve ser em alguma medida preservada permaneceu na literatura da OT, refletida em restrições de diferentes naturezas, como FAITH-ROOT, FAITH-AFFIX (MCCARTHY & PRINCE, 1995); FAITH-NOUN (SMITH, 2001); ALIGN-M, Φ /ANCHOR (MCCARTHY & PRINCE, 1993, 1995); MORPH-REAL

¹ Definimos *expoente* nos termos de Mathews (1974): considerando-se morfemas como categorias, definidas, em termos individuais, por propriedades morfossintáticas abstratas (p. 136), podemos dizer que os traços que identificam essas propriedades podem ser referidos como seus **expoentes** (p. 144).

(SAMEK-LODOVICI, 1993); MAX-ROOT (MCCARTHY, 2011) / MAX-AFFIX (FARWANEH, 2007) / MAX-M (WOLF, 2008); MORPHEME INTEGRITY (MUTAKA & HYMAN, 1990) etc.²

Em fenômenos categóricos (aqui entendidos como de aplicação obrigatória), esse princípio parece ser respeitado em diferentes línguas. No sistema categórico do PB, um bom exemplo de evitação/reparo a apagamento de morfema é o da estabilidade do traço de abertura da VT de 2ª e 3ª conjugações na 1ª pes. do singular do pres. de indicativo, que é espriado para a vogal da raiz, gerando um padrão harmonizado, opaco na superfície (ex. *seg+i+o* > *sig∅o*), conforme representação em (2), adaptada de Wetzels (1995) por Schwindt e Wetzels (2016).

(2) Truncamento com estabilidade do traço de abertura.



Neste texto, sem a pretensão de esgotar o tema, interessa-nos particularmente o efeito de consistência de exponência sobre processos fonológicos variáveis. Assumindo o pressuposto de que tais processos podem acessar informação morfológica (GUY, 1980; LABOV, 1981; KIPARKSY, 1988, 1995; SCHWINDT, 2012), levantamos as questões que seguem, dependentes entre si.

- (i) Expoentes de morfemas (neste recorte, monossegmentais) são protegidos de – ou são menos suscetíveis a – apagamentos variáveis?
- (ii) Pode-se falar em efeitos compensatórios no caso de apagamento de morfemas? Nesse sentido, há, de fato, apagamento, ou se trata de processo gradiente que deixa resíduo fonético?
- (iii) Quais as alternativas para se lidar formalmente com a proteção ao apagamento de morfemas?

² Segundo van Oostendorp (2007), isso se deve ao fato de que esse princípio tem poucas exceções nas línguas do mundo.

Para discutir essas questões, o texto assume a organização que segue. Na seção 2, partindo de dois exemplos clássicos – o apagamento de *t/d* em inglês, descrito por Guy (1980) e o apagamento de *s* em espanhol porto-riquenho (EPR), descrito por Poplack (1980), e em português brasileiro (PB), descrito por Guy (1981) –, apresentamos os dois fenômenos que alimentam nossa análise: a desnasalização de ditongos átonos finais e o apagamento de *r* em sílaba final de palavra em PB. Na seção 3, apresentamos uma análise em moldes otimalistas, focando mais a natureza de restrições em competição do que uma proposta de tratamento da variação em OT. Por fim, em 4, apresentamos nossas ponderações sobre a análise e apontamos para questões residuais.

2. Fenômenos variáveis sujeitos a Consistência de Exponência

Dois exemplos emblemáticos de fenômenos fonológicos variáveis sujeitos a algum tipo de condicionamento morfológico são o apagamento de *t/d* em inglês e o apagamento de *s* em variedades do espanhol e do português. Esses fenômenos, contudo, encerram uma aparente contradição: enquanto Guy (1980) observou que o processo se aplica menos às coronais do inglês que correspondem a um morfema em isolado (a marca de passado), Poplack (1980), para o EPR, e Guy (1981), para o PB, constataram maior incidência de apagamento de *s* final em formas em que esse segmento corresponde a um morfema (a marca de plural), conforme se vê em (3) e (4).

- (3) Apagamento de *t/d* em inglês (GUY, 1980)
- | | | |
|-----------------|-------------------------------|------|
| monomorfemas | <i>mist</i> ~ <i>mis</i> ∅ | 0.64 |
| passado regular | <i>missed</i> ~ <i>miss</i> ∅ | 0.32 |

- (4) Apagamento de *s* no SN em EPR (POPLACK, 1980) e em PB (GUY, 1981)
- | | EPR | | PB | |
|------------------|--------------------------------------|-----|----------------------------------|-----|
| monomorfema | <i>mes</i> ~ <i>me</i> ∅ | 54% | <i>menos</i> ~ <i>meno</i> ∅ | 15% |
| flexão de plural | <i>las casas</i> ~ <i>las casa</i> ∅ | 68% | <i>os dias</i> ~ <i>os dia</i> ∅ | 36% |

Sustentamos que essa contradição é apenas aparente, já que, nos dois casos, a informação morfológica parece sujeita a alguma força preservadora. No caso de *t/d* em inglês, como o locus do processo parece ser a palavra, o segmento associado ao morfema é menos apagado do que segmentos que fecham a raiz (e mesmo do que segmentos finais de verbos

irregulares); no caso do espanhol, por outro lado, o locus do processo parece ser o sintagma, mais do que a palavra, sendo a informação de plural preservada em algum elemento (de modo geral, o primeiro, conforme descreve ampla literatura a respeito).

Poplack e Guy problematizaram esses dois fenômenos frente a um princípio proposto por Kiparsky (1982[1972]) para a diacronia linguística, conhecido como "distintividade contrastiva", segundo o qual estruturas gramaticais redundantes são menos resistentes à mudança. Poplack (1980) constatou que, ainda que com pequena vantagem, o *s* em *EPR* estava mais sujeito a apagamento em ambientes passíveis de desambiguação (*las reinas mandan*) do que em ambientes ambíguos (*arroz con abichuela(s)*). Guy (1996), apesar de constatar que os segmentos *t/d* finais em inglês são mais preservados quando coincidem com morfemas, rejeita "distintividade contrastiva" enquanto argumento funcional para explicar o fato, uma vez que formas do particípio (*have talked*) – que podem ser consideradas redundantes, já que recuperáveis sempre por um verbo auxiliar na sentença – apresentam índices muito semelhantes aos de apagamento de formas de passado simples. O autor assume, em oposição, um argumento formal, baseado em acolchetamento (ou em níveis) para dar conta dos diferentes domínios de aplicação da regra.

Independentemente, contudo, de se concluir em favor de um argumento funcional ou formal para lidar com esses fenômenos, a discussão em torno da hipótese de "distintividade contrastiva" contribui para a crença de que mecanismos preservadores de morfemas podem atuar também no componente variável da linguagem.³

Isso posto, e assumindo-se a necessidade de formalizar esses processos numa arquitetura de gramática plausível, duas questões importantes emergem:

- (i) o morfema, que não é mais identificável na superfície, foi de fato apagado ou simplesmente não foi inserido?
- (ii) quantos processos ou regras estão por trás de outputs idênticos na superfície da língua?

Frente a essa realidade, examinamos, aqui, dois fenômenos fonológicos variáveis bastante estudados em PB, sob diferentes perspectivas: a desnasalização de ditongos finais (DF), que mereceu, entre diversas outras, as análises de Guy (1981), Scherre e Naro (1998), Battisti (2002), Schwindt e Bopp da Silva (2010) e Cristófar-Silva, Fonseca e Cantoni (2012),

³ Sobre detalhes em torno dessa discussão, ver Schwindt (2015).

e o apagamento de *r* em final de palavra (AR), que foi analisado, também entre diversos outros, por Callou et. al. (1996), Monaretto (1997, 2002), Silveira (2010) e Serra e Callou (2015). Para fins desta discussão, vamos nos orientar preponderantemente pelos trabalhos de Schwindt e Bopp da Silva (2010), para tratar de DF, e de Monaretto (2002) e Serra e Callou (2013), para tratar de AR.

2.1 Desnasalização de ditongos átonos em final de palavra

Um processo comum em todas as regiões do Brasil é a desnasalização de ditongos finais átonos, que se aplica tanto a nomes quanto a verbos (ex. *viagem* ~ *viagi*; *pedem* ~ *pedi*). Este fenômeno foi descrito por diferentes autores com resultados semelhantes. A maioria dos estudos aponta para algum papel da morfologia, que diferencia o processo em nomes e verbos, ao lado de outros fatores fonológicos, como a existência de uma consoante palatal no contexto precedente. Essa diferenciação entre classes gramaticais certamente é influenciada pelo fato de que, no âmbito dos verbos, para além do processo fonológico em foco, um aspecto de natureza morfossintática está envolvido, a concordância com as formas de plural, que caracteristicamente se exponenciam por uma nasal em português.

Em (5), a seguir, estão as taxas de aplicação do processo em toda a região sul do Brasil, levantadas por Schwindt e Bopp da Silva (2010): a nasal que corresponde à marca de plural nos verbos mostra-se menos sujeita a apagamentos do que nasais que fazem parte da raiz nominal ou de sufixos nominais.⁴

(5) Redução da nasalidade no sul do Brasil (SCHWINDT; BOPP DA SILVA, 2010)

não-verbos	<i>homem</i> ~ <i>home</i> ∅ / <i>camaradagem</i> ~ <i>camaradage</i> ∅	56%
verbos	<i>pedem</i> ~ <i>pede</i> ∅ / <i>falaram</i> ~ <i>falaru</i> ∅	32%

Nos verbos, encontramos, muitas vezes, evidências de precedência do processo de concordância sobre o de desnasalização, isto é, ao lado de formas como (*eles*) *cantam* ~ *canta* / *pedem* ~ *pedi*, encontramos também formas como (*eles*) *cantu* / *pedi*, em que [u]/[i] finais parecem provir da semivogal do ditongo, segmento que, em princípio, se produz a partir da

⁴ Guy (1981) encontrou resultado inverso: o processo se mostrou mais favorável entre verbos. Deve-se sublinhar, contudo, que o autor valeu-se de dados de adultos em fase de alfabetização – o que sugere que a face morfossintática do processo, aquela que diz respeito à variação na concordância verbal, tinha mais chances de se apresentar.

relação entre a vogal final do tema, neste caso, e a nasal que a segue, o sufixo número-pessoal (aN > ãw̃N > ãw̃Ø > w̃ũ > ɔ / eN > ẽjN > ij̃Ø > ij̃ > ɪ).⁵ Esse fato não apenas se constitui como evidência de precedência do processo de concordância sobre o de apagamento, mas também abre espaço para a hipótese de efeito compensatório no caso de DF em verbos, isto é, permite se postular que, independente de haver uma regra variável de concordância, o processo fonológico implicado preserva a informação morfológica a partir de um resíduo da marca de concordância. Se isso é verdade, por outro lado, casos como os de *(eles) canta* seriam as legítimas ilustrações de um processo genuinamente morfossintático, que poderia resultar de desnasalização, mas que também poderia ser analisado à luz de uma hipótese de reestruturação das marcas flexionais (como resultado de nivelamento morfológico, por exemplo). Além disso, parecem ser essas formas mais influenciadas por fatores sociais como a escolaridade.⁶

Nos não-verbos, DF aplica-se basicamente a ditongos com vogal coronal (*homem* ~ *homi* / *viagem* ~ *viagi* / *órfão* ~ ?*órfu*). Ao contrário do que ocorre com os verbos, há evidências de que essas formas estão sujeitas a algum tipo de neutralização, já que nenhum expoente morfológico está em jogo quando do apagamento (ẽjN > eØ > ɪ). A propósito disso, ao contrário do que ocorre com os verbos, a pronúncia de ditongo homorgânico ou de monotongo nasalizado, no caso dos não-verbos, é mais rara (ex. ?*hom*[ij̃] / ?*hom*[ĩ], ?*órf*[ĩw̃] / ?*órf*[õ]), o que sugere que o processo que desnasaliza verbos, apesar de possuir uma face restrita pela morfologia, é, do olhar da superfície, foneticamente menos restrito do que o que desnasaliza nomes.⁷

Tomado esse quadro descritivo, assumimos tese anteriormente defendida (SCHWINDT, 2015), de que estão em jogo, no caso de DF, dois processos diferentes, que resultam, em nível de superfície, em “outputs convergentes”. O mais restrito desses processos, que só atinge

⁵ Uma discussão importante, mas que omitimos aqui por questões de foco e de espaço, diz respeito ao que se entende, do ponto de vista fonético-fonológico, por desnasalização, já que o segmento nasal não parece ser plenamente pronunciado em português em nenhum caso. Para fins deste texto, consideramos suficiente dizer que só não podem ser consideradas *reduzidas* as formas que combinem ditongo e ressonância nasal. Formas apenas com ditongo, mas sem nasalidade, não ocorrem, provavelmente pela ordem que se estabelece entre os fenômenos envolvidos; formas com monotongo mas que preservam nasalidade parecem mais raras – e, de todo modo, são difíceis de serem diferenciadas, de oitiva, das que se realizam com ditongos nasais homorgânicos.

⁶ Não incluímos nesta afirmação o ditongo com vogal coronal (ex. *(eles) ped*[i]), porque neste caso sempre resta ambígua a análise, uma vez que a vogal final pode ser atribuída tanto à semivogal [j] quanto à vogal final do tema [e] – neste segundo caso não havendo evidência explícita de precedência da desnasalização. Exames de dialetos em que a redução da átona final é menos frequente, como o de Curitiba, por exemplo, em que se pode encontrar a pronúncia *(eles) ped*[e], pode contribuir para confirmar essa dupla possibilidade de análise.

⁷ Isso pode estar simplesmente indicando, numa perspectiva derivacional de gramática, que verbos se sujeitam mais vezes ao processo do que não verbos, à semelhança do que propôs Guy (1991) para o apagamento de *t/d* em inglês.

verbos, tem acesso à morfologia interna da palavra (com efeitos sobre a interpretação de variação de concordância verbal) – por isso mesmo pode preservar expoentes verbais. Um segundo processo, que atinge verbos e não verbos, tem influência do contexto precedente (preponderantemente consoantes palatais), está sujeito a gradiência fonética (a gama de possibilidades de pronúncia da vogal final resultante da redução) e não vê morfologia interna (não diferencia o segmento que pertence à raiz ou ao sufixo).

2.2 Apagamento de r em coda final

Também bastante frequente em todos os dialetos do PB, o processo de apagamento de *r* em coda, como dissemos, foi objeto de diversas análises, sob diferentes vieses. Neste texto, fixamo-nos sobre o contexto de final de palavra tônico: ao contrário de DF, AR ocorre preponderantemente em formas verbais, conforme se pode ver em (6), onde se comparam os resultados das amostras de dois dialetos com comportamentos bastante distintos em relação ao fenômeno, Monaretto (2002) e Serra e Callou (2013).

- (6) AR em Porto Alegre (MONARETTO, 2002) e no Rio de Janeiro (SERRA; CALLOU, 2013)

		POA	RJ
verbos	<i>amar</i> ~ <i>ama</i> ∅	76%	81%
não-verbos	<i>amor</i> ~ <i>amo</i> ∅	2%	66%

O fato de que distinguir verbos e não verbos, mas também o fato de a distinção de aplicação entre essas categorias ser expressiva apenas no dialeto do RJ, se comparado ao RS, podem ser tomados como evidências, novamente aqui, para a proposição de dois processos distintos. O processo mais restrito tem, como no caso de DF, acesso à informação morfológica. O processo mais geral é de natureza fonética e não faz distinção de classe de palavra. Trata-se de um apagamento que resulta da vulnerabilidade de um tipo específico de margem e que está sujeito a gradiência particularmente nos dialetos em que se contata aspiração do *r* ($r > h > \emptyset$) – caso do RJ, mas não no RS. Nesses dialetos, inclusive, o apagamento pode acontecer, apesar de com frequência significativamente menor, em posição medial e/ou átona (ex. *me*∅*mo* / *impo*∅*tado* / *impa*∅).

De um olhar comparativo com DF, a pergunta que se coloca é *por que, no caso de AR, não há, igualmente, preferência pela preservação do morfema – neste caso, o r de infinitivo?*

Uma explicação possível pode ser encontrada, também, na tese de efeito compensatório. Observe-se que o acento, no caso das formas infinitivas do verbo, incide sistematicamente sobre a VT, diferenciando-a de outras formas do paradigma verbal do português (compare-se *áma* a *amáØ*), repetindo, por assim dizer, a informação veiculada pelo morfema de infinitivo *r* e compensando seu eventual apagamento.

Esse argumento, porém, ainda não explica completamente a diferença entre os dois fenômenos no que tange à preservação de morfema, já que, como vimos, na DF também se observa efeito compensatório e ainda assim apagam-se menos nasais que exponenciam morfemas. Uma explicação possível pode residir no fato de que, no AR, já se parte de formas com informação morfológica redundante, enquanto na DF a emergência da vogal final desnasalizada só se explica como legítimo reparo (compensação, em sentido estrito) à redução do ditongo nasal, que carrega junto o morfema de número-pessoa. Esse argumento, porém, que se fundamenta na associação entre infinitivo e VT acentuada, mesmo parecendo verdadeiro em alguma medida, carece de maior reflexão, pois a pauta tônica se apresenta também como a preferida de não verbos, ainda que AR se aplique com muito menor frequência no âmbito dessas palavras. Na ausência de argumentos mais contundentes que sustentem uma análise baseada na consuetude interna da palavra para AR, parece plausível se supor que a gramática do PB disponha também de alguma restrição que, referindo-se à morfologia, seja capaz de distinguir verbos de nomes.

3. Uma tentativa de formalização

Diante do quadro descritivo que apresentamos, procuramos, nesta seção, representar os processos do PB que abordamos na seção anterior na perspectiva da Teoria da Otimidade standard, por acreditarmos que são fenômenos passíveis de tratamento formal.

A escolha pela OT se dá sobretudo por seu caráter não modular, isto é, nesta perspectiva teórica, a sintaxe não precede ou segue necessariamente a morfologia, por exemplo – o que nos isenta de ter de dar explicações sobre como a informação sobre concordância está presente no nível em que se analisa a palavra. Qualquer ordenamento entre níveis de análise tem de ser passível de expressão na natureza das restrições e em sua posição na hierarquia. No que diz

respeito ao mapeamento de morfemas em unidades fonológicas, tomamos como premissa a hipótese de que essas unidades estão amalgamadas no input e nos candidatos.⁸

No domínio da OT, assumimos a perspectiva da teoria da correspondência, conforme apresentado na introdução, isto é, partimos do pressuposto de que GEN produz candidatos que inserem ou apagam porções fonológicas livremente, mesmo sons que, isoladamente, se mapeiam a estrutura morfológica. Procuramos, contudo, mostrar que consistência de exponência está ativa na gramática, por força de restrições que impõem que morfemas se realizem, mesmo no âmbito da fonologia variável, de forma aberta.

Não pretendemos, ainda, neste texto apresentar argumentos para defender um modelo específico sobre o tratamento da variação em OT. Adotamos um modelo muito simplificado de gramáticas em competição, em que determinadas restrições têm sua posição na hierarquia alterada para fazer emergirem determinadas variantes. Assim, quando nos referimos a processos mais gerais ou mais específicos na gramática, estamos falando de processos que resultam de determinada hierarquia.

3.1 Desnasalização de ditongos átonos em final de palavra

Nossa intenção principal não é explicar a gramática mais fiel em se tratando de DF, qual seja, a que realiza um ditongo nasal como mapeamento da nasal, que tomamos como subespecificada em coda, seguindo a tradição de grande parte das análises que se ocupam da nasalidade em português. Sem maior problematização, pois, adotamos como ponto de partida a hipótese de que desnasalização é resultado da competição entre a restrição de fidelidade MAX, mais baixa na hierarquia, que proíbe apagamentos, e CODA-COND(ITION), restrição empregada por Battisti (2003), seguindo a formulação de Kager (1999:131), que se baseia na premissa de que a nasal, sendo subespecificada, não pode ser licenciada em coda. Essa é a gramática mais geral, que dá conta da análise de não verbos e verbos, como se vê no *tableau* em (7).

⁸ A OT também conta com uma versão realizacional, como a da Morfologia Distribuída, a saber a Optimal Interleaving / Serialismo Harmônico – OI/HS (Wolf, 2008 e McCarthy, 2011). Essa hipótese é interessante descritivamente para explicar aspectos relacionados a competição entre alomorfes, mas interpretaria possivelmente apagamentos como os que descrevemos aqui como não inserção – razão por que se descartou essa abordagem por ora. Uma perspectiva realizacional de morfologia não exclui, porém, uma gramática que se aplique, *a posteriori*, sobre formas já mapeadas. Ainda que não tencionemos explorar essa abordagem neste texto, não desejamos descartá-la enquanto perspectiva analítica.

(7) Gramática de redução da nasalidade – verbos e não-verbos (mais geral)

INPUTS		CODA-COND	MAX
pede+NP6IdPr1	a. 'pedem (ẽj/ĩj)	*!	
	☞ b. 'pede/i		*
homeN	a. 'homem (ẽj/?ĩj)	*!	
	☞ b. 'home/i		*

Como afirmamos anteriormente, porém, acreditamos que uma gramática mais específica concorre com esta, atingindo apenas verbos. Nessa gramática é que restrições de natureza morfológica têm papel mais relevante, representando o princípio de consistência de exponência. A primeira delas, mais geral, é MAX-M(SNP) (adapt. de MCCARTHY, 2011), que está numa relação de estringência com MAX geral e milita contra apagamento do morfema de número-pessoa. Esta é a gramática que traduz a preferência pela preservação da nasal que exponencia o morfema verbal em relação à nasal dos não verbos, que é parte da raiz.

(8) Gramática de redução da nasalidade verbos e não verbos – (mais específica)

INPUTS		MAX-M(SNP)	CODACOND	MAX
cata+NP6IdPr1	☞ a. 'catam (ãw/ũw)		*	
	b. 'cata	*!		*
	c. 'catu	*!		*
homeN	a. 'homem (ẽj / ?ĩj)		*!	
	☞ b. 'home/i			*

A segunda restrição que recupera consistência de exponência, mais alta na hierarquia, é MORPH-REAL (SAMEK-LODOVICI, 1993), que impõe que morfemas sejam realizados de forma aberta no output. O papel dessa restrição, aqui, é o de garantir a preferência pela forma que preserva a articulação do glide, compensando o expoente morfológico apagado.

(9) Gramática de redução da nasalidade – verbos

cata+NP6IdPr1	CODACOND	MORPH-REAL	MAX-M(SNP)	MAX
a. 'catam (ãw/ũw)	*!			
b. 'cata		*!	*	*
☞ c. 'catu			*	*

A escolha do candidato b, aquele que não possui qualquer indício do morfema de número-pessoa, em detrimento de c, em (9), envolveria dois expedientes: a demissão de MORPH-



REAL para baixo de MAX-M (SNP) e a atuação de uma restrição que penalizasse a alteração na qualidade da vogal final (do tipo IDENT-I/O ou mesmo MAX-VT). Essa gramática ilustraria o que conseguimos entender como *variação* de concordância propriamente dita, numa perspectiva de expoentes especificados (hipótese em que há sempre apagamento, ao contrário de não inserção).

Se assumirmos, porém, que consistência de exponência é uma condição que se impõe sobre as línguas em geral, como sugere van Oostendorp (2007), e que, portanto, MORPH-REAL deve estar alta na hierarquia do português, somos levados a acreditar que a gramática que licencia formas como as de (9b), *(eles) cata*, é fruto de reestruturação do *input* – o que referimos anteriormente como *reestruturação de marcas flexionais*. Esse tipo de reestruturação poderia resultar, inclusive, de *nivelamento*, como sugerem Christófaros-Silva, Fonseca e Cantoni (2012) para algumas formas envolvendo DF. Também poderia explicar a incidência de não concordância em formas salientes, no dizer de Scherre e Naro (1998).

3.2 Apagamento de r em coda final

Como mencionamos, AR é um processo típico de pauta tônica e incide mais sobre verbos do que sobre nomes. Para dar conta desse padrão, lançamos mão, na falta de restrição mais específica, de uma restrição de borda, M(ARGIN)σ-R (adaptada de De Lacy, 2001), que rejeita segmentos em posição proeminente de acordo com sua sonoridade. Como afirmamos na análise de DF, não é nosso objetivo discutir a natureza dessa restrição aqui, mas problematizar consistência de exponência em casos de fonologia variável. M(ARGIN)σ-R concorre com MAX, restrição que milita contra apagamentos, numa gramática mais geral, que dá conta do fenômeno em verbos e não verbos.

(10) Gramática de apagamento de r – verbos e não verbos (mais geral)

INPUTS		M(ARGIN)σ-R	MAX
ama+r _{Inf}	a. a'mar	*!	
	 b. a'ma		*
amor	a. a'mor	*!	
	 b. a'mo		*

Na gramática mais específica, no caso de AR, diferentemente do que ocorre com DF, são necessárias duas restrições militando a favor da supressão concorrendo com restrições de

fidelidade. Isso é assim porque é preciso que se garanta uma gramática que, interagindo com a morfologia, permita, a um só tempo, apagar o *r* dos verbos e mantê-lo nos não-verbos (situação mais típica no PB). Assim, propõe-se que ALIGN-R(STEM-PWD), restrição que exige que a borda direita do tema verbal coincida com a borda direita de uma palavra fonológica, concorra com MAX-M(INFIN), no sentido de licenciar a não realização do morfema de infinitivo em final de palavra. Isso só é possível neste caso porque o acento é garantidor da informação morfológica, o que se reflete na não violação de MORPH-REAL. Observe-se que a gramática mais geral retratada em (10) situa-se nas duas últimas posições no ranking em (11), caso em que as restrições que se referem à morfologia são vácuas.

(11) Gramática de apagamento de *r* – verbos e não verbos (mais específica)

INPUTS		ALIGN-R (STEM- PWD)	MORPH- REAL	MAX- M(INF)	M(ARGIN)σ-R	MAX
ama+r _{Inf}	a. a'mar	*!			*	
	☞ b. a'ma			*		*
amor	☞ a. a'mor					
	b. a'mo					*!

4. Principais generalizações e questões residuais

Partindo-se do pressuposto de que há evidências suficientes para não se entender mais variação fonológica como inacessível à informação morfológica e lexical, há uma demanda clara para modelagem desse acesso, já que nem todos os processos fonológicos variáveis parecem se comportar da mesma forma nesse quesito.

Há fenômenos fonológicos produzidos a partir de processos diferentes que convergem, em nível de superfície, para outputs idênticos. Muitas vezes, esses processos compartilham propriedades fonológicas, mas se diferenciam por um interagir com a morfologia e outro não. É o caso dos processos de desnasalização final e de apagamento de *r* em final de palavra.

Parece haver uma relação assimétrica entre os processos envolvidos em outputs convergentes que sugere graus de generalidade no tratamento da variação fonológica: o mais restrito, que interage com a morfologia, cabe formalmente nos mesmos condicionamentos do mais geral, de base puramente fonética, mas o contrário não é verdadeiro. Isso recupera, numa medida totalmente redimensionada, a hipótese neogramática. É preciso se responder, ainda, se há correlação necessária entre geral/mais frequente e específico/menos frequente.

No âmbito da interação com a morfologia, pode-se dizer que atua, na gramática, uma força que impede que se apaguem expoentes morfológicos. Essa força, contudo, é violável, quando a língua oferece um mecanismo de compensação para este apagamento. Essa compensação, numa perspectiva de regras, é de difícil formalização, o que leva muitos a entenderem-na como um expediente funcionalista. Numa abordagem de restrições violáveis, contudo, é possível se formalizar a relação entre apagamento/preservação/compensação de modo menos complexo do que numa perspectiva derivacional.

Sobre a proteção de expoentes morfológicos, há muito o que se discutir. Uma das questões é saber se se restringe de fato a morfemas monosssegmentais, como os que tratamos aqui e outros descritos na literatura (como a evitação de certos processos de sândi externo a monomorfemas).

Sobre a análise formal, embora pareça mais ou menos clara a vantagem das restrições sobre as regras, precisa-se dizer mais a respeito da vantagem de uma modelagem paralela (com restrições flutuantes ou gramáticas competindo) sobre a proposição de níveis (um em que a morfologia é visível, outro em que não). Outro caminho analítico desafiador, que constitui etapa da pesquisa em que esta análise se inserir, é o de uma abordagem de morfologia realizacional sobre a de uma morfologia com expoentes definidos no input.⁹

Por fim, há aspectos fonéticos que precisam ser sofisticados. O primeiro deles, grande demais talvez, é o de se definir gradiência, para que se possa usar esse parâmetro na diferenciação dos processos envolvendo outputs convergentes. Observe-se que, aqui, dizer *eles compra* ou *eles compru* está longe de ser considerado gradiência, no sentido tradicional; por outro lado, *amar* ~ *ama[h]* poderia ser considerado um caso típico de gradiência (neste caso, uma restrição da família M(ARGIN)σ-X terá lugar na gramática mais geral, abaixo de M(ARGIN)σ-líq, mostrando a preferência, nessa gramática, pela variante aspirada).

Referências Bibliográficas

BATTISTI, E. A redução dos ditongos nasais átonos. In: BISOL, L.; BRESCANCINI, C. (ed.) **Fonologia e variação**: recortes do português brasileiro. Porto Alegre: EDIPUCRS, 2002.

_____. Ditongos nasais em sílaba átona e fidelidade posicional. In: COLLISCHONN, G.; HORA, D. (ed.) **Teoria linguística**: fonologia e outros temas. João Pessoa: Ed. Universitária/UFPB, 2003.

⁹ Para análise da VT e da flexão de gênero numa abordagem de OT realizacional, ver Schwindt (2011).

CALLOU, D. et al. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: KOCH, I. (ed.). **Gramática do Português Falado**. Campinas, UNICAMP, 1996. v. VI, p. 465-493.

CRISTÓFARO-SILVA, T.; FONSECA, M. A. ; CANTONI, M. A redução do ditongo [ãw] postônico na morfologia verbal do português brasileiro: uma abordagem baseada no uso. **Letras de Hoje** (Impresso), v. 47, p. 283-292, 2012 .

DE LACY, P. Markedness in prominent positions. In MATUSHANSKY, O.; COSTA, A.; MARTIN-GONZALEZ, J.; NATHAN, L.; SZCZEGIELNIAK, A. (eds.) **HUMIT 2000, MIT Working Papers in Linguistics 40**. Cambridge, MA: MITWPL, pp.53-66, 2001. [ROA 432]

FARWANEH, S. Hypocoristics revisited: Challenging the primacy of the consonantal root. **Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic linguistics. Volume XX: Kalamazoo, Michigan, March 2006**, Edited by Mustafa A. Mughazy. [Current Issues in Linguistic Theory, 290], p. 25-49, 2007. **crossref** <http://dx.doi.org/10.1075/cilt.290.05far>

GUY, G. Variation in the group and the individual: the case of final stop deletion. In: LABOV, W. (ed.) **Locating language in time and space**. New York: Academic Press., 1980.

_____. **Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax and language history**. Tese de Doutorado. Universidade da Pennsylvania, 1981.

_____. Explanation in variable phonology: an exponential model of morphological constraints. **Language Variation and Change**, 3, p. 1-22, 1991. **crossref** <http://dx.doi.org/10.1017/S0954394500000429>

_____. Form and function in Linguistic Variation. In: GUY et al. (ed.) **Current Issues in Linguistic Theory** 127, 1996. p. 221-252.

KAGER, R. **Optimality Theory**. Cambridge: CUP, 1999.

KIPARSKY, P. **Explanation in Phonology**. Dordrecht: Foris, 1982[1972]. **crossref** <http://dx.doi.org/10.1515/9783111666242>

KIPARSKY, P. Phonological change. In: NEWMAYER, F. **Linguistics: the Cambridge survey**. vol. 1. Linguistic theory foundations. Cambridge: CUP, 1988.

KIPARSKY, P. The phonological basis of sound change, in John A. Goldsmith (ed.), **The Handbook of Phonological Theory**. Blackwell: Oxford, p. 640-670, 1995.

LABOV, W. Resolving the neogrammarian controversy. **Language** 57:267-309, 1981. **crossref** <http://dx.doi.org/10.2307/413692>

MATHEWS, P. H. **Morphology: an introduction to the theory of word-structure**. Cambridge University Press, 1974.

MCCARTHY, J. Pausal phonology and morpheme realization. In: BOROWSKY, T. KAWAHARA, S., SHINYA, T., SUGAHARA, M. **Prosody Matters: Essays in Honor of Lisa Selkirk**. London: Equinox, 2011.

MCCARTHY, J.; PRINCE, A. **Prosodic Morphology I: constraint interaction and satisfaction**. University of Massachusetts, Amherst and Rutgers Univ., 1993.

_____. Generalized Alignment. In BOOIJ, G.; VAN MARLE, J. (eds.) **Yearbook of Morphology** 1993. Dordrecht: Kluwer. p.79-153, 1995.

MONARETTO, V. N. (2002) A vibrante pós-vocálica em Porto Alegre. In: BISOL, L.; BRESCANCINI, C. (eds.). **Fonologia e variação: recortes do português brasileiro**. Porto Alegre: EDIPUCRS, p. 253-268.

MONARETTO, V. N. **Um reestudo da vibrante: análise variacionista e fonológica**. 1997. Tese de Doutorado. Porto Alegre, Pontifícia Universidade Católica do Rio Grande do Sul.

MUTAKA, N.; HYMAN, L. M. Syllable and morpheme integrity in Kinande reduplication. **Phonology**, 7: 73-119, 1990. **crossref** <http://dx.doi.org/10.1017/S0952675700001123>

POPLACK, S. The notation of the plural in Puerto Rican Spanish: Competing restraints on /s/ deletion. In: LABOV, W. (Ed.) **Locating language in time and space**. New York: Academic Press, 1980, p. 55-67.

SAMEK-LODOVICI, V. **Universal constraints and morphological gemination: a cross-linguistic study**. Manuscrito. Waltham, MA: Brandels University, 1993. [ROA-149]

SCHERRE, M. M. P.; NARO, A. J. Sobre a concordância de número no português falado do Brasil. In RUFFINO, G. (Ed.) **Dialettologia, geolinguística, sociolinguística. Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza**. Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag. 1998. 5, p. 509-523.

SCHWINDT, L. C. Zeros na morfologia nominal portuguesa à luz da Optimal Interleaving Theory. **ReVEL**, edição especial n. 5, 2011. [www.revel.inf.br].

_____. Condicionamento morfológico em fenômenos fonológicos variáveis do PB. **Letras & Letras**. Uberlândia: Editora da UFU, 2012. p. 115-127.

_____. Um output, dois processos. **Revista da ABRALIN**, v. 14, 2015, p. 553-570.

SCHWINDT, L. C.; BOPP DA SILVA, T. Panorama da redução da nasalidade em ditongos átonos finais no português do sul do Brasil. In: BISOL, L.; COLLISCHONN, G. (eds.). **Português do Brasil: variação fonológica**. 1 ed. Porto Alegre: EDIPUCRS, 2010. p. 13-33.

SCHWINDT, L. C.; WETZELS, L. The morphology and phonology of inflection. In: COSTA, J.; MENUZZI, S. M.; WETZELS, L. **Handbbook of Portuguese Linguistics**. Wiley-Blackwell, 2016. **crossref** <http://dx.doi.org/10.1002/9781118791844.ch11>

SERRA, C.; CALLOU, D. A interrelação de fenômenos segmentais e prosódicos: confrontando três comunidades. 26 Textos selecionados, **XXVIII Encontro Nacional da Associação Portuguesa de Linguística**, Coimbra, APL, pp. 585-594, 2013.

SILVEIRA, G. **O apagamento da vibrante na fala do sul do Brasil sob a ótica da palavra**. Dissertação de mestrado. Universidade Federal do Rio Grande do Sul, 2010.

VAN OOSTENDORP, M. V. Derived Environment Effects and Consistency of Exponence. In: BLAHO, S.; BYE, P.; KRÄMER, M. **Freedom of Analysis?** Berlin, New York: Mouton de Gruyter, p. 123-148, 2007. [ROA-853]

WETZELS, L. Mid-vowel neutralization in the Brazilian Portuguese verb. **Phonology**, 12 : 281–304, 1995. **crossref** <http://dx.doi.org/10.1017/S0952675700002505>

WOLF, M. **Optimal Interleaving**: serial phonology-morphology interaction in a constraint-based model. Tese de doutorado. University of Massachusetts Amherst, 2008. [ROA- 996]

Artigo recebido em: 31.10.2015

Artigo aprovado em: 03.04.2016

O mapeamento fonético-fonológico das vogais postônicas finais no português brasileiro

The phonetic-phonological mapping of final unstressed vowels in Brazilian Portuguese

Carmen Lúcia Matzenauer*

RESUMO: O funcionamento das vogais átonas é, no dizer de Câmara Jr. (1977), “um dos problemas mais intrincados da fonêmica portuguesa no Brasil”. Em sílaba átona final, a mais débil da palavra, o triângulo de sete vogais tônicas do sistema fonológico da língua fica reduzido às três vogais periféricas /a, i, u/ (CÂMARA JR., 1970, 1977; BISOL, 1981, 2002, 2003). Essas vogais que se mantêm nas sílabas com maior grau de atonicidade no português brasileiro (PB) são as mais frequentes nos inventários fonológicos das línguas (MADDIESON, 1984) e as primeiras a emergir no processo de aquisição da linguagem pelas crianças (RANGEL, 2002). As formas fonéticas que tais vogais assumem no PB tendem a ser centralizadas ([ɪ, ʊ, ɐ]), havendo variação, no sul do país, entre altas e médias (golp[ɪ] ~ golp[e]; camp[ʊ] ~ camp[o]). Diante da possível assimetria ao se considerarem as vogais átonas finais do PB, o mapeamento entre os níveis fonético e fonológico da língua é o foco do presente estudo. Na discussão, retoma-se pesquisa realizada com quatro informantes do sul do Brasil, submetidos a testes de percepção e de produção linguísticas, cujos dados foram interpretados por meio de restrições com o suporte do Modelo BiPhon (BOERSMA, 2007, 2011). Os resultados foram formalizados de modo a representar a gramática responsável pela representação das vogais átonas finais de falantes nativos de PB.

ABSTRACT: The functioning of unstressed vowels is, according to Câmara Jr. (1977), “one of the most intricate problems of Portuguese phonemics in Brazil”. In final unstressed position, the weakest of the word, the seven stressed vowel triangle of the phonological system of the language is reduced to the three peripheral vowels /a, i, u/ (CÂMARA JR., 1970, 1977; BISOL, 1981, 2002, 2003). These vowels that remain in the syllables with higher degree of stress in Brazilian Portuguese (BP) are the most frequent in the phonological inventories of languages (MADDIESON, 1984) and the first to emerge in children’s language acquisition process (RANGEL, 2002). The phonetic forms that such vowels take on in BP tend to be centralized ([ɪ, ʊ, ɐ]), even though some variation may be found in southern Brazil between high and mid ones (golp[ɪ] ~ golp[e]; camp[ʊ] ~ camp[o]). Taking into account the possible asymmetry when BP final unstressed vowels are considered, the mapping between the phonetic and phonological levels of the language is the focus of this study. In the discussion, a previous investigation that had been carried out in southern Brazil is reviewed: four informants were submitted to linguistic perception and production tests, whose data were interpreted by means of constraints with the support of the BiPhon Model (BOERSMA, 2007). Results were formalized to represent the grammar that is responsible for the PB native speakers’ representation of the final unstressed vowels.

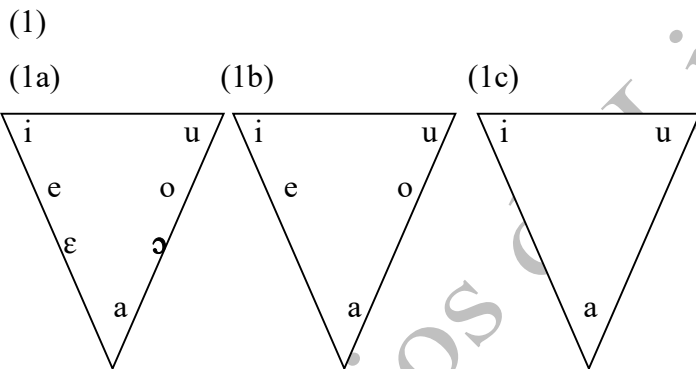
* Professora do Programa de Pós-Graduação em Letras - Universidade Católica de Pelotas (UCPEL). Pesquisadora do CNPq. Email: carmen.matzenauer@gmail.com

PALAVRAS-CHAVE: Vogais átonas finais. Percepção e produção. Níveis fonético e fonológico. Fonologia do português.

KEYWORDS: Final unstressed vowels. Perception and production. Phonetic and phonological levels. Portuguese phonology.

1. Introdução

No dizer de Câmara Jr. (1977), o funcionamento das vogais átonas é “um dos problemas mais intrincados da fonêmica portuguesa no Brasil”. O conjunto de sete vogais que constitui o sistema tônico (/i, e, ε, a, ɔ, o, u/), sendo alvo do processo de neutralização, não consegue sustentar-se em sílabas átonas: no português brasileiro (PB), cinco vogais são fonológicas em posição pretônica (/i, e, a, o, u/) e apenas três, segundo Câmara Jr. (1970, 1977), o são em posição átona final (/i, u, a)¹. Vejam-se as representações em (1).



O subsistema vocálico em (1c), composto por apenas três segmentos vocálicos, segundo Câmara Jr, em sílaba postônica final, é o foco do presente estudo². O objetivo é discutir a constituição do sistema composto pelas vogais /i, u, a/, tecendo-se considerações sobre tipologias de línguas e sobre o processo de aquisição da linguagem, e também o mapeamento entre os níveis fonético e fonológico das vogais átonas finais, trazendo-se dados de produção e de percepção dos segmentos vocálicos por falantes nativos de português do Brasil.

Há interesse particular no estudo desse subsistema vocálico, uma vez que pesquisas sociolinguísticas, como a de Vieira (2002), revelam flutuações no emprego de vogais postônicas finais (*mestre* ~ *mestri*; *livro* ~ *livru*), em algumas variedades do português

¹ Para Câmara Jr., o sistema do PB contém quatro vogais em posição postônica não final: /i, e, a, u/; para Bisol (2003), nessa posição, o sistema ainda se mostra flutuante: com cinco vogais, como na posição pretônica, ou com três vogais, como na posição postônica final.

² O presente artigo integra pesquisa apoiada pelo CNPq – Processo nº 305514/2013-0.

brasileiro, o que é indicativo, segundo a autora e também de acordo com Bisol (2003, 2010), de o processo de neutralização não estar integralizado, ou seja, há ainda variedades do PB que apresentam essa neutralização opcionalmente, até mesmo escassamente. Bisol (2010), ao discutir o sistema vocálico átono final do PB com base em dados provenientes de amostras representativas das três capitais do Sul, refere que o processo de neutralização da átona ainda não chegou à concretude. O tema, portanto, merece ser retomado.

Os dados de percepção e de produção dos segmentos vocálicos que integram o presente estudo e que se mostram relevantes para a discussão do subsistema postônico final são interpretados, sob os enfoques fonético e fonológico, em uma articulação que tem base no Modelo Bidirecional de Fonologia e Fonética (BiPhon), proposto por Boersma (2007, 2011) e Boersma e Hamann (2009), o qual, segundo Boersma, é um modelo de gramática que opera por meio de restrições, nos moldes da Teoria da Otimidade (*Optmality Theory* – OT).

2. A constituição do sistema vocálico átono final do PB – paralelo com tipologias de línguas e com a aquisição da linguagem

A fonologia das vogais átonas finais do PB, ao ser constituída por /i, u, a/, integra segmentos que, no triângulo convencional que representa sistemas vocálicos, ocupam os seus ângulos ou a sua periferia (veja-se (1c)).

Na dimensão fonológica, esses três segmentos vocálicos constituem o conjunto de vogais que contrastam pelos traços que, de acordo com Clements (2001) e Calabrese (2005), poderiam ser considerados os mais robustos na constituição de sistemas vocálicos³. A partir desses autores, propõe-se aqui que esses segmentos vocálicos, apresentados como coocorrências de traços, mostram a configuração em (2).

(2)

/a/ é a combinação dos traços /baixo, dorsal/;

³ Segundo Clements (2001) e Calabrese (2005), é pertinente a proposição desta escala de Robustez de traços constitutivos de segmentos vocálicos:

- a) [Baixo]
- b) [Alto]
- c) [Posterior]
- d) [Arredondado]
- e) [ATR]

Veja-se que o traço [posterior] corresponde, no modelo de Clements (1991) e Clements & Hume (1995), aos traços de ponto [coronal], [dorsal] e [labial].

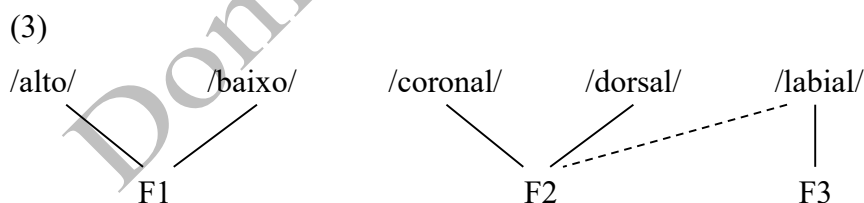
/i/ é a combinação dos traços /alto, coronal/;

/u/ é a combinação dos traços /alto, dorsal, labial/.

Na dimensão fonética, caracterizando-se as vogais por meio de formantes (F1 e F2)⁴, tem-se que a vogal baixa [a] tem o mais alto F1 e as vogais altas [i] e [u] têm o menor F1; quanto ao F2, relevante apenas para a caracterização das duas vogais altas, [i] mostra o mais alto F2, enquanto [u] mostra o menor F2. Essas três são, pois, as “vogais de ponta”, conforme as identifica Maddieson (1984, p.140); são as vogais com os valores mais extremos para altura e posterioridade, que exploram ao máximo o espaço acústico.

O fato de cada uma das três vogais localizar-se em um extremo acústico é fator que facilita a sua percepção, motivando clara distinção entre elas e, como consequência, a sua utilização como unidades da fonologia das línguas. Na realidade, as pistas acústicas de altura e posterioridade, que caracterizam as vogais, são interpretadas como unidades fonológicas – como traços – e, como consequência, o *continuum* fonético é mapeado em unidades discretas, fonológicas.

Atentando-se para o *continuum* das dimensões fonológico-fonéticas, em (3) representa-se a interface que integra tais dimensões, trazendo-se os traços⁵ como unidade fonológica e os formantes como unidades fonéticas⁶.



⁴ O parâmetro de altura é representado acusticamente pelo Formante 1 - F1 e o de anterioridade/posterioridade é representado acusticamente pelo Formante 2 - F2.

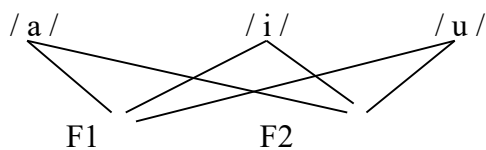
⁵ Apresentam-se, em (3), os traços entre barras, a exemplo de Boersma & Escudero (2003), a fim de formalizar sua natureza fonológica.

⁶ A representação do traço [labial] expressa a relação implicacional com F2, uma vez que se segue Guy (2009), para quem o português “não exibe processos afetando arredondamento; todas as vogais, na língua, mantêm os valores não marcados para arredondamento” (slide n°7). Ainda o autor destaca a relação que há entre o arredondamento e F2: “o efeito acústico de arredondamento sobre F2 cria a percepção de posteriorização (a protrusão dos lábios produz alongamento do tubo ressonante bucal, igual à retração da língua); esse efeito explica a associação universal entre vogais posteriores e arredondamento: aumenta a diferenciação perceptual entre vogais anteriores e posteriores” (slide n°22).

A integração fonética/fonologia mostrada em (3) está apresentada de forma restrita, vinculando apenas duas pistas acústicas a traços, que são responsáveis por contrastes fonológicos. Mesmo assim, já é possível verificar-se que essa integração não é tão simples, já que uma única propriedade fonética ou pista acústica está vinculada a mais de um traço, ou seja, a mais de um contraste fonológico.

A complexidade desse mapeamento talvez se evidencie mais claramente na representação em (4), quando em lugar dos traços se colocam os segmentos vocálicos.

(4)



Pela formalização em (4), é possível visualizar-se que não é qualquer valor de F1 que pode ser interpretado como /a/, por exemplo, na fonologia, uma vez que F1 é pista acústica também vinculada às outras vogais, assim como os outros formantes.

No entanto, ao tratar-se do conjunto das vogais /i, u, a/, essa complexidade de mapeamento vê-se amenizada, porque, havendo no plano fonético a exploração máxima do espaço acústico, no plano fonológico há o estabelecimento de contraste perceptivelmente saliente e, como consequência, de categorização facilitada na gramática.

As características fonéticas e fonológicas das vogais /i, u, a/, como vogais periféricas no espaço acústico e como vogais que contrastam no plano fonológico, têm repercussão na constituição dos inventários fonológicos das línguas do mundo e também na construção gradual dos inventários de segmentos durante o processo de aquisição da linguagem pelas crianças.

2.1 As vogais periféricas em tipologias de línguas

O conjunto das vogais periféricas /i, u, a/ é o de presença mais frequente nos inventários fonológicos⁷ – essa afirmação é de Maddieson (1984, p.136-7) em estudo sobre tipologias de línguas que avalia 317 sistemas linguísticos. Para Maddieson (1984) e Lindblom (1986), essas três vogais constituem o menor inventário completo encontrado com alguma regularidade em

⁷ Segundo Maddieson (1984), das 317 línguas estudadas, apenas duas não apresentavam a vogal /a/ em seu inventário: Cheremis (p.274) e Tatalog (p.340), sendo que, das três vogais, /u/ é a menos frequente.

línguas do mundo. Esse fato leva Maddieson a pressupor que a organização vocálica preferida tende a apresentar dispersão larga e balanceada no espaço fonético. Nesse mesmo sentido, Lindblom (1986) refere que o inventário vocálico formado por /i, u, a/ apresenta dispersão máxima e, portanto, sua manifestação fonética mostra ambiguidade acústica mínima.

A hipótese da dispersão parece confirmar-se na observação de sistemas vocálicos lacunares – as lacunas afetam predominantemente as vogais médias (MADDIESON, 1984; GUY, 2009; MATZENAUER, 2009).

Veja-se que a dispersão, ao considerar o uso eficiente do espaço acústico disponível, é relevante aspecto fonético que tem repercussão direta no funcionamento do nível fonológico das línguas, já que aumenta a distintividade entre os sons e, como efeito, facilita a sua identificação como categoria da gramática.

É nessa linha de entendimento que Crosswhite (2001) vê o fortalecimento do contraste como uma das explicações para a redução vocálica, decorrente de neutralização: essa forma de redução, para a autora, levará os sistemas a eliminarem vogais não periféricas, especialmente vogais médias. O resultado dessa redução, portanto, é que vogais de sílabas átonas fiquem restritas ao subsistema /i, u, a/.

Para Crosswhite (2001), as vogais periféricas /i, u, a/ são especiais por uma conjugação de fatores: pela dispersão (de acordo com MADDIESON, 1984; LINDBOM, 1986), e também pelas características quânticas (STEVENS, 1989) e pela focalização (STEVENS, 1989; SCHWARTZ et al., 1997). Em resumo, a dispersão, que diz respeito ao uso eficiente do espaço acústico disponível, está relacionada à facilidade de percepção; o efeito quântico, que ocorre quando a uma determinada alteração na articulação não corresponde uma grande alteração acústica⁸, está relacionado ao vínculo entre alteração articulatória e alteração acústica⁹ e a focalização, que se refere a padrões estáveis em tarefas de discriminação de sons, está relacionada à saliência acústica e perceptual, reunindo ambas dispersão e focalização¹⁰. Essa convergência de propriedades torna especial o sistema vocálico /i, u, a/.

⁸ Sons da fala mostram efeito quântico quando sua qualidade acústica é mais ou menos consistente em uma larga extensão de articulações. Como consequência, sons não-quânticos mostram grandes alterações na qualidade acústica quando há alterações articulatórias.

⁹ Para Stevens (1989), as línguas preferem sons que mostrem efeitos quânticos, porque se mantêm consistentes em uma extensão maior de articulações – as vogais /i, u, a/ mostram esse tipo de efeito.

¹⁰ Pela híbrida Teoria da Dispersão-Focalização, Schwartz et al. (1997) tratam a focalização como parâmetro que determina padrões mais estáveis em tarefas de discriminação, relacionada à maximização da distância entre vogais e intra-vogais.

Destaca-se que, por terem as qualidades de estabilidade articulatória e acústica, por serem de máxima distinção acústica e de maior saliência perceptual, as vogais /i, u, a/ são menos sensíveis a variação e se mostram como as vogais preferidas não apenas por si mesmas, mas também como um conjunto. Não é, pois, surpreendente o fato de integrarem predominantemente os inventários de diferentes línguas, assim como não causa estranheza ser o subsistema que se mantém em contextos mais débeis, como posições átonas. Por fim, o status dessas vogais permite entender-se por que são as primeiras a emergir no processo de aquisição da linguagem pelas crianças.

2.2 As vogais periféricas na aquisição da linguagem

No gradual processo de constituição do inventário fonológico pelas crianças, as vogais periféricas /i, u, a/ são as primeiras a emergir. Essa constatação se faz evidente nos estudos sobre a aquisição do sistema tônico do português, em que se encontram sete segmentos e em que essas três vogais ocupam os espaços fonético-fonológicos das vogais médias antes da incorporação das vogais /e, o, ε, ɔ/ ao sistema da criança.

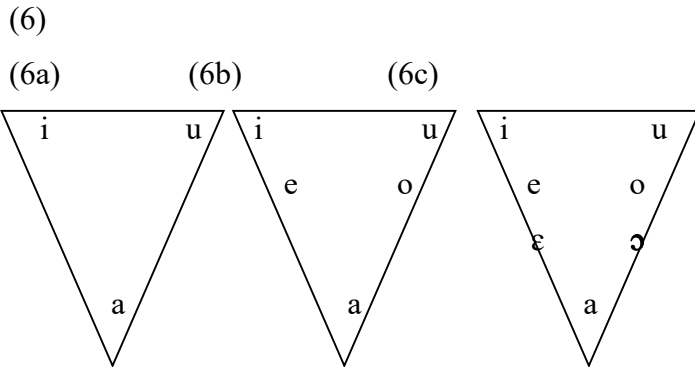
Segundo as pesquisas sobre o processo de aquisição do sistema vocálico do PB por crianças brasileiras (RANGEL, 2002; MATZENAUER & MIRANDA, 2007, 2009; MATZENAUER, 2012), há estágios até a integralização do sistema, em se considerando as sete vogais tônicas da língua. Sintetiza-se esse processo de aquisição em três estágios, seguindo-se os autores acima referidos, conforme se mostra em (5).

(5)

/a, i, u/ > /e, o/ > /ε, ɔ/

Esses três estágios desenvolvimentais implicam um diferente e gradual recorte do espaço vocálico: primeiramente, como em (6a), subsequentemente, conforme (6b) e, por fim, como em (6c)¹¹.

¹¹ Observe-se que os estágios de aquisição do sistema vocálico tônico por crianças brasileiras, representados em (6), apresentam a distribuição inversa dos subsistemas vocálicos pertinentes no PB, representados em (1).



A essência (e a dificuldade) do processo de aquisição da fonologia está em recortar o contínuo fonético em categorias da gramática. O ponto de partida para a construção do sistema fonológico está no *input* linguístico que a criança recebe, ou seja, no *continuum* fonético que a criança ouve: as pistas acústicas têm de ser convertidas em categorias do sistema linguístico que é alvo da aquisição.

O interesse que tem esse processo vale aqui uma pequena digressão, a partir de uma pergunta de especial relevância: como se poderia entender o processo gradual de aquisição do sistema vocálico pelas crianças?

A proposta que se tem é de que o bebê, antes de chegar a categorias fonológicas, tem de estabelecer categorias fonéticas. Nesse posicionamento, concorda-se com Boersma et al. (2003) no sentido de que as crianças detêm a capacidade inata de “calcular as distribuições estatísticas da informação auditiva fonética em seu input linguístico”, sendo que o conhecimento dessa distribuição leva o bebê à criação de categorias fonéticas aos seis/oito meses de idade¹². A formação de categorias fonéticas, com base em evidências distribucionais no input linguístico, é estabelecida de acordo com o que os autores denominam “efeito perceptual magnético”: as experiências auditivas começam a aglutinar-se em torno de um *protótipo*, o qual está na base da constituição de uma categoria fonética. De acordo com essa proposta, portanto, interpreta-se que antes mesmo do nível fonológico já há recortes no *continuum* fonético, pela formação de categorias decorrentes do chamado “efeito perceptual magnético”.

¹² Talvez as categorias fonéticas tenham formação anterior a esse período, segundo pesquisa relatada por Wanrooij, Boersma & Zuijlen (2014).

Indo-se além, entende-se que, diante do *input* linguístico, os bebês são sensíveis ao “efeito perceptual magnético” em frequências naturais de ressonância, as quais vão caracterizar os formantes, dando origem a categorias que poderão ser identificadas com o que, na percepção e na produção linguísticas, se reconhece como o Formante 1 (F1) e o Formante 2 (F2), oferecendo as bases para categorias fonéticas presentes na língua.

Estendendo-se esse entendimento, é possível interpretar-se também que as categorias fonéticas são o fundamento da formação de categorias fonológicas, expressas como traços. A percepção, portanto, teria início em uma aprendizagem dirigida pela audição, por meio de uma organização inicial do espaço perceptual do bebê em categorias fonéticas, em razão de um efeito perceptual magnético. Somente em momento subsequente, quando já houver léxico, emergirão representações abstratas, fonológicas; então, a aprendizagem será também dirigida pelo léxico.

É possível, como situação ilustrativa, entender-se que, no momento da aprendizagem dirigida pela audição, poderia ser constituída uma categoria fonética em torno, por exemplo, de F1, já que essa é uma das pistas auditivas consideradas na formação de categorias fonéticas – poderia ser composta, nesse caso, uma categoria da percepção de F1 [900Hz]¹³; essa categoria fonética seria a base para a ativação, subsequentemente, do traço [baixo] na aquisição fonológica da vogal /a/ – assim, os níveis das representações fonética e fonológica mostram-se em interação. Com esse encaminhamento, seriam constituídas, de início, as categorias fonéticas e, então, se estabeleceria sua relação com categorias fonológicas. Interpreta-se aqui que essa serialidade na relação entre os níveis Fonético e Fonológico ocorre na etapa mais inicial da aquisição fonológica, quando ainda não há léxico, restringindo-se a relação entre as unidades a um caráter distribucional.

Com essa linha de argumentação, deve entender-se que, em havendo léxico, a aquisição será por ele dirigida, sendo que então se manifesta (e tem de manifestar-se) a capacidade de extrair unidades abstratas, constitutivas da gramática da língua. E essa capacidade, no processo de integrar, à fonologia, traços e segmentos, também poderá aproveitar os benefícios da capacidade de cálculo de distribuição que as crianças detêm. Como consequência, categorias

¹³ Esse valor em Hz foi proposto a partir de resultado de pesquisa sobre a produção vogais do PB por pelotenses: Azevedo (2014) identificou, para falantes do sexo feminino, o valor médio de F1 para [a] de 1029Hz, e, para falantes do sexo masculino, o valor médio de F1 para [a] 857Hz. Nesse sentido, assume-se, com Boersma & Escudero (2003), que a percepção apresenta o máximo de verossimilhança com o resultado da produção.

fonológicas são estabelecidas gradativamente, cada uma das quais poderá incluir um fonema¹⁴, como protótipo da categoria, e formas alofônicas. Estando em atividade o nível fonológico, este também condicionará o nível fonético e, como consequência, a percepção será fonológica.

O fato que merece destaque é que, no curso da aquisição, a integração, à fonologia, de traços e de segmentos pode ser vista como resultante da organização de espaços acústicos determinados por categorias de percepção, com base em valores de F1 e F2 (em se tratando de segmentos vocálicos), a partir dos quais vão sendo delineados espaços fonológicos, pelo reconhecimento de um conjunto de conexões entre elementos fonológicos (ex.: fonemas vocálicos) e pistas auditivas (valores de F1 e F2)¹⁵.

Ao considerar-se o primeiro estágio de aquisição do sistema vocálico do PB mostrado em (6a), as conexões que integram pistas fonéticas a segmentos fonológicos poderiam ser representadas da mesma forma como se mostrou em (4), quando se salientou a complexidade do mapeamento do nível fonético no nível fonológico.

Essa aquisição pressupõe, sem dúvida, o mapeamento das pistas acústicas em traços, os quais estão na essência dos contrastes estabelecidos na gramática da língua¹⁶.

A aquisição gradual do sistema vocálico, formalizada em três estágios para crianças falantes de PB conforme é mostrada em (6), suscita uma indagação: haverá estabilidade na organização do espaço acústico das vogais? O que se questiona é se a distância de altura, representada acusticamente por F1, e a distância de anterioridade/posterioridade, representada acusticamente por F2, entre as três vogais em (6a) é a mesma que se mantém nos outros dois estágios, ou se, de início, essa distância é menor e vai se alterando à medida que outros segmentos são incorporados ao sistema, em um exercício de acomodação do espaço estrutural. Essa é pergunta ainda sem resposta definitiva que merece ser objeto de investigação subsequente.

Embora o presente estudo não busque a resposta a essa questão, considera que duas hipóteses são pertinentes ao tratar-se de uma investigação dessa natureza: a primeira é a de que, desde o início, as formas fonéticas que representam as vogais /i, u, a/ ocupam, no espaço

¹⁴ Um determinado fonema deve funcionar, no início da constituição de uma classe de segmentos, como protótipo da classe – é o que ocorre, por exemplo, com o segmento /l/, que se defende ser o protótipo da classe das consoantes líquidas (Matzenauer-Hernandorena, 1990).

¹⁵ O caminho aqui traçado para a categorização de traços e de segmentos na fonologia da criança não é incompatível com o entendimento dos traços como primitivos fonológicos, presentes na GU, se houver o pressuposto de que os traços, como primitivos, precisam ser ativados para que integrem a gramática do falante e de que essa ativação de traços ocorre a partir do input fonético e da formação de categorias fonéticas.

¹⁶ A formalização desse mapeamento já foi feita em (3).

acústico, posições que vão ser mantidas, independentemente da integração de novos segmentos ao sistema – o fundamento para esta hipótese está no fato de que a motivação para a emergência precoce, na aquisição, desses três segmentos, está na distância fonética que os separa, facilitando sua percepção e, ao mesmo tempo, não implicando dificuldade à sua produção. Essa motivação encontra paralelo nas tipologias de línguas, já que, segundo Maddieson (1984), são esses os três segmentos vocálicos mais frequentes nas línguas e, segundo Calabrese (1995, 2005), são essas as vogais consideradas não marcadas. A confirmação dessa hipótese atestaria a manutenção da estabilidade das três vogais periféricas do sistema fonológico do PB, independentemente da integração de novos segmentos à fonologia da criança e, portanto, do aumento do número de segmentos no inventário vocálico.

A segunda hipótese é a de que as representações fonéticas das vogais /i, u, a/ mostram diferenças na distância que mantêm no espaço acústico, dependendo do número e da qualidade dos segmentos vocálicos que vão sendo incorporados ao sistema – a base para a proposição dessa hipótese está na possibilidade de um movimento análogo ao que ocorre no processo de aquisição de uma língua estrangeira (LE) com sistema vocálico menor do que aquele da língua materna do falante. Foi o que constatou Santos (2014), em estudo sobre o processo de aquisição do espanhol como LE, cujo sistema contém cinco vogais (com duas vogais médias: /e, o/), por falantes nativos de PB, cujo sistema conta com sete vogais, incluindo quatro vogais médias: /e, o, ε, ɔ/. Na aquisição do espanhol, os falantes do PB, em lugar de apenas eliminarem as vogais médias baixas [ε, ɔ] do espaço acústico, mantendo [e, o] com os mesmos valores de F1 e de F2, reestruturaram seu sistema, alterando a posição das médias altas no espaço vocálico e, conseqüentemente, redistribuindo os segmentos e a relação fonética estabelecida entre eles.

A confirmação dessa hipótese atestaria o dinamismo da organização estrutural do sistema vocálico durante o processo de aquisição da fonologia da língua, com o estabelecimento de diferentes relações e distâncias acústicas entre os segmentos, à medida que novos processos de fonologização vão ocorrendo e que o inventário segmental vai crescendo.

Um estudo de caso realizado com os dados linguísticos de uma menina falante de PB do sul do Brasil, com três anos de idade (AZEVEDO et al., 2014), revelou, no nível fonético, comportamento que pode ser considerado análogo ao do nível fonológico, no sentido da precocidade com que as vogais periféricas ocupam espaço acústico estável, mostrando valores de F1 muito próximos aos de adultos da mesma região do país, e no sentido da tardia organização do espaço acústico ocupado pelas vogais médias, cujos valores de F1 e de F2 ainda

mostram assimetrias em relação aos dados de adultos. Observe-se a plotagem das vogais, com dados da menina e dos adultos, mostrada na Figura 1. A figura traz, em vermelho, a representação das vogais produzidas por adultos e, em azul, a representação das vogais produzidas pela criança.

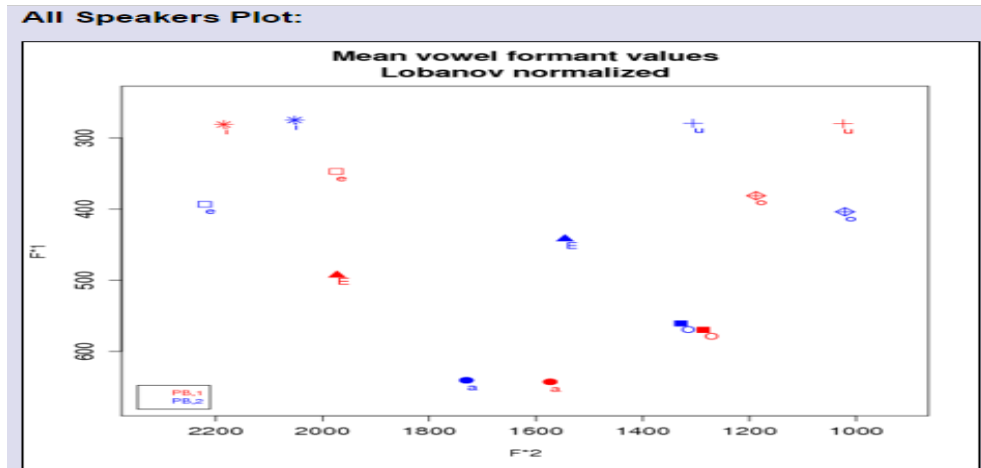


FIGURA 1. Plotagem das vogais de falantes da região sul do Brasil (em azul: vogais da menina; em vermelho, vogais dos adultos).

Fonte: Azevedo et al. (2014).

A disposição das vogais na Figura 1 leva a observar-se que as vogais periféricas [a, i, u], produzidas pela menina, apresentam altura (F1)¹⁷ equivalente às das vogais dos adultos, mostrando alguma diferença (posição mais centralizada) quanto a posterioridade/anterioridade (F2). Esse comportamento das vogais periféricas leva à confirmação da primeira hipótese acima referida: as formas fonéticas das vogais são representadas, no espaço acústico, por posições que vão ser mantidas, independentemente da integração de novos segmentos ao sistema; esse entendimento parece valer para as formas fonéticas referentes às vogais periféricas [i, u, a], embora o F2 ainda não se identifique plenamente com o dos adultos.

Diferentemente do que ocorre com as vogais periféricas, as vogais médias da menina mostram divisão do espaço acústico ainda assimétrico e distante da organização do espaço das vogais dos adultos, ao se considerarem valores de F1 e de F2, especialmente em se tratando das vogais coronais. Dentre as médias, a exceção estaria na vogal média baixa [ɔ]. Essa assimetria entre o espaço acústico das vogais médias da menina e o das vogais dos adultos poderia oferecer subsídios ao fundamento da segunda hipótese acima referida, evidenciando o dinamismo da

¹⁷ Não se apresentam com detalhes os valores de F1 e de F2 por esse não ser o tópico central da discussão proposta no artigo.

organização estrutural do sistema vocálico durante o desenvolvimento linguístico, com diferentes acomodações do espaço acústico até (e, inclusive, um período após) a estabilização dos segmentos em sua forma fonológica.

Validando-se esses fatos em estudos com maior número de informantes, estaria confirmada mais uma diferença importante entre vogais periféricas e vogais médias, aqui sustentada pelo processo de aquisição da linguagem pelas crianças.

Merece registrar-se que, em uma análise de oitava, tais distâncias fonéticas passam despercebidas; capta-se apenas o fato de o inventário fonológico da menina já integrar as sete vogais da língua.

Esse processo gradual de aquisição, tanto no plano fonológico (veja-se (5)), como no plano fonético (veja-se Figura 1), pressupõe o estabelecimento de etapas na medida em que ocorre a gradual organização acústica e a sucessiva organização, na gramática, de segmentos vocálicos e de traços.

A interação entre as dimensões fonética e fonológica no processo de aquisição da linguagem pode ser visualizada no Modelo Bidirecional de Processamento em Língua Materna – BiPhon, proposto por Boersma (2007, 2011); Boerma & Hamann (2009).

2.2.1 O Modelo BiPhon como suporte para a análise dos dados

O Modelo Bidirecional de Processamento em Língua Materna – BiPhon, proposto por Boersma (2007, 2011); Boerma & Hamann (2009), apresentado em uma versão simplificada na Figura 2, caracteriza-se por ser bidirecional, em virtude de incluir a compreensão e a produção da língua: mostra à esquerda a tarefa do ouvinte (Módulo da Compreensão: movimento ascendente na figura, a partir da [forma fonética]) e, à direita, a tarefa do falante (Módulo da Produção: movimento descendente na figura, a partir da |forma subjacente|).

Em ambas as direções, o modelo tem o funcionamento determinado pela interação de restrições da Teoria da Otimidade, o que o faz, também, um modelo de gramática. O modelo representa, então, o processamento (formalizado, na Figura 2, por meio dos mecanismos ligados pelas flechas) e também a gramática (formalizado, na Figura 2, por meio das restrições).

Também tem o mérito de explicitar a interface fonética/fonologia no processo de compreensão e de produção linguísticas: por contar com três níveis de representação (dois fonológicos e um fonético: |forma subjacente|, /forma fonológica de superfície/ e [forma

fonética]), mostra a conexão entre eles e, conseqüentemente, a conexão entre o *continuum* da dimensão fonética e o discreto da dimensão fonológica.

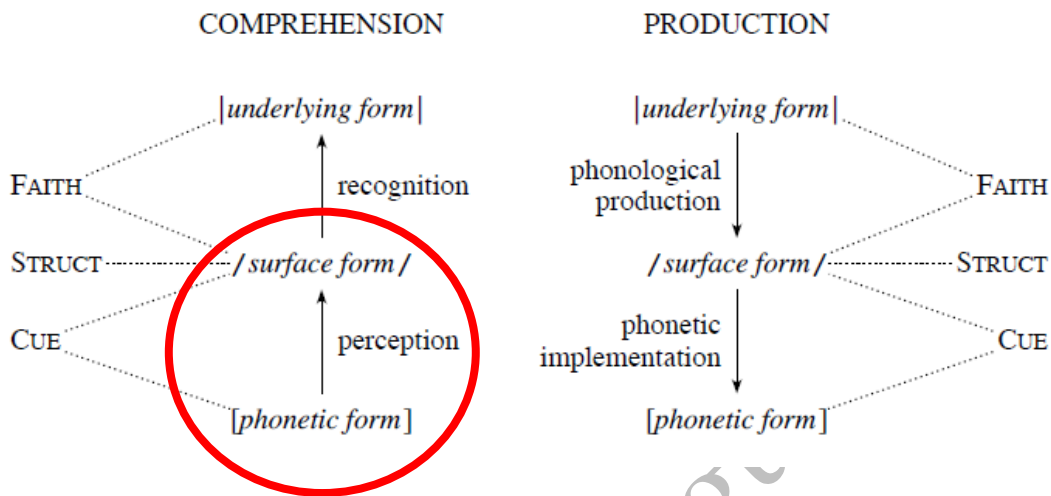


Figura 2. Modelo de processamento de L1 e de gramática (*BiPhon Model*).
Fonte: Boersma & Hamann (2009)¹⁸.

No Modelo BiPhon, a transformação de pistas acústicas em categorias da gramática é movimento que, no processo de aquisição da linguagem, tem a ocorrência formalizada no Módulo da Compreensão, na comunicação entre o nível da [Forma Fonética] e o nível da /Forma Fonológica de Superfície/ (no espaço circundado na Figura 2), sendo que essa comunicação é representada na gramática pela interação entre restrições de pista (que interpretam/avaliam pistas acústicas) e restrições de estrutura (que interpretam/avaliam estruturas das línguas em termos de traços, segmentos, sílabas e pés).

Retomando-se a discussão sobre a aquisição de segmentos vocálicos, nesse jogo interativo, tem-se que as categorias fonéticas constituídas em torno das pistas acústicas que definem as vogais [i, u, a] são, portanto, as primeiras a ser fonologicamente categorizadas em traços e nos segmentos fonológicos /i, u, a/ – tais pistas acústicas são oferecidas às crianças no *input* linguístico que recebem.

Assim, informações advindas tanto da organização de sistemas vocálicos de diferentes tipologias de línguas, como da construção do inventário de vogais da língua pelas crianças no

¹⁸ Traduzem-se aqui os elementos fundamentais do BiPhon para o presente estudo, que são os níveis de representação: [Forma Fonética] ([Phonetic Form]), /Forma Fonológica de Superfície/ (/Surface Form/), |Forma Subjacente| (|Underlying Form|), bem como os três tipos de restrições que articulam esses níveis: restrições de pista (Cue), restrições de estrutura (Structure) e restrições de fidelidade (Faith).

processo de aquisição da linguagem atribuem *status* diferenciado e especial ao sistema de vogais tripartido em /i, u, a/.

3. O sistema vocálico átono final do PB – representações fonológica e fonética

O *status* especial que as vogais /i, u, a/ parecem apresentar deve estar na motivação de a gramática do português as ter escolhido para a constituição do sistema que é distintivo na posição mais débil da palavra: a posição átona final.

Conforme já foi salientado, para Câmara Jr. (1970, 1977), o sistema átono final do português é composto pelas vogais /i, u, a/. Também para Bisol (2003, 2010), esse é o inventário de vogais átonas finais na fonologia da língua, embora a autora reconheça que pesquisas sociolinguísticas apontam, em algumas variedades do PB do sul do país, também a presença das vogais médias altas, encaminhando-se para o sistema de três vogais. É possível, portanto, pensar-se que os falantes podem categorizar cinco vogais no sistema átona final¹⁹. Considerando-se tal fato, reporta-se, aqui, estudo sobre vogais átonas finais no sul do Brasil.

3.1 Relato de um estudo sobre vogais átonas finais no sul do Brasil²⁰

Para a discussão do comportamento das vogais átonas finais no português e do mapeamento fonético-fonológico que seu funcionamento implica, retoma-se pesquisa proposta por Matzenauer, Neuschrack, Carniato e Azevedo (2015), realizada com quatro brasileiros do sul do país, dois homens e duas mulheres, monolíngues, falantes nativos de PB, com idade entre 25 e 37 anos, nascidos e estabelecidos em Pelotas - RS, com escolaridade universitária. Com foco no sistema vocálico átono final, no estudo os informantes foram submetidos a testes de percepção e de produção linguísticas, tendo sido os resultados analisados por meio de restrições, com o suporte do Modelo BiPhon (BOERSMA, 2007, 2011), nos moldes da Teoria da Otimidade Estocástica. As restrições foram formalizadas com a utilização do Algoritmo de Aprendizagem Gradual – GLA (BOERSMA; HAYES, 2001).

3.1.1 Metodologia do estudo sobre vogais átonas finais

¹⁹ A categorização fonológica de cinco vogais no sistema átono final poderia também ser influenciada pelo código escrito.

²⁰ Esse estudo, de autoria de Matzenauer, Neuschrack, Carniato e Azevedo, foi recentemente publicado na Revista da ABRALIN (2015).

O *corpus* de produção foi obtido pela leitura de frases de um instrumento²¹, contendo logatomas com as vogais-alvo em sílaba átona final. Utilizaram-se 10 logatomas para cada vogal (ortograficamente apresentadas como a, e, i, o, u) com cinco estruturas fonológicas (pVpV / tVkV / kVkV / fVfV / sVsV), o que somou 100 *tokens* para cada sujeito; todos os logatomas eram, portanto, dissílabos e paroxítonos. Os dados, após a coleta em cabine acústica, foram segmentados por meio do software Praat (BOERSMA; WEENINK, 2012) e organizados em arquivo de áudio, tendo sido as vogais postônicas finais submetidas a uma análise acústica no Praat e descritas com foco nos Formantes 1 e 2, suficientes para a definição das vogais do PB.

O *corpus* de percepção das vogais estudadas foi obtido por meio de um teste de identificação, proposto a partir do instrumento anteriormente aplicado para avaliar a produção das vogais. Os estímulos para a percepção continham as vogais átonas finais [ɪ, ʊ, ɐ] – que foram as formas fonéticas presentes no teste de produção realizado –, sendo que os informantes tinham a possibilidade de identificar o que ouviam como uma das cinco vogais ortograficamente apresentadas como a, e, i, o, u. Os arquivos de áudio (estímulos) foram editados e normalizados no *software* Praat (BOERSMA; WEENINK, 2012) e apresentados aos informantes no *software* TP (RAUBER et al, 2012), sem contexto frasal²².

A análise estatística dos resultados, realizada com o uso do *software* IBM SPSS versão 20.0, teve a finalidade de verificar a acurácia da percepção e da produção das vogais em posição postônica final. Os resultados foram analisados de forma a lidar em conjunto com os dados relativos à percepção e à produção das vogais estudadas. O suporte teórico para a análise foi o Modelo Bidirecional de Processamento e de Gramática – BiPhon, de Boersma (2007, 2011), pela interação entre diferentes tipos de restrições: *restrições de pista*, *restrições de estrutura* e *restrições de fidelidade*, seguindo os pressupostos da Teoria da Otimidade Estocástica.

3.1.2 Resultados do estudo sobre vogais átonas finais

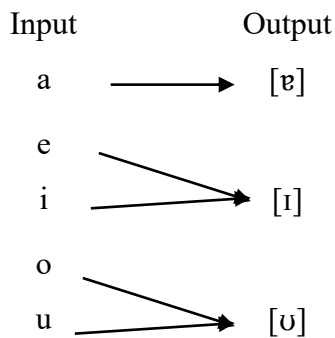
3.1.2.1 Resultados dos dados de produção

²¹ O instrumento para o presente estudo foi criado a partir de uma proposição de Rauber (2008), referente ao estudo das vogais do português em posição tônica.

²² Maiores detalhes sobre a metodologia do estudo podem ser obtidos em Matzenauer, Neuschrack, Carniato e Azevedo (2015).

A leitura de logatomas com as vogais átonas finais ortograficamente apresentadas como a, e, i, o, u por todos os informantes resultou em três outputs – [ɪ, ʊ, ɐ] –, com a correspondência mostrada em (7).

(7)



O resultado expresso em (7) leva a entender-se que, para os falantes do sul do Brasil da região focalizada pelo estudo, mesmo que haja a representação fonológica de cinco vogais átonas finais /a, e, i, o, u/, apenas três são as formas fonéticas produzidas: [ɪ, ʊ, ɐ].

Salienta-se que os dados de produção das vogais átonas finais foram submetidos a uma análise acústica (F1 e F2), sendo que os testes estatísticos confirmaram a relação mostrada em (7). Na Figura 3, mostra-se a plotagem das três vogais foneticamente realizadas nessa posição.

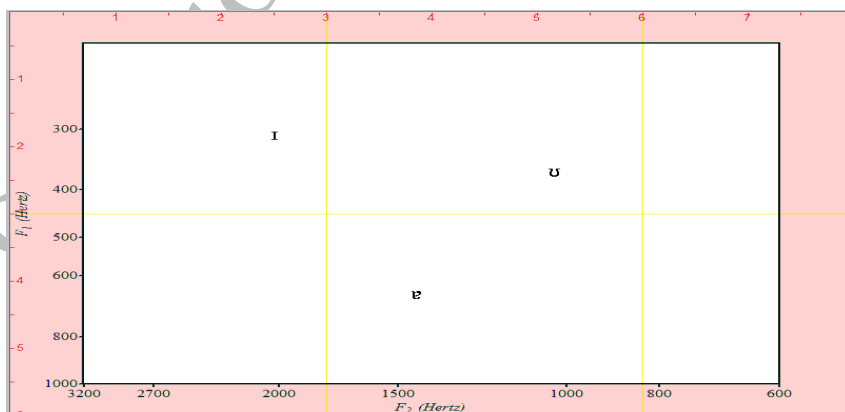


Figura 3. Plotagem das vogais em posição átona final.
Fonte: Matzenauer, Neuschrack, Carniato e Azevedo (2015).

A manifestação fonética das vogais periféricas, na posição átona final, é diferente daquela que as mesmas vogais mostram em posição tônica: mais centralizada, essa

manifestação forma um conjunto estruturalmente mais compacto em posição átona, se comparada com a sua forma em posição tônica (veja-se Figura 4). A comparação entre as Figuras 3 e 4 evidencia que a forma fonética das vogais altas átonas é mais baixa do a das vogais altas tônicas²³. Vê-se, assim, que, foneticamente, a forma das vogais altas átonas se aproxima da forma das vogais médias altas quando ocupam a posição tônica.

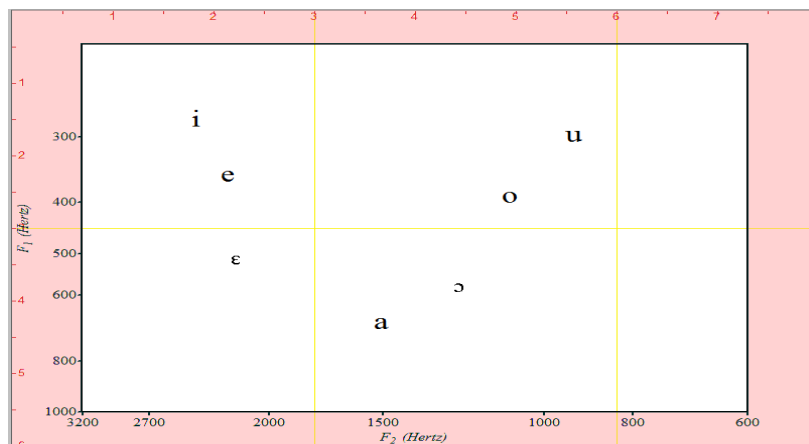


Figura 4. Plotagem das vogais em posição tônica.
Fonte: Matzenauer, Neuschrack, Carniato e Azevedo (2015).

3.1.2.2 Resultados dos dados de percepção

No teste de percepção, os informantes ouviram estímulos, produzidos por voz feminina e por voz masculina, pela leitura dos logatomas com três vogais átonas finais: [ɪ, ʊ, ɐ], conforme já foi referido. Esses estímulos foram identificados/categorizados pelos informantes conforme se mostra em (8).

(8)

Estímulo	Identificação/ Categorização	Percentual
[ɐ]	→ /a/	100%
[ɪ]	→ /e/	1%
	→ /i/	99%
[ʊ]	→ /o/	26%
	→ /u/	74%

²³ As vogais em posição tônica aqui analisadas também foram obtidas pela leitura de logatomas.

Pelos resultados em (8), observa-se que houve a inequívoca categorização da forma fonética [ɐ] como /a/, a quase total categorização da forma fonética [ɪ] como /i/ e a variável categorização da forma fonética [ʊ] ora como /o/, ora como /u/, sendo esta a predominante.

Essa variabilidade na percepção das formas fonéticas apresentadas aos informantes em posição átona final pode ser interpretada como indicativa de que, no sul do país, o PB mostra o processo de neutralização ainda em andamento, com a possível categorização de cinco vogais na posição átona final. Pelos percentuais observados, no entanto, os dados evidenciam o encaminhamento do processo, no PB da Região Sul, para o conjunto de três vogais átonas finais /i, u, a/, sendo a alta posterior a última a estabilizar-se nesse sistema.

3.1.3 Formalização dos dados de produção e de percepção das vogais átonas finais

O sistema de três vogais (/i, u, a/) na posição átona final é visto como resultado do processo de neutralização, ou seja, como decorrência da anulação de traços que estabelecem contraste entre segmentos da fonologia da língua. A formalização desse fenômeno considerou apenas os níveis mais baixos do modelo BiPhon, tanto para a percepção, como para a produção (veja-se Figura 2). Nesse Modelo de Processamento e de Gramática, a relação entre os níveis da [Forma Fonética] e a /Forma Fonológica de Superfície/ é estabelecida por meio de restrições de pista e restrições de estrutura.

No estudo de Matzenauer, Neuschrack, Carniato e Azevedo (2015), pela interpretação de que a preferência pelas vogais periféricas /i, u, a/ em posição átona final decorre de seu maior teor de contraste, ou seja, pela distinção claramente estabelecida a partir da mínima ambiguidade acústica que as formas fonéticas dessas vogais apresentam, foram utilizadas as restrições de estrutura apresentadas em (9a) ((a) e (b)) e as restrições de pista mostradas em (9b) (de (c) a (m)).

As duas restrições de estrutura constituem-se em restrições de licenciamento e de aumento de contraste; foram inspiradas em Crosswhite (2001) e em Steriade (1994).

(9a) Restrições de Estrutura:

(a) *Licence Noncorner²⁴

²⁴ A restrição *Licence Noncorner, inspirada em Crosswhite (2001) e em Steriade (1994), diz respeito ao não licenciamento de traços em posições em que pistas acústicas estejam ausentes ou diminuídas. No tratamento dado

(b) *[+tense]

As restrições de pista já apareceram em trabalhos sobre vogais em línguas diferentes do português, como Boersma (1998, 2000), Escudero & Boersma (2003) e Pater (2004). O asterisco que apresenta a restrição de pista não implica proibição, tal como acontece com as restrições de estrutura, mas significa uma negação, o que permite que possam ser lidas e utilizadas na produção e na percepção ao mesmo tempo. A leitura da restrição, para a produção e para a percepção, respectivamente, obedece à seguinte estrutura: *não produza /x/ fonológico como [y] fonético e, ao mesmo tempo, não perceba [y] fonético como /x/ fonológico.*

(9b) Restrições de Pista:

- (c) */+baixo, dorsal - a/ [F1 - 648hz / F2 - 1432hz - ə],
- (d) */+baixo, dorsal - a/ [F1 - 673hz / F2 - 1505hz - a],
- (e) */-alto, coronal - e/ [F1 - 320hz / F2 - 2018hz - i],
- (f) */-alto, coronal - e/ [F1 - 354hz / F2 - 2219hz - e],
- (g) */+alto, coronal - i/ [F1 - 281hz / F2 - 2400hz - i],
- (h) */+alto, coronal - i/ [F1 - 306hz / F2 - 2020hz - i],
- (i) */-baixo, dorsal - o/ [F1 - 360hz / F2 - 1036hz - o],
- (j) */-baixo, dorsal - o/ [F1 - 392hz / F2 - 1090hz - o],
- (k) */-baixo, dorsal - o/ [F1 - 648hz / F2 - 1432hz - ə],
- (l) */+alto, dorsal - u/ [F1 - 300hz / F2 - 928hz - u],
- (m) */+alto, dorsal - u/ [F1 - 363hz / F2 - 1027hz - u].

Em virtude de as restrições de pista representarem a interface do *continuum* para o discreto, todas as possibilidades de resultados fonéticos obtidos no levantamento de dados do estudo precisavam estar contempladas nas restrições, o que justifica a presença de cada uma das restrições de pista listadas acima²⁵.

pelo presente estudo às vogais, essa restrição faz referência a segmentos não periféricos na ocupação do espaço acústico.

²⁵ Todos os valores em hz registrados na formulação das restrições foram obtidos pela produção das vogais dos informantes, tanto em posição tônica, como em posição átona final. Para mais detalhes, veja-se o estudo de Matzenauer, Neuschrack, Carniato e Azevedo (2015).

Segundo a formalização proposta no estudo de Matzenauer, Neuschrack, Carniato e Azevedo (2015), o *Tableau* 1 representa com uma única hierarquia de restrições (ou seja, com uma única gramática) o fenômeno estudado: um possível sistema de cinco vogais /a, e, i, o, u/, no nível fonológico, é neutralizado, em posição átona, em três vogais fonéticas na produção [ɐ, ɪ, ʊ]; na percepção, as três vogais fonéticas são categorizadas em possíveis cinco vogais fonológicas.

O *Tableau* 1²⁶ apresenta, portanto, os resultados obtidos no confronto entre as restrições de estrutura e de pista entendidas como pertinentes para a formalização do fenômeno estudado. Destaca-se que a parte do *tableau* identificada pela chave em azul representa o processo de “produção”, cujos outputs são as formas fonéticas [ɐ, ɪ, ʊ]; a parte do *tableau* identificada pela chave em vermelho representa o processo de “percepção”, cujos *outputs* são predominantemente os segmentos /a, i, u/, sendo que os dois últimos mostram variação, respectivamente, com /e/ e /o/.

Tableau 1. Produção e Percepção das vogais do PB em posição átona final

	ranking value	disharmony	plasticity
*+tense	112.417	115.864	1.000000
*/a/[a]	110.308	111.146	1.000000
*/e/[e]	106.878	107.984	1.000000
*/o/[o]	103.653	106.674	1.000000
*/u/[u]	100.916	102.865	1.000000
*/i/[i]	100.970	102.073	1.000000
*/ɔ/[ɐ]	100.000	100.857	1.000000
*/e/[ɪ]	99.040	99.593	1.000000
*/o/[ʊ]	98.450	98.871	1.000000
LNoncorner	95.837	97.516	1.000000
*/u/[ʊ]	96.971	94.880	1.000000
*/a/[ɐ]	89.692	91.788	1.000000
*/i/[ɪ]	93.112	90.795	1.000000

²⁶ As restrições de pista refletem a passagem do contínuo para o discreto, assim, devem conter apenas pistas acústicas no nível contínuo e informação discreta no nível fonológico. Entretanto, para facilitar a representação das restrições no *tableau* apresentaremos as restrições com o segmento apenas - ex.: */a/[a].

	112.417	110.308	106.878	103.653	100.916	100.970	100.000	99.040	98.460	95.837	96.971	89.692	93.112
pep/a/	*+tense	*/a/[a]	*/e/[e]	*/o/[o]	*/u/[u]	*/i/[i]	*/o/[ɤ]	*/e/[ɪ]	*/o/[ʊ]	*Licence Noncorner	*/u/[ʊ]	*/a/[ɤ]	*/i/[ɪ]
/a/ pep[a]		*!											
↻ /a/ pep[ɤ]										.		.	
pep/e/													
/e/ pep[e]	*!		.							.			
↻ /e/ pep[ɪ]								.		.			
pup/i/													
↻ /i/ pup[ɪ]						.				.			.
/i/ pup[i]	*!												
pup/o/													
/o/ pup[o]	*!		.							.			
↻ /o/ pup[ʊ]									.	.			
pop/u/													
↻ /u/ pop[ʊ]					.					.	.		
/u/ pop[u]	*!												
pep[ɤ]													
↻ [ɤ] pep/a/												.	
[ɤ] pep/o/	*!						.			.			
pup[ɪ]													
[ɪ] pup /e/	.							*!		.			
↻ [ɪ] pup /i/	.											.	
pop[ʊ]													
[ʊ] pup /o/	.								*!	.	.		
↻ [ʊ] pup /u/	.										.		

Fonte: Matzenauer, Neuschrack, Carniato e Azevedo (2015)

Observem-se, inicialmente, os dados de produção no *Tableau* 1 (conforme já foi referido, no modelo de Boersma, a produção é formalizada no Módulo da Produção, responsável pelo mapeamento do *input* fonológico no *output* fonético):

- a vogal /a/ não sofre neutralização; o seu *output*, na produção, como [ɤ] tem motivação fonética; é, portanto, uma restrição de pista que determina a escolha da forma [ɤ] para o segmento /a/ (é determinante a restrição (c) em (9b): */+baixo, dorsal - a/ [F1 - 648hz / F2 - 1432hz - ɤ]);
- para as outras vogais, há neutralização; o mapeamento: /i/ e /e/ → [ɪ] e /o/, /u/ → [ʊ] é determinado pela interação entre restrições de estrutura e de pista, em que se mostra decisiva a restrição */+tense] (as restrições de pista relevantes são (f), (g), (j) e (l) em (9b)).

Vejam-se, agora, os dados de percepção no *Tableau* 1 (como já foi destacado, no modelo de Boersma, a percepção é formalizada no Módulo da Compreensão, responsável pelo mapeamento do *input* fonético no *output* fonológico):

- a vogal /a/ é o *output* fonológico para [ɤ] porque, em competição com outros *outputs* possíveis, não viola as restrições de estrutura e de pista que, na gramática, são decisivas para o licenciamento do sistema periférico para as átonas finais;

- b) as vogais /i/ e /u/, diferentemente, são respectivamente os *outputs* fonológicos para [ɪ, ʊ] pela interação entre restrições de estrutura e de pista: não violam as restrições de estrutura e de pista que, na gramática, são decisivas para o licenciamento do sistema periférico para as átonas finais; ao violarem as restrições de pista “*não perceba [ɪ] fonético como /e/ fonológico*” e “*não perceba [ʊ] fonético como /o/ fonológico*”, os *outputs* acrescentam os segmentos /e/ e /o/ ao sistema de vogais átonas finais.

No *Tableau* 1 está também formalizada a variação que há na categorização fonológica das formas fonéticas [ɪ, ʊ]: [ɪ] → /i/ ~ /e/; [ʊ] → /u/ ~ /o/; tal variação já aparece registrada em (8). Essa possibilidade de variação é representada pelos valores centrais (*ranking value*) das restrições que estão relacionadas à disputa do candidato ótimo: a diferença entre tais valores é inferior a 10 pontos²⁷. Veja-se o esquema em (10).

(10)

(a) candidatos em variação:

candidato [ʊ] → /u/ (percentual de percepção: 74%)

candidato [ʊ] → /o/ (percentual de percepção: 26%)

(b) restrições em conflito para a escolha dos dois outputs fonológicos em variação (diferença inferior a 10 pontos, o que oferece possibilidade de alteração de posição na hierarquia – veja-se nota 27):

*/o/[ʊ] – valor central: 98.460

*/u/[ʊ] – valor central: 96.971

*Licence Noncorner – valor central: 95.837

Observe-se ainda que o valor da restrição que interpreta [ʊ] → /o/ é superior àquela que interpreta [ʊ] → /u/; esse fato evidencia a maior frequência da escolha do candidato a *output* ótimo /u/. Leitura da mesma natureza está estabelecida, no *Tableau* 1, para a variação [ɪ] → /i/ ~ /e/, com prevalência para a escolha do *output* /i/.

²⁷ Na OT Estocástica, diferença do valor central de restrições inferior a 10 pontos é indicativo da possibilidade de alteração de posição, entre essas restrições, na hierarquia.

O que merece destaque é que, no estudo de Matzenauer, Neuschrack, Carniato e Azevedo (2015), o *Tableau 1* foi capaz de formalizar, em uma única gramática, a percepção e a produção das vogais átonas do PB falado no sul do país: representou o mapeamento, na percepção, de três formas fonéticas para possíveis cinco fonológicas, com supremacia dos segmentos /i, u, a/, bem como representou, na produção, o mapeamento de cinco possíveis segmentos fonológicos em três formas fonéticas [ɪ, ʊ, ə].

4. Considerações finais

Com foco no sistema de vogais átonas finais do PB, o presente artigo foi orientado por objetivo que incluía a discussão de dois pontos: (a) a constituição do sistema de vogais /i, u, a/ (que no PB opera, na posição átona final, como um subsistema das sete vogais da língua) e (b) o mapeamento entre os níveis fonético e fonológico dessas vogais, trazendo-se dados de produção e de percepção dos segmentos por falantes nativos de português do Brasil.

A constituição do sistema vocálico da posição mais débil da palavra – a átona final – por apenas três vogais /i, u, a/, quando na posição tônica a fonologia da língua integra sete segmentos, encontra amparo no contraste máximo preferido pelos inventários fonológicos. Quatro tipos de evidências podem oferecer suporte para a escolha desse sistema vocálico átono final:

1º) essas três vogais, em suas formas fonéticas, apresentam grande dispersão no triângulo acústico, ocupando a sua periferia e, assim, atendem à preferência das línguas; essa preferência, segundo Maddieson (1984), aponta para dispersão larga e balanceada no espaço fonético; para Lindblom (1986), o inventário formado por essas vogais apresenta dispersão máxima e, portanto, sua manifestação fonética mostra ambiguidade acústica mínima; por sua estabilidade articulatória e acústica, diz Crosswhite (2001) que as vogais [i, u, a] são preferidas não apenas por si mesmas, mas também como um conjunto;

2º) na dimensão fonológica, é o conjunto de vogais que contrastam pelos traços que, segundo Clements (2001) e Calabrese (2005), poderiam ser considerados os mais robustos na constituição de sistemas vocálicos (/baixo, alto, coronal, dorsal/);

3º) em estudo de tipologias de línguas, o conjunto das vogais /i, u, a/ é o de presença mais frequente nos inventários fonológicos; para Maddieson (1984) e Lindblom (1986), essas três

vogais constituem o menor inventário completo encontrado com alguma regularidade em línguas do mundo;

4º) em pesquisas sobre o processo de aquisição da linguagem, as vogais /i, u, a/ são as primeiras a emergir (RANGEL, 2002; MATZENAUER & MIRANDA, 2007, 2009; MATZENAUER, 2012), sendo que ocupam os espaços fonético-fonológicos das vogais médias antes da incorporação das vogais /e, o, ε, ə/ ao sistema da criança.

Quanto ao mapeamento entre os níveis fonético e fonológico das vogais átonas finais, particularmente no sul do Brasil, com o suporte de dados de produção e de percepção dos segmentos por falantes nativos de PB, o estudo realizado por Matzenauer, Neuschrank, Carniato e Azevedo (2015) concluiu que, na produção, há o registro de três vogais na posição átona final [ɪ, ʊ, ɐ] e que, na percepção, a partir dessas três formas fonéticas, os falantes de PB mapeiam predominantemente as vogais /i, u, a/, podendo também categorizar cinco vogais, em razão de alguma variação ao representar segmentos a partir das vogais [ɪ, ʊ]: a partir do input [ɪ], a categorização diferente de /i/ chega a apenas 1%; a partir do input [ʊ], a categorização diferente de /u/ chega a 26%.

A categorização majoritária das vogais átonas finais como os três segmentos /i, u, a/, assim como a sua representação fonética como [ɪ, ʊ, ɐ], pode dizer-se, vem ao encontro do que sustenta Bisol (2003, p.6-7): “é possível afirmar que no português brasileiro, como um todo, a neutralização da átona final é um processo em andamento no que diz respeito à opção pela vogal alta, uma vez que a variação permanece em certas comunidades”. A autora destaca que “a neutralização entendida como perda do traço distintivo entre vogais médias e altas é uma regra geral nesta posição, e que a preferência à realização da vogal alta tende a generalizar-se. O contraste fonológico fica restringido a três vogais, independentemente do alofone que se realiza”. Salienta-se, pois, que o sistema não utiliza, embora pudesse fazê-lo, a oposição entre vogais altas e médias em posição átona final, a não ser em casos como *júri* ≠ *jure* (CÂMARA JR., 1970). É exatamente esse funcionamento da fonologia da língua que abre espaço para a neutralização das vogais átonas finais.

Os dados aqui discutidos e formalizados vêm ao encontro do que diz Bisol (2010, p.45), ao analisar dados representativos das três capitais do Sul do Brasil: “O sistema de três vogais da átona final e o de cinco são características do português brasileiro [...] algumas variedades ainda não chegaram à concretude da neutralização da átona final...”. Observe-se que a afirmação de Bisol (2010) tem base em dados de produção de vogais átonas finais e as

conclusões do presente artigo, compatíveis com a de Bisol, têm base também em dados de percepção de vogais átonas finais por falantes de PB.

Referências Bibliográficas

AZEVEDO, R. Q.; PEDONE, M.; KOHLS-RIBEIRO, D. O processo de aquisição do sistema vocálico tônico do PB: estudo de caso. **Salão Universitário da UCPEL**. Pelotas: UCPEL, 2014.

BISOL, L. **Harmonização vocálica**: uma regra variável. Tese (Doutorado em Letras). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1981.

BISOL, L. Epílogo. In: BISOL, L.; BRESCANCINI, CR. (Eds.): **Fonologia e Variação**. Recortes do português brasileiro. Porto Alegre: EDIPUCRS, 2002. p.303-312.

BISOL, L. Neutralização das átonas. **D.E.L.T.A.** São Paulo: v.19, n.2, 2003.

BISOL, L. A Simetria no Sistema Vocálico do Português Brasileiro. **Linguística - Revista de Estudos Linguísticos da Universidade do Porto**. Porto: v. 5, p.45-52, 2010.

BOERSMA, P. Prototypicality judgments as inverted perception. In: FANSELOW, G., FÉRY, C.; SCHLESEWSKY, M. and VOGEL, R. (eds.) **Gradience in grammar**. Oxford: Oxford University Press, 2006. **crossref**
<http://dx.doi.org/10.1093/acprof:oso/9780199274796.003.0009>

BOERSMA, P. Cue constraints and their interactions in phonological perception and production. **Rutgers Optimality Archive** 944, 2007.

BOERSMA, P. Emergent ranking of faithfulness explains markedness and licensing by cue. **Rutgers Optimality Archive** 954, 2008.

BOERSMA, P. A programme for bidirectional phonology and phonetics and their acquisition and evolution. 2009. **crossref** <http://dx.doi.org/10.1075/la.180.02boe>

BOERSMA, P. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: BENZ, A. & MATTAUSCH, J. (eds.) **Bidirectional Optimality Theory**, 33-72. Amsterdam: John Benjamins, 2011.

BOERSMA, P.; ESCUDERO, P.; HAYES, R. Learning Abstract Phonological from Auditory Phonetic Categories: An Integrated Model for the Acquisition of Language-Specific Sound Categories. **Rutgers Optimality Archive** 585, 2003.

BOERSMA, P.; ESCUDERO, P. Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In: ARUNACHALAM, S.; KAISER, E.; WILLIAMS, A. (eds.) **Proceedings of the 25th Annual Penn Linguistics Colloquium**. (Penn Working Papers in Linguistics 8.1), 2003.

BOERSMA, P.; HAMANN, S. In: CALABRESE, A.; WETZELS, W. L. (eds.) **Loanword phonology**. Amsterdam: John Benjamins, 2009.

BOERSMA, P.; CHLÁDKOVÁ, K. Asymmetries between speech perception and production reveal phonological structure. In: **ICPhS XVII - Proceedings**. Hong Kong: 2011.

BOERSMA, P.; HAYES, B. Empirical Tests of the Gradual Learning Algorithm. **Linguistic Inquiry** 32, p. 45-86, 2001. **crossref** <http://dx.doi.org/10.1162/002438901554586>

BOERSMA, P.; WEENINK, D. **PRAAT**: doing phonetics by computer [computer program], 2012. Disponível em <http://www.praat.org/>.

CALABRESE, A. A constraint-based theory of phonological markedness and simplification procedures. **Linguistic Inquiry**, v. 26, n.3, p.373-463, 1995.

CALABRESE, A. **Markedness and economy in a derivational model of phonology**. New York: Mouton de Gruyter, 2005. **crossref** <http://dx.doi.org/10.1515/9783110197600>

CÂMARA JR, J. M. **Estrutura da Língua Portuguesa**. Petrópolis: Vozes, 1970.

CÂMARA JR, J. M. **Para o estudo da fonêmica portuguesa**. Rio de Janeiro: Padrão, 1977.

CLEMENTS, G. N. Representational economy in constraint-based phonology. In: HALL, A. (ed.) **Distinctive Feature Theory**. Berlin: Mouton de Gruyter, 2001. **crossref** <http://dx.doi.org/10.1515/9783110886672.71>

CROSSWHITE, K. Vowel Reduction. In: HAYES, B.; KIRCHNER, R.; STERIADE, D. (eds.) **Phonetically Based Phonology**. Cambridge: Cambridge University Press. 2001. p. 191-232

GUY, G. As Vogais: Perspectivas Analíticas e Teóricas. In: **II SIS-Vogais** – conferência de encerramento. Belo Horizonte: UFMG, 2009.

LINDBLOMB, B. Phonetic universals in vowel systems. In: OHALA, J. J. & JAEGER, J. J. (eds.) **Experimental Phonology**, Orlando, Florida: Academic Press, 1986, p.13-44.

MADDIESON, I. **Patterns of Sounds**. Cambridge: Cambridge University Press, 1984. **crossref** <http://dx.doi.org/10.1017/CBO9780511753459>

MATZENAUER-HERNANDORENA, C. L. **Aquisição da fonologia do Português: estabelecimento de padrões com base em traços distintivos**. Tese (Doutorado em Letras). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1990.

MATZENAUER, C. L. B. Aquisição das vogais do "PB" e tipologias de línguas. In: LEE, S. H. (org.). **Vogais além de Belo Horizonte**. Belo Horizonte: Faculdade de Letras da Universidade Federal de Minas Gerais, 2012.

MATZENAUER, C. L. B.; MIRANDA, A. R. M. Traços distintivos e a aquisição das vogais do PB. **I SIS-VOGAIS**. João Pessoa: UFPB, 2007.

MATZENAUER, C. L. B.; MIRANDA, A. R. M. Traços distintivos e a aquisição das vogais do PB. In: HORA, D. da. (org.). **Vogais no Ponto mais Oriental das Américas**. João Pessoa: Ideia, 2009, p. 45-63.

MATZENAUER, C. L. B.; NEUSCHRANK, A.; CARNIATO, M. C.; AZEVEDO, R. Q. Vogais em posição postônica final: percepção e produção (no Sul do Brasil). **Revista da ABRALIN**, v. 14, n.1, p.17-42, 2015.

PATER, J. Bridging the gap between receptive and productive development with minimally violable constraints. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (eds.) **Constraints in Phonological Acquisition**. Cambridge: Cambridge University Press. 2004. p.219-244.

RANGEL, G. de A. **Aquisição do sistema vocálico no português brasileiro**. Tese (Doutorado em Letras). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

RAUBER, A. et al. **TP: perception tests and perceptual training with immediate feedback**, versão 3.1. 2012. Disponível em: http://www.worken.com.br/tp_regfree.php. Último acesso em 28 fev. 2015.

SANTOS, G. R.dos. **Percepção e produção das vogais médias do espanhol por falantes do português brasileiro**. Tese (Doutorado em Letras). Universidade Católica de Pelotas, Pelotas, 2014.

STERIADE, D. Complex Onsets as Single Segments. In: COLE, J. & KISSEBERTH, C. (eds) **Perspectives in Phonology**, Stanford: CSLI. 1994. p. 203-291.

STEVENS, K. On the quantal nature of speech. **Journal of Phonetics**. v.17, p. 3-45, 1989.

SCHWARTZ, J. L.; BOË, L. J.; VALLÉE, N.; ABRY, C. The Dispersion-Focalization Theory of vowel systems. **Journal of Phonetics**. v.25, n.3, p. 255-286, 1997. **crossref** <http://dx.doi.org/10.1006/jpho.1997.0043>

VIEIRA, M. J. B. As vogais médias postônicas. Uma análise variacionista. In: BISOL, L.; BRESCANCINI, C. R. (eds.): **Fonologia e Variação**. Recortes do português brasileiro. Porto Alegre: EDIPUCRS, 2002. p.127-159.

VIEIRA, M. J. B. As vogais médias átonas nas três capitais do sul do País In: BISOL, L.; COLLISCHONN, G. (orgs) **Português do Sul do Brasil: variação fonológica**. Porto Alegre: EDIPUCRS, 2009. p.50-72.

WANROOIJ, K.; BOERSMA, P.; VAN ZUIJEN, T. L. Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study. **Frontiers in Psychology**, 2014. **crossref** <http://dx.doi.org/10.3389/fpsyg.2014.00077>

Artigo recebido em: 14.09.2015

Artigo aprovado em: 05.03.2016

Uma abordagem fonológica para as postônicas médias não-finais

A Phonological approach to non-final post-tonic mid-vowels

Arthur Pereira Santana*

RESUMO: Este artigo analisa as vogais médias postônicas não-finais. Especificamente, a motivação fonológica para a emergência das formas altas e, em dialetos em que isso é possível, para a emergência de médias-baixas. Para tanto, realizou-se um experimento controlado com 40 indivíduos de duas localidades, São Paulo e São Luís, que resultou em um corpus balanceado de 4720 palavras. Após análise estatística, observou-se, principalmente, a correlação das formas assumidas na postônica não-final com a vogal átona final. Por meio da Geometria de traços (CLEMENTS, 1985) e assumindo a proposta de Wetzels (2011) de neutralização como um mecanismo de mudança no valor do traço, propõem-se duas regras fonológicas: (i) a primeira, para a emergência das vogais altas (ex.: prót[i]se, díal[u]go), um espriamento do nó vocálico da átona final para a postônica não-final; (ii) e a segunda, para a emergência das médias-baixas câm[ɛ]ra; agríc[ɔ]la), um espriamento de [+aberto3] da átona-final para a postônica não-final.

ABSTRACT: This article discusses mid-vowels in non-final post-tonic position in Brazilian Portuguese, specifically, the phonological motivation for the emergence of high vowels and, in dialects that allows it, mid-low vowels in the referred position. Thus, we carried out a controlled experiment with 40 individuals from two Brazilian capitals, São Paulo (Southern dialect) and São Luís (northern dialect), which resulted in a balanced corpus of 4720 tokens. After statistical analysis, it was observed the correlation of the form that emerged in non-final post-tonic position with the final unstressed vowel. Assuming the Feature Geometry framework (CLEMENTS, 1985) and Wetzels' (2011) proposal for neutralization as a mechanism of feature change, we propose two phonological rules: (i) one for the emergence of high vowels, as a result of vocalic node spreading from the final unstressed vowel to the non-final post-tonic vowel; (ii) and a second rule for the emergence of low-mid vowels resulting from [+open3] spreading, also from the unstressed final position to the non-final post-tonic position.

PALAVRAS-CHAVE: Neutralização. Vogais médias. Postônica não-final.

KEYWORDS: Neutralization. Mid-vowels. Non-final post-tonic.

1. Introdução

Câmara Jr. (1970) foi o primeiro autor a classificar, para o Português Brasileiro, a elevação do grau de altura das vogais médias em posições átonas como resultado de um processo de neutralização. Para ele, as sete vogais tônicas /a, ɛ, e, i, ɔ, o, u/ são reduzidas a cinco na pretônica /a, e, i, o, u/, a quatro na postônica não-final /a, e, i, u/ e a três na átona final

* Doutorando em Linguística no Programa de Pós-graduação em Semiótica e Linguística Geral da Universidade de São Paulo. Bolsista do CNPq (processo 140139/2015-0). arthurpereirasantana@gmail.com

/a, i, u/. Por privilegiar as formas menos marcadas e já existentes no inventário da língua, a neutralização tende a resultar em um subsistema mais simples, o que pode facilmente ser observado nas línguas do mundo, uma vez que as vogais altas tendem a ser mais comuns entre os inventários das línguas. Característica dos dialetos nordestinos, a emergência de médias-baixas em posições átonas chama atenção por (i) não condizer com a tendência geral de que formas menos marcadas¹ são sempre as que emergem após regras de neutralização; (ii) não ser recorrente em outros dialetos do País; (iii) apresentar comportamento variável e (iv) por nem sempre parecer que exista uma motivação aparente que condicione tal configuração.

Enquanto diversos estudos a respeito do vocalismo no PB tratam das vogais em posição pretônicas, as postônicas não-finais carecem de análises que consigam capturar em termos formais o fato de (i) haver alternância entre médias-altas e altas em todos os dialetos do País e de, (ii) especificamente em dialetos do nordeste, as médias-baixas também emergirem na referida posição (cf. Silva, 2010; Santana 2015). É este, pois, o objetivo desta análise.

2. Neutralização e vogais no PB

Como dito anteriormente, no PB e em diversas línguas do mundo, as vogais se neutralizam a depender da posição em que ocorrem. Em termos autosegmentais, a noção clássica de neutralização é a de que se trata de um processo por meio do qual um traço responsável pela distinção de dois fonemas é perdido, isto é, é apagado da configuração (CLEMENTS, 1985). Assim, no PB, por exemplo, os traços [coronal] e [-aberto1] são suficientes para diferenciar a vogal /i/ das outras vogais na átona final, uma vez que [aberto2] e [aberto3] são neutralizados nesta posição.

Por sua vez, Wetzels (2011) argumenta a favor de que a regra de neutralização não seja entendida como a perda de um traço, mas sim como um mecanismo pelo qual o valor distintivo de um traço seja substituído por seu valor oposto na camada em que a distinção é definida. Dessa forma, voltando ao exemplo do /i/ na átona final, a neutralização, em vez de capturada pela perda de [aberto2] e de [aberto3], deveria ser entendida como a substituição de [+aberto2] e [+aberto3] por [-aberto2] e [-aberto3] – veja que o valor positivo é o que assegura a distinção entre as médias-baixas e as demais vogais e é por isso que, após a neutralização, os traços de

¹ Um dos motivos para tais formas serem consideradas menos marcadas se dá pelo fato de elas serem as primeiras a sofrerem neutralização dentre as vogais médias.

abertura assumem o valor oposto. Assim, assumindo esta proposta, a vogal /i/ na átona final possuiria em sua configuração [coronal], [-aberto1], [-aberto2] e [-aberto3].

A respeito das postônicas não-finais, há um impasse no que diz respeito à configuração do subsistema, tendo em vista que não há um consentimento sobre o *status* das vogais médias: Câmara Jr. (1977) defende que somente a média-alta anterior /e/ faz parte do subsistema, tendo sido perdida a distinção entre /o/ e /u/. Bisol (2003), por outro lado, argumenta contrariamente à hipótese de uma configuração assimétrica e propõe que devido ao alto nível de aplicação do fenômeno de alçamento das médias na postônica não-final em dialetos do Sul, o subsistema desta posição está em vias de mudança para uma configuração mais simples, tal qual a da átona final, constituída por três vogais. Em suma, Câmara Jr. (1977) assume um subsistema para as postônicas não-finais formado por /a, e, i, u/ e Bisol (2003) acredita em uma mudança iminente para um sistema constituído por /a, i, u/. Em termos autosegmentais, captura-se a proposta de Câmara Jr. (1977) por meio da neutralização de [+aberto2] somente para as labiais na postônica não-final; já a proposta de Bisol (2003) é capturada por meio da neutralização de [+aberto2] para ambas as vogais médias.

Entretanto, as médias na postônica não-final não são alvo de apenas um fenômeno de altura que resulta na emergência de vogais altas na posição. Em dialetos nordestinos há casos de emergência de médias-baixas, tal qual em abób[ɔ]ra e câm[ɛ]ra, fatos descritos por Silva (2010) e Santana (2015) e que precisam ser explicados.

Se assumirmos a proposta de Câmara Jr. (1977) ou a de Bisol (2003), temos, então, que explicar o fato de as vogais médias serem as que mais emergem nos dialetos controlados neste estudo – como veremos adiante. Além disso, nos casos em que as vogais emergem como médias-baixas em dialetos do nordeste, se deveria assumir um processo de abaixamento de dois níveis de altura na língua (de /u/ para [ɔ] e de /i/ para [ɛ]), o que não encontramos em nenhum outro processo do PB. Por fim, ainda há casos em que a vogal alta parece ser bloqueada, como em *óp[i]ra ou *hét[i]ro, o que não deveria acontecer caso a vogal alta fosse de fato a forma subjacente.

Dessa forma, propomos que o subsistema postônico não-final seja /a,e,i,o,u/, tal qual o pretônico. Isto é, composto de uma vogal baixa, das altas anterior e posterior e também de ambas as vogais médias-altas, que podem se concretizar como [e, i, ɛ] e [o, u, ɔ]. Assumindo este subsistema, faz-se necessário explicar como se dá a emergência das vogais altas e das

médias-baixas. Neste artigo, com base nas observações empíricas que serão descritas a seguir, propomos duas regras distintas.

3. Metodologia

Uma vez que os resultados conhecidos de neutralização em postônica não-final são, em sua maioria, de dialetos do sul/sudeste, e os dialetos do norte/nordeste parecem apresentar outra direção de neutralização, optamos por analisar dados de dois dialetos: São Paulo (SP) e São Luís (MA), de forma a comparar os resultados encontrados, buscando similaridades e diferenças que pudessem nos ajudar a compreender melhor o fenômeno em questão.

Tendo em vista os desafios metodológicos dos estudos que tem por objeto as palavras proparoxítonas, como o fato de as palavras, em geral, serem menos usuais e a consequente dificuldade de encontra-las em número satisfatório e com boa distribuição de contextos para análises estatísticas em *corpus* de fala espontânea, optamos por um experimento controlado a fim de que pudéssemos contar com uma elevada quantidade de dados para posterior análise estatística.

Neste artigo, optamos por discutir somente as variáveis de caráter fonológico, a saber: o *Ponto de Articulação da Vogal tônica*, o *Ponto de Articulação da Vogal átona final*, o *Ponto de Articulação do Contexto fonológico precedente*, o *Ponto de Articulação do Contexto fonológico seguinte*, a *Altura da Vogal Tônica* e a *Altura da Átona Final*. Os segmentos foram classificados com base no ponto de articulação (Labial, Coronal ou Dorsal). Com relação à altura da tônica, classificamos as vogais em Baixa, Média-Baixa, Média-alta ou Alta, e para a altura da átona final, em Baixa ou Alta.

Selecionamos um total de 118 palavras que foram organizadas da seguinte forma: 59 palavras com vogal anterior em posição postônica não-final – 30 usuais (ex.: *cérebro*) e 29 não usuais (ex.: *diamantífero*); 59 palavras com vogal média posterior em posição postônica não-final – 30 usuais (ex.: *árvore*) e 29 não usuais (ex.: *necrópole*).² A fim de contar com um

² Classificar vocábulos de acento antepenúltimo no Português com base no grau de usualidade não é uma tarefa fácil. Primeiramente, porque a noção de usualidade não é igualmente compartilhada por todos; em segundo lugar, porque a usualidade de determinado item lexical também pode não ser compartilhada igualmente por diferentes indivíduos; e em terceiro lugar, porque são muito escassas listas de frequência ou de usualidade dos itens lexicais do Português. Então, para que tal classificação não fosse feita somente a partir do julgamento dos pesquisadores envolvidos neste trabalho, optou-se por fazer uso do buscador de frequência do Projeto ASPA – Avaliação Sonora do Português Atual (CRISTÓFARO-SILVA *et. al*, 2009).

distribuição balanceada de contextos fonológicos, fez-se uso de logatomas que foram classificadas como palavras não-usuais.

Foram gravados vinte informantes de cada localidade, homens e mulheres, com idade entre vinte e trinta anos, todos universitários ou com ensino superior completo, naturais de São Paulo e de São Luís e que não tivessem se ausentado das localidades por mais de um ano. A produção das 118 palavras pelos quarenta informantes proporcionaram um corpus de 2236 proparoxítonas com vogal anterior na postônica não-final e 2236 com vogal posterior, totalizando um corpus total de 4720 palavras.

A definição da vogal produzida pelo informante (alta, média-alta, média-baixa) foi feita por meio de uma análise de medição de formantes³ com uso do software PRAAT (BOERSMA & WEENINK, 2013). Após esta etapa de classificação, foram realizados dois testes estatísticos: o primeiro, de Qui-Quadrado, que foi feito por meio do software Stata; e o segundo, que se vale de técnicas de Regressão Logística, feito por meio do software Varbrul. O uso de dois testes estatísticos buscou observar se certas características da emergência das formas altas e médias-baixas (no dialeto em que isso é possível) não capturadas por meio de uma técnica tornar-se-ia evidente com outra.

4. Resultados

Uma vez que os dados foram codificados com base na produção efetiva dos falantes, as vogais médias anterior e posterior poderiam assumir três formas: a de vogal média-alta [e, o], a de vogal alta [i, u] e, como será observado nos dados de São Luís, a de vogal média-baixa [ɛ, ɔ]. Não houve registro de caso de vogais médias-baixas na postônica não-final nos dados de São Paulo, como se pode ver na Tabela 1 a seguir.

Tabela 1: Distribuição Geral das vogais médias em São Paulo e São Luís.

		[i]	[e]	[ɛ]	[u]	[o]	[ɔ]	Total
SP	Número de casos	149 (12,6%)	1031 (87,4%)	0 (0%)	149 (12,9%)	1031 (87,4%)	0 (0%)	2360
SL	Número de casos	107 (9%)	755 (64%)	318 (27%)	100 (8,4%)	831 (70,4%)	249 (21,2%)	2360
								4720

Fonte: Dados da Pesquisa.

³ Utilizamos, para classificação, a tendência observada e descrita nas análises de Ecudeiro *et. al* (2009) para as tônicas e de Machado (2010) para as pretônicas. Entretanto, observamos, também, a tendência individual de cada informante e dos dados do *corpus* da pesquisa.

Nos dados de São Paulo, a vogal anterior foi produzida como média-alta 87% das vezes (ex.: ['sɛlebrɪ] ‘célebre’, ['kãmerɐ] ‘câmera’ e [dʒi'ẽmetɾu] ‘diâmetro’), e em 12% dos casos como alta (ex.: ['ɔspɪdʒɪ] ‘hóspede’, ['prɔtʃɪzɪ] ‘prótese’ e ['ĩgrɪmɪ] ‘íngreme’); enquanto que a vogal posterior em 83% das vezes ocorreu como média-alta (ex.: [a'grɪkolɐ] ‘agrícola’, ['sĩkopɪ] ‘síncope’ e [e'kivokʊ] ‘equivoco’) e em 16% dos casos como alta (ex.: [ʃe'nɔfɔbu] ‘xenófobo’, [me'tɾɔpɔli] ‘metrópole’ e [o'rɔskɔpu] ‘horóscopo’).

A respeito dos dados de São Luís, chama atenção o fato de, nos dados coletados para esta pesquisa, em conformidade com o que foi observado por Silva (2010) e Santana (2013), haver mais casos de derivadas médias-baixas na postônica não-final do que de derivadas altas. Observamos que enquanto os casos de derivadas altas anterior e posterior não chegam a 10%, a presença de vogal média-baixa corresponde a 21,2% dos casos para a anterior e a 27% para a vogal posterior (ex. ['nadɛgɐ] ‘nádega’, [a'grɪkolɐ] ‘agrícola’). Além disso, há uma predominante hegemonia em ambas as pautas da forma média-alta, de 64% dos casos para a vogal anterior e de 70,4% dos casos para a vogal posterior.

Os resultados dos testes estatísticos para cada localidade serão apresentados nas subseções a seguir.

4.1 São Paulo

Os resultados obtidos com o Varburl apontaram correlação entre a emergência da vogal alta [i] com as variáveis *Ponto de Articulação da Vogal Átona Final*, *Ponto de Articulação do Contexto Fonológico Precedente* e *Ponto de Articulação do Contexto Fonológico Seguinte*, como apresentado na Tabela 2.

Tabela 2: Correlações com a emergência de [i] – SP.

	Favorecedor	Neutro	Inibidor
Átona final	Coronal (0.83)	–	Labial (0.34) Dorsal (0.29)
Contexto precedente	Coronal (0.74)	Dorsal (0.45)	Labial (0.28)
Contexto seguinte	Coronal (0.65)	Dorsal (0.46)	Labial (0.36)

Como podemos ver, os fatores que favorecem a emergência de [i] são: vogal coronal na átona final (ex.: ['ɔspɪdʒɪ] ‘hóspede’), consoante dorsal no contexto precedente (ex.: ['kɔxɪgʊ] ‘córrego’) e consoante coronal no contexto seguinte (ex.: [ko'laʒɪmʊ] ‘colágeno’).

No que se refere aos resultados das vogais posteriores, observou-se a partir dos dados de São Paulo que as variáveis *Ponto de Articulação da Vogal Átona Final*, *Ponto de Articulação do Contexto fonológico seguinte* e *Ponto de Articulação da Vogal tônica* apresentaram correlação à emergência da vogal alta. Assim como observado para a vogal média anterior, o *Ponto de Articulação da Vogal Átona Final* foi selecionado pela análise como a variável que mais se correlacionou à forma [ʊ] em posição postônica não-final. Tais resultados foram dispostos na Tabela 3 a seguir, que mostra que a emergência de [ʊ] estava correlacionada a uma vogal labial na átona final (ex.: pẽ'tagonʊ] ‘pentágono’), a uma vogal labial na tônica (ex.: [a'bõburɛ] ‘abóbora’) e a uma consoante labial no contexto fonológico seguinte (ex.: [ẽ'tzilopi] ‘antílope’).

Tabela 3: Correlações com a emergência de [ʊ] – SP.

	Favorece	Neutro	Desfavorece
Vogal átona final	Labial (0.66)	–	Coronal (0.44) Dorsal (0.38)
Contexto seguinte	Labial (0.58)	Dorsal (0.51)	Coronal (0.40)
Tônica	Labial (0.57)	Coronal (0.51)	Dorsal (0,41)

Uma observação rápida dos dados nos mostra que se correlacionaram à emergência das formas altas segmentos que possuíam o mesmo ponto de articulação que o da vogal postônica não-final. Além disso, os resultados apresentados também chamaram atenção pelo fato de a variável *Altura da Átona Final* e *Altura da Vogal Tônica* não terem apresentado correlação nos resultados obtidos com o Varbrul, já que esperávamos que uma regra que afeta a altura de uma vogal fosse motivada por traços de altura das vogais adjacentes, e não por traços de ponto. Por conta disso, fizemos uso do teste de Qui-quadrado para observar detalhadamente a frequência de ocorrência e a possível correlação entre a emergência das formas das vogais médias e o grau de altura das vogais tônica e átona final.

A Tabela 4 mostra que a vogal anterior emergiu como alta mais vezes quando na tônica também havia uma vogal alta. Entretanto, o mesmo não aconteceu para a posterior, que emergiu mais vezes como vogal alta quando na posição de acento havia uma média-baixa.

Tabela 4: Comparação Altura da Vogal Tônica – SP.

Tônica	Anterior		Posterior	
	[i]	[e]	[o]	[o]
Alta	19,6%	80,4%	13,8%	86,2%
Média-alta	7%	93%	15%	85%
Média-baixa	17,7%	82,3%	22,6%	77,4%
Baixa	5,5%	94,5%	14,2%	85,8%
	<i>p-value</i> < 0.001		<i>p-value</i> = 0.001	

A respeito da *Altura da Vogal Átona Final*, a distribuição dos dados mostra a tendência geral de as vogais altas emergirem na postônica não-final quando na átona também havia uma vogal alta, como podemos ver na Tabela 5.

Tabela 5: Comparação Altura da Átona Final – SP.

Átona Final	Anterior		Posterior	
	[i]	[e]	[o]	[o]
Alta	18,5%	81,5%	19,2%	80,7%
Baixa	4,7%	95,3%	10,5%	89,5%
	<i>p-value</i> < 0.001		<i>p-value</i> = 0.001	

As tendências observadas nos resultados obtidos com os dados de São Paulo também foram observadas quando analisados os dados de São Luís, como apresentaremos na seção a seguir.

4.2 São Luís

Os resultados obtidos com o Varbrul mostraram que se correlacionaram à emergência da vogal anterior alta as seguintes variáveis: o *Ponto de Articulação do Contexto fonológico Precedente*, o *Ponto de Articulação da Vogal Tônica*, o *Ponto de Articulação do Contexto Fonológico Seguinte* e o *Ponto de Articulação da Vogal Átona Final*. Ou seja, todas as variáveis escolhidas para esta análise, como esquematizado na Tabela 6 a seguir.

Tabela 6: Correlações com a emergência de [ɪ] – SL.

	Favorece	Neutro	Desfavorece
Contexto Precedente	Coronal (0.88)	–	Dorsal (0.30) Labial (0.21)
Tônica	Coronal (0.66)	–	Labial (0.44) Dorsal (0.35)
Contexto seguinte	Coronal (0.72)	–	Labial (0.38) Dorsal (0.37)
Átona final	Coronal (0.70)	–	Labial (0.39) Dorsal (0.40)

Como podemos ver, a forma alta da vogal anterior está associada a segmentos coronais no contexto precedente (ex.: ['sɛlɪbrɪ] ‘célebre’, na tônica (ex.: [ɪ'tɛhpɪtʃɪ] ‘intérprete’, na átona final (ex.: [abo'ɾɪʒɪŋɪ] ‘aborígene’) e no contexto fonológico seguinte (ex.: [alusi'nɔʒɪnɔ]). Já os segmentos dorsais e labiais nas referidas posições estão correlacionados à inibição da emergência da vogal alta, tendo em vista os pesos abaixo do ponto de neutralidade.

Já para a vogal posterior, os resultados apontaram como variáveis correlacionáveis o *Ponto de Articulação do Contexto Fonológico Seguinte* e o *Ponto de Articulação da Vogal Átona Final*, como é possível ver na Tabela 7, que mostra que a vogal emerge como [ʊ] principalmente quando há uma vogal labial na átona final (ex.: [ʃɛ'nɔfɔbʊ] ‘xenófobo’) e uma consoante labial no contexto fonológico seguinte (ex.: [gas'trɔ̃nɔmɐ] ‘gastrônoma’).

Tabela 7: Resultado do teste para a emergência de [ʊ] – SL.

	Favorece	Neutro	Desfavorece
Contexto seguinte	Labial (0.58)	Coronal (0.48)	Dorsal (0.42)
Átona final	Labial (0.56)	Coronal (0.55)	Dorsal (0.37)

Já para a emergência da vogal anterior como média-baixa, mostrou-se correlacionável segmentos labiais e dorsais no contexto fonológico precedente (ex.: ['kãmɛɾɐ] ‘câmera’, ['prɔtɛzi] ‘prótese’) e por segmentos coronais na átona final (ex.: ['sĩtɛzi] ‘síntese’). Em ambos os contextos, segmentos coronais parecem inibir a emergência da forma (ex.: [ĩn'dʒizẽnɐ] ‘indígena’, [alje'ɲizẽnɐ] ‘alienígena’), como vemos na Tabela 8 abaixo.

Tabela 8: Resultado para a emergência de [ɛ] – SL.

	Favorece	Neutro	Desfavorece
Contexto Precedente	Labial (0.67) Dorsal (0.62)	–	Coronal (0.23)
Átona final	Dorsal (0.65)	Labial (0.50)	Coronal (0.33)

Por sua vez, a variável fonológica que se correlacionou à emergência da forma média-baixa da vogal posterior foi o *Ponto de Articulação da Átona Final* quando dorsal, tal qual para a média-baixa anterior, como podemos ver na Tabela 9 abaixo.

Tabela 9: Resultado para a emergência de [ɔ] – SL.

	Favorece	Neutro	Desfavorece
Átona final	Dorsal (0.68)	–	Coronal (0.44) Labial (0.23)

Similar ao que observamos nos dados de São Paulo, o resultado obtido com o uso do Varbrul não apontou correlação entre o grau de altura da vogal postônica não-final e o grau de altura das vogais adjacentes a partir dos dados de São Luís. Assim como fizemos com os dados da capital paulista, utilizamos o teste do Qui-quadrado para observar mais detalhadamente a frequência de ocorrência das formas das vogais nos dados da capital maranhense.

Como podemos observar na Tabela 10, a vogal anterior na postônica não-final emergiu como alta mais vezes quando havia uma vogal alta na tônica, enquanto que a posterior emergiu mais vezes como alta quando havia na tônica uma vogal média-baixa. Por sua vez, a postônica não-final emergiu mais vezes como [ɛ] quando na tônica havia uma vogal média-alta e como [ɔ] quando na posição de acento havia uma vogal alta.

Tabela 10: Comparação Altura da Vogal Tônica – SL.

Tônica	Anterior			Posterior		
	[i]	[e]	[ɛ]	[o]	[o]	[ɔ]
Alta	23,1%	67,1%	17,8%	6,5%	69,8%	23,7%
Média-alta	4,1%	43,3	52,6%	6,6%	79,2%	14,2%
Média-baixa	7,4%	71,6%	21%	11,9%	68,4%	19,7%
Baixa	3%	66%	31%	6,7%	70,3%	23%
	<i>p-value</i> < 0.001			<i>p-value</i> < 0.001		

A respeito da altura da átona final, observamos a partir dos dados de São Luís a mesma tendência observada nos dados de São Paulo, isto é, tanto na pauta anterior quanto na posterior as vogais emergiram mais vezes como altas quando na átona final também havia uma vogal alta, como se pode ver na Tabela 11. Além disso, os dados de São Luís também mostram o fato de as vogais médias-baixas emergirem mais vezes na postônica não-final quando, na átona final, havia uma vogal baixa.

Tabela 11: Comparação Altura da Átona Final – SL.

Átona Final	Anterior			Posterior		
	[i]	[e]	[ɛ]	[o]	[o]	[ɔ]
Alta	10%	67,8%	22,2%	9,4	76,2%	14,4%
Baixa	7,2%	56,3%	36,5%	6,5%	59%	34,5%
	<i>p-value</i> < 0.001			<i>p-value</i> < 0.001		

Uma observação rápida dos resultados obtidos a respeito das variáveis *Altura da tônica* e *Altura da átona final* nos leva a crer que, de fato, não parece haver indícios suficientes para apontar uma correlação entre a emergência das formas na postônica não-final e a altura da vogal tônica, principalmente quando se observa os dados da vogal posterior. Por outro lado, a altura da átona-final, em ambas as pautas, e com base nos dados das duas localidades, parece estar correlacionada à emergência das formas na postônica não-final. Discutiremos esse e todos os outros resultados apresentados anteriormente na seção a seguir.

5. Análise

Os resultados apresentados no tópico anterior serão utilizados agora como indícios para julgarmos duas hipóteses que serão levantadas com base nos estudos já realizados, bem como na teoria adotada para a análise, a Geometria de traços (CLEMENTS & HUME, 1995). As hipóteses são sobre: (i) a frequência de ocorrência e a relação entre anterioridade, posterioridade e grau de altura fonética das médias; (ii) e a emergência das formas altas e médias-baixas como resultado de uma regra fonológica.

5.1 A frequência de ocorrência e a relação entre anterioridade, posterioridade e grau de altura fonética das médias

Em diversos trabalhos, a frequência de ocorrência das vogais médias como derivadas altas na postônica não-final é apresentada a fim de discutir a hipótese levantada por Bisol (2003) a respeito de este subsistema estar em vias de mudança para uma configuração tal qual a da átona final, de só três vogais. Como discutido no tópico de metodologia, optamos por adotar um experimento controlado para o estudo pois se julgou necessária uma nova abordagem que viabilizasse maior controle das variáveis fonológicas. Além disso, diferentemente do que foi feito pela maioria dos estudos prévios, classificamos as vogais a partir de critérios acústicos, e não auditivos. É, pois, por conta da diferença de tipo de experimento, de coleta e de análise dos dados que comparações diretas entre a frequência de ocorrência das vogais a partir do

corpus deste estudo e a frequência observada por outras análises não serão utilizadas para nortear a análise, e sim a tendência geral observada.

Como vimos, a emergência da forma média-alta da vogal anterior e da posterior em ambas as localidades é predominante. Em São Paulo, na postônica não-final, a vogal anterior foi pronunciada como [e] em 87,4% dos casos, e como [o] em 83,7% dos casos. Em São Luís, como [e] em 64% das vezes e como [o] em 70,4% das vezes. Entretanto, o fato de a frequência de médias-altas em São Luís ser menor que a de São Paulo não significa que há mais casos da forma alta das vogais na localidade. Ao contrário, enquanto em São Luís a vogal anterior foi realizada como [i] em 9% dos casos e a posterior como [u] em 8,4% dos casos, em São Paulo a vogal anterior emergiu como [i] em 12,6% das vezes e a posterior como [u] em 16,3% das vezes.

Há menos casos de médias-altas na capital nordestina porque além das duas formas que emergem na posição em São Paulo, ainda existe a possibilidade de que médias-baixas sejam realizadas na postônica não-final. Segundo os resultados apresentados, a anterior emergiu como [ɛ] em 27% dos casos e a posterior como [ɔ] em 21,2% dos casos. Notamos, no caso das vogais anteriores, que as média-baixas apareceram três vezes mais do que as altas, enquanto que para as vogais posteriores, as média-baixas foram duas vezes e meia mais frequentes do que as altas. Tais resultados nos mostram que no dialeto nordestino analisado na postônica não-final, há mais casos de vogais médias-baixas [ɛ, ɔ] do que de formas altas [i, u].⁴ A pequena diferença entre as vogais altas em São Paulo e São Luís, aliada ao fato de que os fatores que se correlacionam à emergência das referidas formas são sempre os mesmos nas duas cidades são indícios de que o surgimento de médias-baixas se dá, em São Luís, em contextos em que essas vogais são produzidas como médias-altas em São Paulo.

Diversos estudos que tratam do vocalismo no PB defendem que há maior tendência à emergência das formas altas na pauta posterior do que na pauta anterior. As pesquisas que trataram das postônicas não-finais, em sua plenitude, argumentam a favor desta hipótese (cf. VIEIRA, 2002; RIBEIRO, 2007, SANTOS (2010), SILVA (2010)). Uma das hipóteses levantadas para explicar o porquê de haver mais casos de alçamento na pauta posterior se refere a questões de caráter articulatório. Bisol (2003) remete à proposta das vogais cardeais de Daniel Jones para argumentar que a hipótese de que há maior tendência de alçamento para as

⁴ Acreditamos que tal tendência seja geral para os dialetos nordestinos, com base no que também descreve Silva (2010).

posteriores do que para as anteriores se daria por conta do menor espaço articulatorio na parte posterior da cavidade oral, se comparada à parte anterior. Isto é, uma vez que o espaço é menor, as vogais médias e altas tendem a ser articuladas mais próximas umas das outras e a alternância média-alta/alta ocorre mais facilmente com /o/ do que com /e/.

A este respeito, se o menor espaço articulatorio é responsável por uma melhor distribuição entre as formas média-alta e alta na pauta posterior, em dialetos que admitem as formas médias-baixas na posição, a alternância em favor da forma mais aberta também deveria ser maior para a pauta posterior. Em outras palavras, se o menor espaço articulatorio influencia a alternância entre as formas, deve-se esperar mais casos tanto de [u] quanto de [ɔ] se comparados a [i] e a [ɛ], respectivamente.

Entretanto, a Tabela 12 abaixo, que esquematiza a frequência de ocorrência entre as formas em ambos os dialetos, mostra o contrário. Além de [u] ter emergido mais vezes do que [i] somente em São Paulo, no dialeto de São Luís, que também admite as formas médias-abertas na postônica não-final, [ɛ] foi mais recorrente que [ɔ].

Tabela 12: Disposição geral das frequências de ocorrência.

	[i]	[u]	[e]	[o]	[ɛ]	[ɔ]
SP	12,6%	16,3%	87,4%	83,7%	–	–
SL	9%	8,4%	64%	70,4%	27%	21,2%

Além disso, fazer uso de critérios exclusivamente articulatorios para avaliar um subsistema vocálico, bem como regras de neutralização, não parece ser apropriado, tendo em vista o caráter fonológico de ambos. No que diz respeito à fonologia, não há nada que esteja correlacionado, por si só, à posterioridade/anterioridade (ou ainda entre labialidade, coronalidade e dorsalidade, na perspectiva da Geometria de Traços adotada neste estudo) e o grau de abertura de uma vogal.

Finalmente, a partir dos resultados apresentados na Tabela 12, refutamos a hipótese de que vogais posteriores na postônica não-final tendem a emergir mais como vogal alta do que as anteriores, tendo em vista que a diferença em São Paulo é muito reduzida e, em São Luís, não chega a ser observada.

5.2 A emergência das formas altas em São Paulo e São Luís

Como vimos anteriormente, os resultados encontrados apontaram que, para o dialeto de São Paulo, a átona final (coronal), o contexto seguinte (coronal) e o contexto precedente

(coronal) apresentaram correlação à emergência de [i]; já com base nos dados de São Luís, se correlacionaram à forma alta da vogal anterior a átona final (coronal), o contexto seguinte (coronal), a vogal tônica (coronal) e o contexto precedente (coronal). Como podemos observar, todos os segmentos que apresentaram correlação à emergência de [i] possuem ponto de articulação coronal, o mesmo da vogal anterior alta. A única diferença observada entre os dialetos foi o fato de a variável *Ponto de Articulação da Vogal Tônica* ter apresentado correlação em São Luís, mas não em São Paulo. O fator selecionado, entretanto, da mesma forma que nas outras variáveis, foi o coronal.

Os resultados da estatística para os dados de São Paulo também mostraram que a vogal átona final (labial), o contexto precedente (labial) e a vogal tônica (labial) se correlacionaram à emergência de [u]; já com base nos dados de São Luís, a átona final (labial) e o contexto fonológico seguinte (labial) se correlacionaram à forma alta da vogal posterior. Dessa forma, vemos que somente a vogal labial quando na átona final apresentou correlação à vogal [u] em ambos os dialetos.

Quando observamos resultados estatísticos para que, a partir deles, possamos entender a motivação de um fenômeno fonológico, buscamos regularidade, isto é, no caso desta análise, que a mesma tendência para uma pauta também seja observada na outra e em ambos os dialetos (mesmo que a proporção de aplicação pudesse variar de acordo com razões fonéticas ou de outras ordens). Desse modo, poderíamos dizer que a emergência da forma alta da vogal anterior e da posterior estão correlacionadas a segmentos adjacentes quando estes possuem o mesmo ponto de articulação que o da postônica não-final – especialmente a átona final, variável que apresentou correlação todas as vezes, para ambas as vogais e com base nos dados dos dois dialetos investigados. Já a vogal dorsal, quando átona final, foi o fator que mais se correlacionou à não emergência da forma alta de ambas as vogais e em ambos os dialetos.

Entretanto, como ressaltado anteriormente, não há nada a respeito do ponto de articulação de um segmento que, para o modelo adotado, por si só possa influenciar o grau de abertura de uma vogal. Dessa forma, decidimos observar mais atentamente o grau de altura tanto da vogal tônica quanto da átona final e a possível correlação com a emergência das formas das vogais médias.

Ao observar os dados, a não correlação entre o grau de altura da postônica não-final e a *Altura da Vogal Tônica* ficou ainda mais evidente. Isto é, os resultados mostraram que as vogais altas não são as que por mais vezes estão na tônica quando a vogal postônica não-final posterior

emerge como alta e nem são as vogais baixas que estão na posição de acento quando as vogais de ambas as pautas emergem como média-baixa, o que seria esperado para a influência do grau de altura da vogal. Aliado a isso, há o fato de que uma influência deste tipo (ou seja, uma possível regra de assimilação de altura entre a postônica não-final e a tônica) ir contra a tendência do Português, que não apresenta outras regras progressivas na língua.

A respeito da átona final, é necessário que ressaltemos que a posição admite três tipos de segmentos com relação ao ponto de articulação (coronal /e/, dorsal /a/ e labial /o/) e dois tipos de segmentos com relação à altura (altas /i,u/ e baixa /a/). Tendo em vista que segmentos coronais e labiais na átona-final necessariamente são altos, é relevante saber se o principal fator para a emergência das formas altas é a altura de uma vogal qualquer (ou seja, com qualquer ponto de articulação), ou a altura de uma vogal específica (isto é, a altura de uma vogal que tem um ponto de articulação específico). O fato de o teste feito pelo Varbrul ter selecionado especificamente as vogais labiais na átona final para a emergência de [ʊ] e a vogal coronal na átona final para a emergência de [ɪ] é um indício que a correlação não é com uma vogal alta qualquer, e sim com uma vogal alta que compartilha o mesmo ponto de articulação que a vogal postônica não-final, caso contrário, esperaríamos que segmentos labiais (que são altos na átona-final) também apresentassem correlação para a emergência de [ɪ] e que segmentos coronais (que são altos na átona-final) também apresentassem correlação para a emergência de [ʊ], o que não ocorre. Outro fato que aponta para esta direção é que a altura da vogal não foi selecionada pela regressão logística realizada pelo Varbrul. Isso ocorre porque na variável *Altura da Átona Final*, as altas incluíam tanto as vogais coronais quanto as labiais. Assim, mesmo que as altas coronais sempre favoreçam o alçamento, o fato de as altas labiais estarem também sendo computadas no mesmo fator – e de estas não favorecerem o alçamento – faz com que a soma de altas coronais e labiais não seja correlacionável à emergência das vogais altas em postônica não final. Porém, a própria análise de Ponto de Articulação da Átona Final faz esta separação, como vimos acima.

Como mencionado anteriormente, as variáveis *Ponto de Articulação da Átona final* e *Altura da Átona Final* foram as que apresentaram correlação à emergência da vogal alta de forma mais constante, isto é, do mesmo modo, para ambas as pautas e em ambos os dialetos. Entretanto, não foram as únicas variáveis que se correlacionaram à emergência das vogais [ɪ] e [ʊ]. É necessário, portanto, que se analise as demais variáveis que, segundo os resultados obtidos, apresentaram algum tipo de correlação, a saber, o *Ponto de Articulação do Contexto*

Fonológico Precedente, o *Ponto de Articulação do Contexto Fonológico Seguinte* e o *Ponto de Articulação da Vogal Tônica*. Apresentamos, resumidamente na Tabela 13 a seguir, uma disposição dos resultados de correlação para estas variáveis.

Tabela 13: Resultados das variáveis fonológicas.

	Contexto Precedente	Contexto Seguinte	Vogal Tônica
[i] SP	Coronal (0.74)	Coronal (0.65)	–
[i] SL	Coronal (0.88)	Coronal (0.75)	Coronal (0.66)
[u] SP	Labial (0.58)	–	Labial (0.57)
[u] SL	–	Labial (0.8)	–

Como podemos ver, as coronais no contexto precedente e no contexto seguinte estão correlacionadas à emergência de [i] tanto em São Paulo, quanto em São Luís. A vogal tônica coronal, entretanto, só está correlacionada à emergência de [i] em São Paulo, mas não em São Luís. Com relação à posterior, o contexto precedente labial e a vogal tônica labial estão correlacionados à emergência de [u] em São Paulo, mas não em São Luís; já o contexto seguinte, quando labial, se correlaciona à emergência de [u] em São Luís, mas não em São Paulo.

Dessa forma, observamos que, diferentemente do que ocorre para a correlação do Ponto de Articulação da Átona Final e a Altura da Átona Final, que apresentam a mesma correlação para ambas as vogais altas e em ambos os dialetos, a estabilidade de correlação para a vogal tônica e para os contextos adjacentes não é a mesma, variando entre os dialetos. A exceção fica por conta dos segmentos coronais para a emergência de [i] no contexto precedente e no seguinte, que é classificado como relevante em ambos os dialetos e com peso elevado. Por outro lado, como ressaltado anteriormente, não há nada no Ponto de Articulação de um segmento adjacente que possa influenciar o grau de altura de uma vogal. Assim, decidimos averiguar diretamente nos dados a motivação para a possível correlação indicada pelos resultados do Varbrul. Apresenta-se, na Tabela 14, palavras que ilustram os contextos fonológicos que se mostraram correlacionáveis.

Tabela 14: Exemplificação dos resultados.

	Contexto Precedente	Contexto Seguinte	Vogal Tônica
[i]	Piogênese Aborígene <i>Erógeno</i> <i>Indígena</i>	Hóspede Intérprete <i>Nêspera</i> <i>Láguecha</i>	Aférese Íngreme <i>Conífera</i> <i>Trêfego</i>
[o]	Xenófobo Horóscopo <i>Necrópole</i> <i>Tômbola</i>	Autódromo Tecnófobo <i>Agrônoma</i> <i>Síncope</i>	Autódromo Horóscopo <i>Códope</i> <i>Úpobe</i>

Em cada célula da Tabela 14 observamos duas palavras grafadas normalmente e duas palavras em itálico. As palavras em itálico são as que apresentam a configuração que se mostrou correlacionável à emergência da vogal alta (um segmento coronal no contexto precedente, por exemplo), mas que possuem baixa (ou nenhuma) ocorrência de emergência da referida forma; já as palavras grafadas normalmente exibem a configuração que apresentou correlação com a vogal alta e que apresentaram elevado grau de emergência da referida forma das vogais.

O que é notamos é que as palavras nas quais mais emergiram vogais altas, na verdade, também apresentam o contexto para a aplicação da regra de assimilação do Nó Vocálico, ou seja, uma vogal alta na átona final que compartilha o mesmo Ponto de Articulação que o da vogal média na postônica não-final, como em ‘Piogênese’ e ‘Xenófobo’. Todas as palavras em itálico, isto é, todas as palavras que apresentam a configuração que apresentou correlação segundo os resultados da estatística, mas que não apresentaram elevado grau de emergência das vogais altas, não possuem o contexto para a aplicação da regra, como em ‘Indígena’ e ‘Necrópole’. Acreditamos que este fato é um indício de que a correlação das variáveis *Ponto de Articulação da Tônica*, *Ponto de Articulação do Contexto Fonológico Precedente* e *Ponto de Articulação do Contexto Fonológico Seguinte*, na verdade, está associada às variáveis que, de fato, são as mais constantes e formalizam a regra – *A altura da Átona Final* e o *Ponto de Articulação da Átona Final*.

5.3 A emergência das médias-baixas em São Luís

A respeito da emergência das médias-baixas em São Luís, os resultados da estatística mostraram que as variáveis que se correlacionaram à emergência de [ɛ] foram as labiais e dorsais em contexto precedente e a vogal dorsal na átona final. Em ambas as posições as vogais coronais são inibidoras da forma. Já para [ɔ], as dorsais na átona final mostraram correlação. A

fim de tornar a leitura mais simples, rerepresentamos tais resultados na Tabela 15 abaixo, que também mostra que segmentos coronais e dorsais na átona-final inibem a forma média-baixa da vogal posterior.

Tabela 15: Resultados para as vogais médias-abertas.

Vogal	Variável	Favorece	Neutro	Desfavorece
[ɛ]	Contexto Precedente	Labial (0.67) Dorsal (0.62)	–	Coronal (0.23)
	Átona final	Dorsal (0.65)	Labial (0.50)	Coronal (0.33)
[ɔ]	Átona final	Dorsal (0.68)	–	Coronal (0.44) Labial (0.23)

A respeito dos resultados obtidos quando selecionamos as médias-baixas posterior e anterior como variáveis dependentes, o *Ponto de Articulação da Átona Final* foi o fator que apresentou correlação para ambas as vogais. Entretanto, diferentemente do que vimos para a emergência das vogais altas, para as quais a correlação se deu com a vogal que partilhava o mesmo ponto de articulação com a postônica não-final; para as médias-baixas, o ponto de articulação que se correlacionou à emergência de [ɛ] não foi o coronal e o que se correlacionou a emergência de [ɔ] não foi o labial, mas sim o dorsal para ambas as vogais.

Interpretar esse resultado como evidência para a correlação das vogais médias-baixas com o ponto de articulação dorsal seria equivocado, uma vez que um traço de ponto estaria licenciando uma regra de altura, e mais estranho ainda seria o fato de o traço em questão não ser compartilhado pelas vogais alvo da regra. Por isso, acreditamos que a correlação não se deu com /a/ na átona final pelo fato de a vogal ser dorsal, mas sim pelo fato de ela ser baixa.

Entretanto, é necessário que entendamos, então, o porquê de a correlação ter se dado com o *Ponto de Articulação da Átona Final* e não com a variável *Altura da Átona final*, caso o licenciador da regra seja, de fato, um traço de altura. Como ressaltamos anteriormente ao tratar das emergências das vogais altas, a premissa de uma Regressão Logística, principal técnica utilizado pelo Varbrul para as análises, é a de que uma variável que se correlaciona a determinada forma, necessariamente, seja inibidora da forma concorrente. Por conta de termos agrupado em um mesmo fator “Alto” vogais altas coronais e labiais e a correlação só existir entre aquelas que possuem o mesmo ponto de articulação que o da postônica não-final, a

correlação não pôde ser atestada para as altas e, conseqüentemente, também não pôde ser atestada para as médias-baixas.

Um indício de que esta interpretação dos resultados é a correta é o fato de o teste de Qui-quadrado, cujos resultados foram apresentados no tópico 4 deste artigo, ter indicado associação entre a *Altura da átona final* e a altura das vogais médias na postônica não-final. Isto é, como vimos na Tabela 11, [ɛ] e [ɔ] emergem mais vezes quando na átona-final há uma vogal baixa (ou seja, a vogal /a/), o que é condizente com o que se esperava e com nossa interpretação dos resultados obtidos com o Varbrul.

A correlação, todavia, não se deu apenas com a vogal átona final. Os resultados obtidos também apontaram correlação entre a emergência de [ɛ] a segmentos dorsais e labiais no contexto precedente, isto é, consoantes dorsais e labiais no ataque da sílaba postônica não-final. Uma vez que, fonologicamente para este caso, não parece haver nenhuma motivação aparente para a alteração da altura de uma vogal por conta de uma consoante que a precede, decidimos fazer uma análise direta dos dados, tal qual fizemos anteriormente para analisar a correlação de consoantes adjacentes à emergência das vogais altas.

Tabela 16: Palavras com segmentos labiais e dorsais precedentes a [ɛ].

Emergência da média-baixa	Ponto de Articulação do Contexto precedente	
	Labial	Dorsal
+ [ɛ]	Câmera Ômega	Párrega Váquega
- [ɛ]	Áspero Tráfego	Íngreme Tíquete

Na Tabela 16, as palavras são apresentadas em duas categorias: aquelas que tiveram alto nível de emergência da média-baixa, representado por [+ɛ], e as que tiveram baixo nível de emergência da média-baixa, representado por [-ɛ]. Dispostas dessa forma, percebe-se que as palavras em que mais emergiram à média-baixa são as que possuem uma vogal /a/ na átona-final, como em ‘Câmera’ e ‘Ômega’. Por outro lado, as palavras que possuem a mesma configuração (segmentos labiais e coronais no contexto precedente), mas que não possuem a vogal dorsal na átona-final, quase não apresentaram casos de emergência de [ɛ], como em ‘Áspero’ e ‘Tráfego’. Tal fato é um indício de que a correlação da variável *Ponto de Articulação do Contexto Fonológico Precedente*, na verdade, está associada à variável que, de fato, parece estar correlacionada à emergência das médias-baixas, isto é, a *Átona Final*.

Em suma, acreditamos que a variável *Altura da Átona Final* não se mostrou correlacionável, no teste rodado no Varburl, à emergência das altas porque ambas as vogais, labial e coronal, foram agrupadas em um mesmo fator, o *alto* e, por consequência, tendo em vista que a análise é binária e que a correlação à emergência de uma forma implica necessariamente na correlação contrária para a emergência da forma concorrente, a não-correlação para as altas impedem que se observe a associação entre a emergência das médias-baixas com a vogal baixa na átona-final, o que ocorreu, como vimos detalhadamente na análise dos dados.⁵ Além disso, mostramos que os dados utilizados nesta pesquisa, bem como os resultados obtidos com ambos os testes estatísticos, corroboram a hipótese inicial de que a emergência das formas altas e médias-baixas na postônica não-final para ambas as pautas é correlacionado fonologicamente, especificamente por uma vogal alta na posição átona final que compartilha o mesmo ponto de articulação que a vogal média na postônica não-final e por uma vogal baixa na átona final.

5.4 A formalização das regras

No início deste artigo, apresentamos duas propostas para a regra de neutralização, a que chamaremos de clássica e remetemos a Clements (1985) e a de Wetzels (2011). A noção clássica é a de que a neutralização é uma regra que apaga traços responsáveis por distinções em um dado contexto. Exemplificaremos na Figura 2 abaixo a configuração das vogais do PB na átona final.

Nó de abertura	i/u	e/o	ɛ/ɔ	a
[aberto1]	-	-	-	+

Figura 1: Átona final do PB assumindo Clements (1985).

Como podemos ver, apenas [aberto1] aparece na configuração, uma vez que [aberto2] e [aberto3] foram neutralizados (i.e., apagados), e, uma vez que todas as vogais médias possuem

⁵ Alguns fatos, no entanto, ficaram por ser explicados. As únicas palavras que possuem alto grau de frequência de médias-baixas na postônica e não apresentam um padrão específico de vogal baixa na átona-final são Célebre (0% [ɪ]; 30% [e]; 70% [ɛ]), Diâmetro (0% [ɪ]; 30% [e]; 70% [ɛ]), Cônego (0% [ɪ]; 35% [e]; 65% [ɛ]), Córrego (0% [ɪ]; 40% [e]; 60% [ɛ]) e Trólebus (0% [ɪ]; 35% [e]; 65% [ɛ]). Todas elas, como se pode ver, com a vogal coronal na postônica não-final. Não foram encontradas palavras com alta frequência de emergência de médias-baixas posterior e que não tivessem a vogal /a/ na átona-final.

o mesmo valor para o traço de abertura, a distinção entre elas é perdida a favor da vogal alta, que emerge com mais frequência.

Já assumido a noção de neutralização proposta por Wetzels (2011), de que se trata de um mecanismo pelo qual muda-se o valor do traço na camada em que a distinção é garantida, os traços não são apagados, e sim permanecem na configuração com um valor invertido. Nesses termos, a configuração da átona final seria como podemos ver abaixo.

Nó de abertura	i/u	e/o	ɛ/ɔ	a
[aberto1]	-	-	-	+
[aberto2]	-	-	-	+
[aberto3]	-	-	-	+

Figura 2: Átona final do PB assumindo Wetzels (2011).

Ambas as propostas dão conta do fato de, na átona final, a distinção entre as médias ser perdida, mas de formas diferentes. Na primeira, fazendo uso de apenas um traço de abertura, enquanto que a segunda mantém todos os traços de abertura necessários para distinguir as vogais do PB. Entretanto, apenas a proposta de Wetzels (2011) é capaz de capturar a correlação entre as postônicas finais e não-finais apresentadas neste estudo.

Parece ser evidente que a correlação entre as vogais átonas postônicas é resultado de uma regra de assimilação de traço – de altura e de ponto para a emergência das altas, e apenas de altura para a emergência das médias-baixas.

Assumindo que para que uma vogal seja produzida como alta na postônica não-final é necessário que haja [-aberto2] em sua configuração, não haveria como [-aberto2] ser espreado para a vogal postônica não-final, tal qual ocorre da tônica para a pretônica assumindo a proposta clássica, tendo em vista que para esta visão este traço não está mais na configuração. Ou seja, o processo de alçamento da postônica não-final não poderia ser entendido como uma regra de assimilação de traço da átona final. Assumindo a proposta de Wetzels (2011), por outro lado, o traço [-aberto2] existente em uma vogal alta na átona final e poderia ser espreado para a postônica não-final. O mesmo para a produção de uma vogal média-aberta, que na posição exigira em sua configuração [+aberto3], que também não mais existiria na átona-final, segundo a proposta clássica.

Entretanto, assumir que a emergência das formas altas na postônica não-final são consequência de um regra de espreado de [-aberto2] não conseguiria explicar o porquê de

os resultados estatísticos terem sempre apresentado, invariavelmente e para ambos os dialetos, correlação especificamente com o ponto de articulação para a emergência de [ɪ] e de [ʊ].

Na análise clássica a respeito da harmonia vocálica do PB, Bisol (1981) observa que a vogal anterior, na pretônica, tende a alçar mais vezes quando na tônica há uma vogal alta e com o mesmo ponto de articulação, isto é, um [i]. Battisti (1993), a respeito da vogal /o/, observou que o alçamento ocorre mais vezes quando há no contexto precedente e no contexto seguinte uma consoante labial, havendo uma vogal alta contígua à pretônica.⁶ Tais fatos do PB mostram que regras que afetam a altura das vogais da língua tendem a ser mais frequentes quando há ação conjugada de causas. Como se pode ver, há respaldo em outros fatos do Português para que não se desconsidere a ação conjunta de fatores para a aplicação de uma regra fonológica, especificamente, uma regra que afete a altura de uma vogal.

Seguindo este raciocínio, a melhor forma de capturar a tendência observada nos dados desta pesquisa de que as vogais médias-baixas estão correlacionadas à presença de uma vogal baixa na átona final é por meio de uma regra de assimilação de [+aberto3], e de que a emergência das formas altas das vogais médias está correlacionada principalmente à altura e ao ponto de articulação da átona-final é por meio de uma regra de associação de Nó Vocálico, tendo em vista que o constituinte domina tanto o nó de abertura como o nó de ponto de articulação.

Tal hipótese consegue capturar, também, o fato de os resultados estatísticos terem apontado que o principal fator correlacionável ao desfavorecimento das formas altas de ambas as vogais é a vogal dorsal na átona final. Isso porque como se sabe a vogal /a/ na átona final é caracterizada como [dorsal, +aberto1, +aberto2, +aberto3], ou seja, além de não compartilhar o ponto de articulação com nenhuma das duas vogais médias, também não é [-aberto2] e [-aberto3], por ser uma vogal baixa.

As regras podem ser esquematizadas como se vê na Figura 4, que exemplificam a emergência de [u] e de [ɔ] na postônica não-final, flechas azul e vermelha, respectivamente.

⁶ É importante ressaltar que estes não são as únicas variáveis correlacionadas ao alçamento das médias na pretônica. Para a análise completa, ver Bisol (1981) e Battisti (1993).

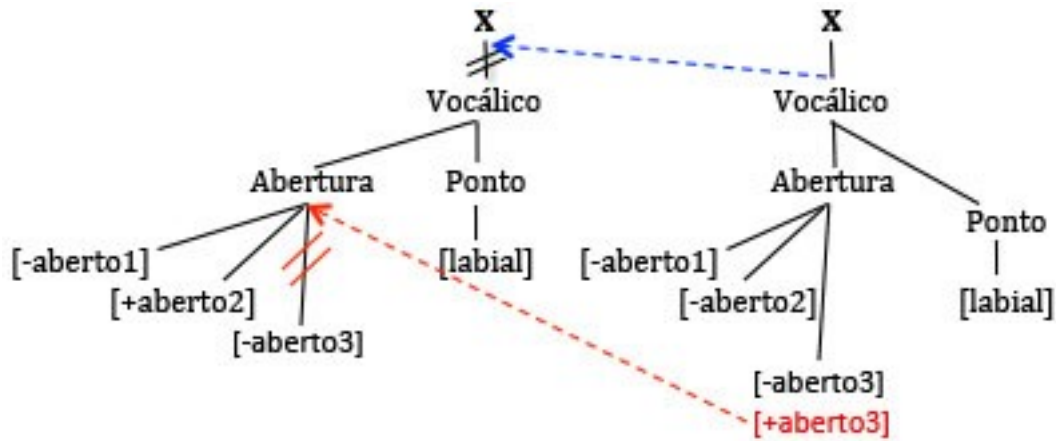


Figura 3: formalização das regras.

Wetzels (1992) propõe que o domínio da regra de neutralização da postônica não-final seja o pé métrico, tendo em vista que as vogais nesta posição estão sempre na parte fraca de um pé, ou na posição forte de um pé degenerado. Uma vez que as regras aqui propostas são de assimilação dos traços da átona final pela postônica não-final, ambas as vogais devem estar no mesmo domínio. Dessa forma, o domínio de aplicação das regras aqui propostas deve ser a palavra fonológica – obviamente, restrito a palavras de acento antepenúltimo para que o contexto de aplicação, i.e., as postônicas não-finais, possa existir.

6. Conclusão

O objetivo deste estudo foi analisar as vogais médias em posição postônica não-final para explicar em termos formais a alternância entre vogais médias baixas (para dialetos nordestinos) médias-altas e vogais altas, o que até então não havia sido feito. Para tanto, decidimos adotar uma metodologia diferente daquelas que haviam sido utilizadas pelos estudos anteriores a fim de obter maior número de dados e maior número de combinações de contextos fonológicos, o que conseguimos por meio de um experimento controlado.

Os resultados mostraram a correlação de diversas variáveis à emergência das formas médias-baixas e altas na referida posição, entretanto, ao observar os dados detalhadamente, pudemos observar que tais correlações, na verdade, reproduziam uma tendência geral: a associação da postônica não-final à altura da átona final (no caso das médias-baixas em São Luís) e à altura e ao ponto de articulação (no caso das vogais altas para ambas as localidades). Tal associação só pode ser capturada se assumirmos a noção de neutralização proposta por

Wetzels (2011), no qual os três traços de abertura são mantidos em todas as posições átonas e somente os valores adotados por eles é que mudam.

A partir daí foi possível formalizar duas regras: a que tem por resultado as vogais altas na postônica não-final acontece por meio de um espriamento do nó vocálico da átona final para a postônica não-final; e a que tem por resultado as médias-baixas acontece por meio de um espriamento de [+aberto3] da átona final para a postônica não-final.

Propomos, ainda, que somente é possível capturar os fatos observados em ambos os dialetos assumindo um subsistema postônico não-final de cinco vogais – do qual as médias-altas fazem parte. Dessa forma, não há uma lógica para a emergência das vogais médias como médias-altas porque não se trata do resultado de uma regra fonológica, mas sim da realização fonética de formas existentes na subjacência.

Em suma, assumindo Wetzels (2011), conseguimos explicar a emergência de vogais média-baixas, média-altas e altas em São Luís e também a emergência de vogais média-altas e altas em São Paulo em termos de regras sendo aplicadas nestes dialetos. Como vimos também, em ambos os dialetos, tratam-se de regras de aplicação opcional (o que aponta que outras variáveis, talvez extralinguísticas, podem também estar correlacionadas à aplicação da regra). A principal contribuição deste artigo se dá por mostrar como linguisticamente é possível gerar as vogais média-baixas a partir de um mesmo conjunto de vogais para os dois dialetos.

Referências

BISOL, L. **Harmonia vocálica**: uma regra variável. 1981. Tese (Doutorado em Linguística) – Faculdade de Letras. Universidade Federal do Rio de Janeiro, Rio de Janeiro

_____. A neutralização das átonas. **Revista Letras** (Curitiba), Curitiba - Paraná, v. 61, p. 273-283, 2003.

CÂMARA JR., Joaquim Matoso. **Estrutura da língua portuguesa**. Petrópolis: Vozes, 1970.

BOERSMA, P., and Weenink, D. 2013. **Praat**: doing phonetics by computer Version MAC OS 10.7. Disponível em: <http://www.praat.org/>

CLEMENTS, N. Vowel height assimilation in Bantu languages. In: K. HUB-BARD (Ed.) **BLS 17S: Proceedings of the Special Session on African Languages Structures**: 25-64. Berkeley Linguistic Society, 1991. **crossref** <http://dx.doi.org/10.3765/bls.v17i2.1662>

_____; HUME, E. The internal organization of speech sounds. In: GOLDSMITH, J. **The handbook of phonological theory**. Oxford: Blackwell, 1995, p. 245-306

CRISTÓFARO-SILVA, T.; ALMEIDA, L. S.; OLIVEIRA-GUIMARAES, D. M. L.; MARTINS, R. M. F. **Corpus do e-Labore** (Laboratório Eletrônico de Oralidade e Escrita). Disponibilizado online em: www.projetoaspa.org/elabore. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais. 2009

ESCUDEIRO, P; BOERSMA, P. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. **Journal of the Acoustics Society of America**. v. 126, p. 1379-1393, 2009. **crossref** <http://dx.doi.org/10.1121/1.3180321>

RAMOS, A. P. **Descrição das vogais postônicas não-finais na variedade do noroeste paulista**. São José do Rio Preto, 2009. 175 f. Dissertação (Mestrado em Linguística), Programa de pós-graduação em Estudos Linguísticos, Universidade Estadual Paulista “Júlio de Mesquita Filho”.

RIBEIRO, D. F. S. **Alçamento das vogais postônicas não-finais no português de Belo Horizonte: uma abordagem difusionista**. 2007. 275 f. Dissertação de mestrado – Universidade de Federal de Minas Gerais, Belo Horizonte.

SANTOS, A. P. **Vogais médias postônicas não-finais no falar do Estado do Rio de Janeiro**. 162f. Rio de Janeiro, 2010. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro.

SANTANA, A. **Análise das Postônicas não-finais em São Paulo e São Luís**. 152f. São Paulo, 2015. Dissertação (Mestrado em Linguística). Programa de Pós-graduação em Semiótica e Linguística Geral, Universidade de São Paulo.

SILVA, A P. da. **Vogais postônicas não-finais: do sistema ao uso**. 216f. João Pessoa, 2010. Tese (Doutorado em Linguística) Programa de Pós-Graduação em Linguística, Universidade Federal da Paraíba.

VIEIRA, M. J. B. As Vogais Médias Postônicas: uma análise variacionista. In: BISOL, L.; BRESCANCINI, C. (orgs) **Fonologia e Variação: Recortes do Português Brasileiro**. Porto Alegre: EDIPUCRS, 2002, p. 127-159.

WETZELS, W. L.. Harmonização vocálica, truncamento, abaixamento e neutralização no sistema verbal do português: uma análise auto-segmental. In.: **Cadernos de Estudos Linguísticos**. Campinas: UNICAMP, 1991, 25-58.

_____. Mid-vowel neutralization in Brazilian Portuguese. **CEL** 23, p. 19-55, 1992.

_____. Mid Vowel Alternations in the Brazilian Portuguese Verb, in: **Phonology** 12. pp.281-304, 1995. **crossref** <http://dx.doi.org/10.1017/S0952675700002505>

_____. The representation of Vowel height neutralization in Brazilian Portuguese (Southern Dialects). In: GOLDSMITH, J.; HUME, E.; WETZELS, L. (eds.), **Tones and Features**. Berlin. Walter De Gruyter, 2011, p. 331 – 359. **crossref** <http://dx.doi.org/10.1515/9783110246223.331>

Artigo recebido em: 26.10.2015

Artigo aprovado em: 25.04.2016

A estrutura silábica em esperanto

The syllabic structure in Esperanto

Karina Gonçalves de Souza de Oliveira*

RESUMO: Este artigo tem como objetivo comentar a bibliografia disponível sobre a estrutura das sílabas em esperanto, baseada em conceitos da teoria fonológica sobre as possíveis estruturas silábicas existentes nas mais variadas línguas. Primeiramente veremos o que pode-se dizer sobre a sílaba, sua estruturação e seus constituintes, para após verificar o que já foi escrito em esperanto e sobre o esperanto em relação ao assunto. Por fim, formulamos regras estruturais de acordo com a discussão feita ao longo do trabalho. Há poucos estudos relacionados a essa temática sobre o esperanto. Nossas principais fontes são artigos de Oostendorp (1999) e Bavant (2006), que fez uma compilação de vários livros que citam a noção de "sílaba" e tentam defini-la, além de ter analisado, por meio de um programa computacional, todas as palavras (mais de 47 mil) que compõem o *PIV – Plena Ilustrita Vortaro* (Dicionário Completo ilustrado), para achar os constituintes possíveis da sílaba e suas frequências na língua.

ABSTRACT: This paper has the purpose of commenting on the available literature about the structure of syllables in Esperanto, based on concepts of the phonological theory about the possible syllabic structuring in a wide range of languages. Firstly we will verify what can be said about the syllable, its structure and its constituents, in order to, afterwards, examine what has been written in Esperanto and about Esperanto on the subject. Finally, we formulate structural rules in accordance with the discussion in the paper. There are only a few studies related to the topic about Esperanto. Our main sources are articles by Oostendorp (1999) and Bavant (2006), who made a compilation of several books mentioning the notion of "syllable" and who try to define it, besides having analyzed, using a computer program, all the words (more than 47,000) that compose the *PIV – Plena Ilustrita Vortaro* (Complete Illustrated Dictionary), to find the possible constituents of the syllable and their frequency in the language.

PALAVRAS-CHAVE: Fonologia. Sílaba. Esperanto.

KEYWORDS: Phonology. Syllable. Esperanto.

1. Introdução

Línguas planejadas existem em grande quantidade, e o motivo pelos quais são planejadas varia bastante. Algumas têm o intuito de serem línguas auxiliares na comunicação internacional; outras, de serem parte de um mundo artístico (filmes, seriados, etc.); outras, ainda, são criadas simplesmente por diversão. O esperanto, planejado com o intuito de servir para língua auxiliar internacional, é a que mais teve sucesso, a que formou a maior comunidade

* Mestranda do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas, da Universidade de São Paulo (FFLCH-USP).

linguística, que usa a língua de forma corrente (ECO, 1996, p. 302-303), e tem, inclusive, muitos falantes nativos (FRAWLEY, 2003). Os falantes de esperanto estão espalhados pelo mundo, literalmente, o que torna o estudo da língua um tanto quanto difícil, no que tange à variação fonética ou lexical, por exemplo. Entretanto, podemos estudar sua estrutura fonológica baseada nos postulados escritos por Zamenhof, seu criador, e por estudos posteriores feitos por acadêmicos.

O presente artigo discute a estrutura silábica da língua, baseando-se em estudos teóricos gerais sobre a sílaba e estudos específicos sobre o esperanto.

2. O que é sílaba?

Blevins (1995) afirma que várias escolas de teorias fonológicas reconhecem a sílaba como item fundamental para a análise. Para ela, a sílaba pode ser entendida como as unidades estruturais que organizam melodicamente a cadeia sonora da fala, que leva em consideração o grau de sonoridade dos segmentos fonológicos para se organizar. Em resumo, “a sílaba então é a unidade fonológica que organiza as melodias segmentais no que diz respeito à sonoridade; os segmentos silábicos são equivalentes aos picos de sonoridade dentro dessas unidades organizacionais” (BLEVINS, 1995, p. 207).

Para sustentar a afirmação acima, Blevins apresenta quatro argumentos de que a sílaba é um constituinte fonológico. São eles:

→ sílaba como domínio: há certas restrições e/ou processos fonológicos que usam a sílaba como unidade de aplicação. Esses processos acontecem em um ambiente que contém um pico de sonoridade, maior que o segmento e menor que a palavra. Exemplos desses processos são a faringalização em dialetos do árabe e do berbere e processos que envolvem o acento e o tom.

→ fronteira de sílaba como lugar: as sílabas correspondem às barreiras de palavras em todas as línguas, e sem essa noção deveriam ser formuladas muitas regras para explicar fenômenos que acontecem entre as palavras.

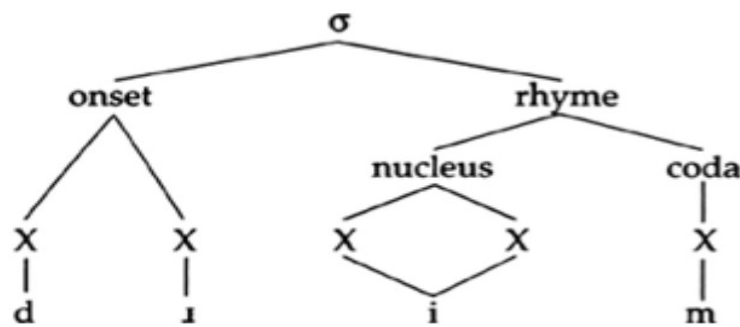
→ sílabas como estruturas-alvo: a noção de sílaba é usada em jogos linguísticos ou alvos prosódicos em processos morfológicos, como a reduplicação.

→ intuição de falantes nativos: os falantes das línguas normalmente tem uma intuição bastante clara que quantas sílabas há nas palavras, e onde estão suas divisões internas. (BLEVINS, 1995, p. 207-210)

Para continuar a discussão sobre os constituintes da sílaba, a autora fala sobre a noção de sonoridade, afirmando que as sílabas são formadas baseadas em algo que tem picos sonoros.

De forma bastante geral, a escala de sonoridade é a seguinte: vocoides > líquidas > nasais > obstruintes, que podem ainda ser divididas em várias subclassificações. Para a construção de sílabas válidas segundo essa escala, deve-se verificar que a sonoridade aumenta até o pico (núcleo) da sílaba e depois decresce.

Várias propostas foram feitas sobre a estrutura interna da sílaba, e a mais usada delas, a qual também usaremos para este trabalho, é a de braço binário com rima. Nesta estruturação, a sílaba se divide em ataque e rima, e esta, por sua vez, se divide em núcleo e coda (BLEVINS, 1995, p. 212), como pode ser visto na estruturação abaixo.



(BLEVINS, 1995, p. 216)

Uma sílaba, portanto, se constitui do núcleo e de suas margens. Zec (2007) chama a atenção para o fato de que “os segmentos que tipicamente ocorrem no núcleo são representados por V, e aqueles tipicamente nas margens por C. V não necessariamente se refere a uma 'vogal'. Em algumas línguas, a posição V também pode ser ocupada por uma consoante [...]” (p. 163). A autora ainda nos traz uma tipologia básica da estrutura da sílaba: CVC, sílaba com as três principais partes; CV, sílaba com ataque e núcleo; VC, sílaba com núcleo e coda e V, sílaba com apenas o núcleo (ZEC, 2007, 163), podendo ainda os ataques e codas serem simples, com apenas um constituinte, como nos exemplos dados, ou complexos, com dois segmentos ocupando tal posição. Os constituintes da sílaba não são interdependentes, “se uma língua requer ataques, isso não significa que bane ou requeira codas, e vice-versa” (ZEC, 2007, p. 164). Na tabela a seguir pode-se observar os possíveis formatos de sílabas existentes nas línguas:

Tabela 1: Tipos estruturais possíveis de formação silábica.

ataque	coda	ataque complexo	coda complexa	inventário	língua
obrigatório	opcional	opcional	opcional	(C)CV(C)(C)	totonaca
			proibida	(C)CV(C)	dacota
		proibida	opcional	CV(C)(C)	klamath
			proibida	CV(C)	temiar
	proibida	opcional	-	(C)CV	arabela
		proibida	-	CV	senufo
opcional	opcional	opcional	opcional	(C)(C)V(C)(C)	inglês
			proibida	(C)(C)V(C)	espanhol
		proibida	opcional	(C)V(C)(C)	finlandês
			proibida	(C)V(C)	turco
	proibida	opcional	-	(C)(C)V	pirahã
		proibida	-	(C)V	fiji

Adaptado de Zec (2007, p. 165).

Mas como saber, em uma sequência VCCV, se os segmentos mediais CC fazem parte de uma coda complexa da primeira sílaba (VCC.V), de um ataque complexo da segunda sílaba (V.CCV), ou ainda se de uma coda simples da primeira sílaba e um ataque simples da segunda sílaba (VC.CV)? Zec (2007) explica que os constituintes têm uma ordem para serem incorporados à sílaba, e que “a maximização do ataque é capturada pela regra de ordenação: a regra do ataque é sempre ordenada antes da regra da coda, então a consoante intervocálica na sequência VCV é invariavelmente incluída no ataque” (ZEC, 2007, p. 166). É comum que a posição de ataque tenha poucas restrições segmentais nas línguas em geral, enquanto a posição de coda costuma permitir um grupo restrito de segmentos para preenchê-la (BLEVINS, 1995).

3. A sílaba em esperanto

Há poucos estudos relacionados a essa temática sobre o esperanto. Nossas principais fontes são artigos de Oostendorp (1999) e Bavant (2006), que fez uma compilação de vários livros que citam a noção de "sílabas" e tentam defini-la, além de ter analisado, por meio de um programa computacional, todas as palavras (mais de 47 mil) que compõem o *PIV – Plena Ilustrita Vortaro* (Dicionário Completo ilustrado), para achar os constituintes possíveis da sílaba e sua frequência na língua.

Oostendorp (1999) faz uma descrição das restrições silábicas que as palavras possuem em esperanto em posição inicial de palavra. Ele afirma que “[...] é importante observar que nem toda combinação de dois segmentos pode ser empregada em esperanto como ataque. O primeiro segmento tem que ser um elemento do grupo {b, d, f, g, k, p, s, ŝ, t, v} e o segundo um elemento do grupo {r, l, n}” (OOSTENDORP, 1999, p. 57, tradução nossa)¹. A tabela a seguir, formulada pelo autor, mostra as combinações possíveis e aquelas que são ou não atestadas em palavras existentes na língua (sem considerar as sequências possíveis de consoante+semivogal):

Tabela 2: Ataques complexos existentes em esperanto segundo Oostendorp (1999).

[br]	<i>bruna</i> ‘marrom’, <i>brako</i> ‘braço’, <i>branĉo</i> ‘galho’
[bl]	<i>blua</i> ‘azul’, <i>blago</i> ‘espécie de piada’, <i>bloko</i> ‘bloco’
[bn]	não atestado
[dr]	<i>drinki</i> ‘beber bebida alcoólica’, <i>droni</i> ‘afogar-se’
[dl]	não atestado
[dn]	apenas em nomes geográficos (<i>Dnepro</i> ‘Dniepre’, nome de rio da Rússia e <i>Dnestro</i> ‘Dniestre’, nome de rio da Ucrânia)
[fr]	<i>franca</i> ‘francês’, <i>fraŭlo</i> ‘homem solteiro’
[fl]	<i>flava</i> ‘amarelo’, <i>Flandrio</i> ‘Flandres’, região da Bélgica
[fn]	não atestado
[gr]	<i>granda</i> ‘grande’, <i>griza</i> ‘cinza’
[gl]	<i>glaso</i> ‘copo’, <i>gliti</i> ‘deslizar’
[gn]	<i>gnomo</i> ‘gnomo’, <i>gnuo</i> ‘gnu’
[kr]	<i>kreteno</i> ‘cretino’, <i>krepo</i> ‘panqueca’
[kl]	<i>klera</i> ‘pessoa culta’, <i>klara</i> ‘claro’
[kn]	<i>knabo</i> ‘menino’, <i>knedi</i> ‘amassar’
[lr]	não atestado
[ll]	não atestado
[ln]	não atestado
[mr]	não atestado
[ml]	não atestado
[mn]	não atestado
[nr]	não atestado
[nl]	não atestado
[nn]	não atestado
[pr]	<i>preni</i> ‘pegar’, <i>profiti</i> ‘aproveitar’
[pl]	<i>plano</i> ‘plano’, <i>plori</i> ‘chorar’
[pn]	<i>pneŭmonio</i> ‘pneumonia’
[sr]	não atestado
[sl]	<i>slipo</i> ‘ficha (de cartão)’, <i>slango</i> ‘gíria’
[sn]	<i>snobo</i> ‘esnobe’, <i>snufi</i> ‘fungar’

¹ “[...] it is important to observe that not every combination of two segments can serve as an Esperanto onset. The first segment always has to be an element of the set {b, d, f, g, k, p, s, ŝ, t, v} and the second one an element of {r, l, n}”.

[ŝr]	<i>ŝraŭbo</i> ‘parafuso’, <i>ŝranko</i> ‘armário’
[ŝl]	<i>ŝlifi</i> ‘erodir’, <i>ŝlosilo</i> ‘chave’
[ŝn]	<i>ŝnuro</i> ‘corda’
[tr]	<i>trajno</i> ‘trem’, <i>tri</i> ‘três’
[tl]	não atestado (exceto em <i>tlaspo</i> ‘espécie de erva’)
[tn]	não atestado
[vr]	<i>vrako</i> ‘destroço’, <i>vringi</i> ‘escorrer’
[vl]	nos nomes <i>Vladimiro</i> e <i>Vladivostoko</i>
[vn]	não atestado
[zr]	não atestado
[zl]	<i>zloto</i> ‘zloty’ – moeda da Polônia
[zn]	não atestado

Segundo o autor, o estudo da estrutura da rima nas palavras em esperanto é um pouco mais complicado, pois

[...] não faz sentido estudar as rimas das últimas sílabas das palavras quando se trata da fonologia do esperanto. A razão para isto é que a maioria das palavras em esperanto termina em uma vogal gramatical; o número desses finais é pequeno e há provavelmente mais rimas possíveis do que terminações gramaticais possíveis (OOSTENDORP, 1999, p. 68, tradução nossa).²

E ainda nos mostra na seguinte tabela quais são as terminações possíveis para a rima das sílabas (os acréscimos explicativos entre parênteses são nossos):

Tabela 3: Rimais possíveis em esperanto, segundo Oostendorp (1999).

Terminações gramaticais	<i>-a</i> (adjetivo), <i>-aj</i> (adjetivo plural), <i>-ajn</i> (adjetivo plural no acusativo), <i>-am</i> (correlativo de tempo), <i>-an</i> (adjetivo no acusativo), <i>-as</i> (tempo presente), <i>-aŭ</i> (ou), <i>-e</i> (advérbio), <i>-el</i> (correlativo de modo), <i>-en</i> (advérbio de direção), <i>-es</i> (correlativo de posse), <i>-i</i> (verbo no infinitivo), <i>-is</i> (tempo passado), <i>-o</i> (substantivo), <i>-oj</i> (substantivo plural), <i>-ojn</i> (substantivo plural no acusativo), <i>-on</i> (substantivo no acusativo), <i>-om</i> (correlativo de quantidade), <i>-os</i> (tempo futuro), <i>-u</i> (imperativo), <i>-us</i> (tempo condicional)
Itens de classe fechada	<i>unu</i> (um), <i>du</i> (dois), <i>tri</i> (três), <i>kvar</i> (quatro), <i>kvin</i> (cinco), <i>ses</i> (seis), <i>sep</i> (sete), <i>ok</i> (oito), <i>naŭ</i> (nove), <i>dek</i> (dez), <i>cent</i> (cem), <i>mil</i> (mil);

² “[...] it does not make sense to study the rhymes of the last syllables of words when studying Esperanto phonology. The reason for this is that most Esperanto words end in a grammatical vowel; the number of these endings is small and there probably are more possible rhymes than possible grammatical endings”.

	<p><i>el</i> (de dentro de), <i>al</i> (para), <i>ĉe</i> (perto de), <i>da</i> (de, relativo a quantidade), <i>de</i> (de), <i>dum</i> (durante), <i>ekster</i> (fora de), <i>en</i> (em), <i>far</i> (longe de), <i>ĝis</i> (até), <i>inter</i> (entre), <i>je</i> (preposição com sentido indefinido), <i>krom</i> (além de), <i>kun</i> (com), <i>per</i> (por meio de), <i>plus</i> (mais), <i>po</i> (à razão de), <i>por</i> (por), <i>post</i> (depois), <i>preter</i> (pelo lado de), <i>pri</i> (a respeito de), <i>pro</i> (por causa de), <i>sen</i> (sem), <i>sub</i> (embaixo de), <i>super</i> (em cima de), <i>sur</i> (acima), <i>tra</i> (através), <i>trans</i> (através de);</p> <p><i>ĉar</i> (porque), <i>do</i> (então), <i>kaj</i> (e), <i>nek</i> (nem), <i>sed</i> (mas), <i>tamen</i> (embora);</p> <p><i>ke</i> (que), <i>kvankam</i> (entretanto), <i>se</i> (se); <i>ajn</i> (qualquer que), <i>nur</i> (apenas), <i>eĉ</i> (até mesmo), <i>des</i> (tanto), <i>tuj</i> (logo), <i>jes</i> (sim), <i>ne</i> (não), <i>nu</i> (ora), <i>ek</i> (começo de), <i>la</i> (artigo definido)</p>
--	---

Oostendorp (1999, p. 69).

Estas palavras da tabela são invariáveis, mas ainda nos sobra a possibilidade de estudar as rimas das sílabas internas dos radicais. A lógica de Oostendorp foi elencar sistematicamente os segmentos possíveis em começo de palavras para achar os ataques e aqueles possíveis em final de palavra para obter as codas, mas o autor ignora as porções mediais das palavras, o que torna o estudo parcial apenas, mas, de qualquer forma, seus resultados devem ser elencados entre as possibilidades estruturais da língua.

Bavant (2006) faz uma análise mais ampla, comparando vários autores que, de uma forma direta ou indireta, tentam definir a sílaba em esperanto, e faz ressalvas sobre as afirmações de Oostendorp. Sigamos o artigo por partes para melhor entendemos as afirmações do autor.

Bavant diz que a noção de sílaba é muito pouco explorada em relação ao esperanto. A gramática mais popular do esperanto na atualidade, a *Plena Manlibro de Esperanta Gramatiko (PMEG)* – Manual completo da gramática do esperanto – evita usar a noção de sílaba, mas há uma frase em que o autor diz que “o acento cai sempre na penúltima sílaba” (WENNERGREN, apud BAVANT, p. 2). No geral, Wennergren (2005) não fala sobre sílabas acentuadas, mas sim vogais acentuadas.

O *Fundamento*, livro oficial da gramática do esperanto, elaborado pelo seu criador, L. L. Zamenhof (1905), usa a noção de sílaba na regra número 10 (há 16 regras gramaticais apresentadas no livro), que consiste em “o acento cai na penúltima sílaba da palavra”, mas não define nem explica como fazer a separação silábica. Ainda no mesmo livro encontramos

exercícios de leitura que trazem as palavras divididas por hífen, supostamente nos limites silábicos delas. Estes exercícios são os números 2 e 3, que consistem em:

§2

Ekzerco de legado.

Al. Bá-lo. Pát-ro. Nú-bo. Cé-lo. Ci-tró-no. Cén-to. Sén-to. Scé-no. Sci-o. Có-lo. Kó-lo. O-fí-cí-ro. Fa-cí-la. Lá-ca. Pa-cú-lo. Ĉar. Ĉe-mí-zo. Ĉi-ká-no. Ĉi-é-lo. Ĉu. Fe-lí-ĉa. Cí-a. Ĉí-a. Pro-cé-so. Sen-ĉé-sa. Ec. Eĉ. Ek. Da. Lú-do. Dén-to. Plén-di. El. En. De. Té-ni. Sen. Vé-ro. Fá-li. Fi-dé-la. Trá-fi. Gá-lo. Grán-da. Gén-to. Gíp-so. Gús-to. Lé-gi. Pá-go. Pá-ĝo. Ĝis. Ĝús-ta. Ré-ĝi. Ĝar-dé-no. Lón-ga. Rég-no. Síg-ni. Gvar-dí-o. Lín-gvo. Ĝu-á-do. Há-ro. Hi-rún-do. Há-ki. Ne-hé-la. Pac-hó-ro. Ses-hó-ra. Bat-hú-fo. Hó-ro. Ĥó-ro. Kó-ro. Ĥo-lé-ro. Ĥe-mí-o. I-mí-ti. Fí-lo. Bír-do. Tró-vi. Prin-tém-po. Min. Fo-í-ro. Fe-í-no. I-el. I-am. In. Jam. Ju. Jes. Ju-ris-to. Kra-jó-no. Ma-jés-ta. Tuj. Dó-moj. Ru-í-no. Prúj-no. Ba-lá-i. Pá-laj. De-í-no. Vėj-no. Pe-ré-i. Mál-plej. Jús-ta. Ĵus. Ĵé-ti. Ĵa-lú-za. Ĵur-nálo. Má-jo. Bo-ná-ĵo. Ká-po. Ma-kú-lo. Kés-to. Su-ké-ro. Ak-vo. Ko-ké-to. Li-kvó-ro. Pac-ká-po.

§3

Ekzerco de legado.

Lá-vi. Le-ví-lo. Pa-ró-li. Mem. Im-plí-ki. Em-ba-rá-so. Nó-mo. In-di-fe-rén-ta. Inter-na-cí-a. Ol. He-ró-i. He-ro-í-no. Fój-no. Pí-a. Pál-pi. Ri-pé-ti. Ar-bá-ro. Sá-ma. Stá-ri. Si-gé-lo. Sis-té-mo. Pe-sí-lo. Pe-zí-lo. Sén-ti. So-fís-mo. Ci-pré-so. Ŝi. Pá-ŝo. Stá-lo. Ŝtá-lo. Vés-to. Vêŝ-to. Dis-ŝí-ri. Ŝan-cé-li. Ta-pí-ŝo. Te-o-rí-o. Pa-tén-to. U-tí-la. Un-go. Plú-mo. Tu-múl-to. Plu. Lú-i. Kí-u. Ba-lá-u. Tra-ú-lo. Pe-ré-u. Ne-ú-lo. Fráŭ-lo. Paŭ-lí-no. Láŭ-di. Eŭ-ró-po. Tro-ú-zi. Ho-dí-aŭ. Vá-na. Vér-so. Sól-vi. Zór-gi. Ze-ní-to. Zo-o-lo-gí-o. A-zé-no. Me-zú-ro. Ná-zo. Tre-zó-ro. Mez-nók-to. Zú-mo. Sú-mo. Zó-no. Só-no. Pé-zo. Pé-co. Pé-so. Ne-ní-o. A-dí-aŭ. Fi-zí-ko. Ge-o-gra-fí-o. Spi-rí-to. Lip-há-ro. In-díg-ni. Ne-ní-el. Spe-gú-lo. Ŝpí-no. Né-i. Ré-e. He-ró-o. Kon-scí-i. Tra-e-té-ra. He-ro-é-to. Lú-e. Mó-le. Pá-le. Tra-í-re. Pa-sí-e. Me-tí-o. In-ĝe-ni-é-ro. In-sék-to. Re-sér-vi. Re-zér-vi. (Zamenhof, 1905, pp. 29-31)

O acento agudo nas palavras indica a vogal que deve ser acentuada. Segundo Bavant (2006), pode-se concluir o seguinte destes exercícios de leitura:

- Não confunda/misture as letras *c*, *ĉ*, *k*; *h*, *ĥ*; *g*, *ĝ*, *s*, *z*, *c*
- A letra *h* deve ser falada aparte, até mesmo depois de outra consoante: não se trata de dígrafos (*pac-hó-ro*, *ses-hó-ra*, *bat-húfo*)
- Vogais dobradas não equivalem a uma vogal longa (*ré-e*, *he-ró-o*, *kon-scí-i*)
- Os grupos vocálicos *ae*, *oe*, *ue* não apresentam nem ditongos nem metafonia (*tra-e-té-ra*, *he-ro-é-to*, *lú-e*)
- As vogais *i* e *u* sempre formam sílabas em contraste com *j* e *ŭ* (*ba-lá-i*, *ru-*

i-no/prúj-no, in-ĝe-ni-é-ro; ba-lá-u, tra-ú-lo/fráŭ-lo, ĝu-á-do).

Ainda que de forma mais sutil, ainda aparecem as seguintes regras sobre silabificação:

- Cada sílaba contém precisamente uma vogal
- Se duas vogais separam uma consoante da outra, o limite silábico se encontra antes da consoante (*si-gé-lo*); isso também é válido se a consoante é africada (*pro-cé-so, fe-lí-ĉa, pá-ĝo*)
- Se duas vogais separam duas consoantes umas das outras, a primeira sendo oclusiva e a segunda sendo líquida (*l* ou *r*), o limite silábico se encontra antes da oclusiva (*ci-tro-no, ci-pre-so, ge-o-gra-fi-o*), mas em outros casos com duas consoantes distintas ele se encontra após a primeira (*sis-té-mo, vér-so, láŭ-di*)
- Palavras compostas não obedecem a regras específicas para a silabificação (*ar-bá-ro, ko-ké-to, bo-ná-ĵo, pa-cú-lo, ju-ris-to, Paŭ-lí-no*).

(Bavant, 2006, p. 3-4)

O próprio autor reconhece que extrapolou os exemplos para formular a penúltima regra estabelecida, pois não há nenhum exemplo com *l*.

Continuando sua análise sobre os livros em esperanto que falam sobre sílaba, Bavant examina a *Plena Analiza Gramatiko (PAG)* – Gramática analítica plena –, de Kalocsay e Waringhien (1985). Ele diz que “com o atrativo título 'Sílabas', encontra-se apenas a seguinte informação: 'cada vogal, exceto *ŭ*, corresponde a uma sílaba aparte'. Segue uma definição de sílabas longas e curtas, mas nada mais sobre a definição da sílaba em si!” (BAVANT, 2006, p. 5). Ele afirma ainda que é de conhecimento geral que o livro *PAG* tem a teoria de que *ŭ* não é uma vogal. Sobre isso, deve-se considerar que há uma certa não uniformidade sobre como se enquadrariam, dentro de uma especificação fonológica, os grafemas *ŭ* (/w/) e *j* (/j/), que são aproximantes, e formam ditongos com as vogais em esperanto. Alguns afirmam ser semivogais, outros, semiconsoantes. O importante é que tais segmentos são glides, e enquadrados como aproximantes. De fato, eles nunca aparecem sozinhos entre duas consoantes, sempre estão ao lado de uma vogal, e formam ditongos.

Como o *PAG* é uma obra de referência até os dias atuais sobre a gramática do esperanto (assim como o *PMEG*), muitos autores passivamente aceitaram a noção de sílaba presente no livro e a reproduziram, ou até mesmo já era a noção que se tinha antes, de forma não muito analítica. Bavant alega que pode ser o caso do livro *La tuta Esperanto – O Esperanto completo* – (SEPPK, 1987), no qual a definição de sílaba é “sílaba é um som ou grupo de sons que é falado por meio de uma expiração” (SEPPK, 1987, apud BAVANT, 2006, p. 7). Segundo Bavant, a afirmação, embora imprecisa, pelo menos informa que sílaba é um agrupamento

sonoro à parte (p. 8). Seppik segue dando exemplos, e afirma que “j e ŭ não são vogais e por causa disso não podem formar sílabas autônomas” (SEPPIK, 1987, p. 5). A separação silábica, segundo ele, “acontece geralmente segundo as mesmas regras que nas outras línguas; *cit-ro-no* ou *ci-tro-no*, *prob-le-mo* ou *pro-ble-mo*” (SEPPIK, p. 6). Essa definição (assim como as outras discutidas no artigo de Bavant) são muito mais prescritivas do que descritivas, e, portanto, pouco científicas, mas interessantes para analisarmos o que existe, na literatura em esperanto, sobre o assunto aqui tratado. Seppik faz ainda uma distinção entre o nível fonológico e morfológico (sem usar esses termos), pois diz que não se deve misturar afixos, radicais e terminações gramaticais (SEPPIK, p. 6). Bavant diz que esta afirmação é um absurdo, pois determina que deveríamos separar a palavra *facila* (fácil) como *fa-ci-la*, mas *tranĉila* (*tranĉ* – radical para “corte”, *il* – afixo para formar instrumento) como *tranĉ-il-a*. Assim como Seppik, Wüster (1923) considera que os morfemas não devem se misturar aos radicais na separação silábica de palavras compostas. Por fim, Bavant analisa o dicionário oficial da Academia de Esperanto, o *PIV*, e chega a conclusão de que a definição nele encontrada é inútil, pois há vários contraexemplos que não são explicados por tal definição, e também cita a partitura do hino do esperanto, *La Espero*, que também faz diferenciação entre os níveis do radical e os afixos e outras derivações.

Bavant resume os livros e fontes analisadas na seguinte tabela.

Tabela 4: Estruturas silábicas possíveis encontradas por Bavant (2006).

Fundamento	PAG	PIV	Wüster	Seppik	"La Espero"
fa-ci-la	fa-ci-la	fa-ci-la	fa-ci-la	fa-ci-la	fa-cil-a
ju-ris-to	(falta regra)	ju-ris-to	ju-ris-to	*jur-ist-o	*jur-ist-o
??	(falta regra)	fa-ci-la-ni-ma ba-na-nar-bo	fa-cil-a-ni-ma ba-nan-ar-bo	*fa-cil-a-nim-a ba-nan-arb-o	*fa-cil-a-nim-a *ba-nan-arb-o
ci-tro-no	ci-tro-no	ci-tro-no	ci-tro-no	ci-tro-no <i>aŭ</i> cit-ro-no	*ci-tron-o
pat-ro	pa-tro	pa-tro	pa-tro	pa-tro <i>aŭ</i> pat-ro	*patr-o
ak-vo	(falta regra)	ak-vo	a-kvo	??	*akv-o

li-kvo-ro	(<i>falta regra</i>)	lik-vo-ro	li-kvo-ro	??	??
kon-sci-i	kon-sci-i	kon-sci-i	kons-ci-i	??	??
??	ob-sti-na	ob-sti-na	obs-ti-na	??	obs-tin-a
lin-gvo	ling-vo	ling-vo	lin-gvo	??	*lingv-o
??	tung-ste-no	tun-gste-no	tungs-te-no	??	??

Adaptado de Bavant (2006, p. 13).

Dos livros analisados por Bavant, e pelas próprias considerações do autor, pode-se afirmar que o estudo da estrutura silábica do esperanto não tem bases firmes ou referências bibliográficas uniformes sobre o tema.

Sobre o trabalho de Oostendorp (1999), citado anteriormente, Bavant se questiona se realmente as regras postuladas podem ser consideradas corretas, uma vez que o autor não levou em consideração muitas palavras com ataques diferentes daqueles por ele elencados (palavras de baixa frequência na língua), e considerou apenas os ataques em começo de palavras, o que não engloba a totalidade de ataques que podem aparecer na porção medial das palavras também (BAVANT, 2006, p. 21). Como exemplo, Bavant cita os ataques *tl* e *dl*, os quais Oostendorp considera como não atestados, a não ser pela palavra *tlaspo*, bastante infrequente na língua. Entretanto, existem palavras como *atlantiko*, *atlaso*, *atleto*, *kotleto*, entre outras, que tem o ataque composto *tl* em sua composição.

4. Considerações finais

Vimos que a estruturação silábica segue regras que se relacionam com o grau de sonoridade dos segmentos e com a posição estrutural que eles podem ocupar. A partir disso, podemos montar uma escala para o esperanto, e tirarmos nossas primeiras conclusões sobre as possíveis sílabas que podem aparecer na língua, em palavras não derivadas, ou seja, radicais simples. A escala, segundo o grau de sonoridade, é: a > e o > i u > j w > l r > m n > v z > f s > ĵ h x > dʒ > tʃ ts > b d g > p t k (ortograficamente: a > e o > i u > j ŭ > l r > m n > v z ĵ > f s ŝ h ĥ > ĝ > ĉ c > b d g > p t k)³.

³ vogais baixas > vogais médias > vogais altas > glides > líquidas > nasais > fricativas sonoras > fricativas surdas > africada sonora > africadas surdas > oclusivas sonoras > oclusivas surdas. Oostendorp (1999: 59) também elenca tal escala de sonoridade, mas de forma menos especificada.

Como conclusão parcial, podemos afirmar que a estrutura silábica máxima em esperanto é (C)(C)V(C)(C). Exemplos:

- Ataque e coda simples: ĵurnalo (jornal) – CVC.CV.CV
- Coda complexa: post (depois) – CVCC
- Ataque complexo: granda (grande) – CCVC.CV
- Sílabas sem ataque e sem coda: ĉielo (céu) – CV.V.CV

A especificação sobre quais elementos podem ou não preencher cada posição na sílaba se torna um pouco mais complicada, e é preciso um maior aprofundamento no tema para se ter conclusões sobre o assunto. Por exemplo, além dos ataques complexos elencados por Oostendorp (1999) e citados neste trabalho, há outros, menos frequentes, como *gneto* e *kŝatrio*, citados por Bavant (2006, p. 39). Este autor também efetuou um estudo muito interessante por meio de um programa computacional, utilizando a linguagem de programação *perl*, que listou as sílabas mais frequentes no esperanto (de acordo com o *corpus* presente no dicionário estudado). As dez primeiras da lista são: **to** (que apareceu 6.287 vezes no *corpus*), **o** (5.552), **lo** (4.367), **no** (4302), **a** (3.828), **ti** (3.448), **do** (3.212), **ro** (3.130), **ko** (3.062) e **ta** (2.987) (Bavant, 2006, s/ pg). Por estes resultados, podemos afirmar que as sílabas mais frequentes em esperanto tem o padrão V ou CV. A sílaba fechada que primeiro apareceu na lista é **kon**, na posição número 54, ou seja, as 53 sílabas mais frequentes são abertas, e nenhuma delas têm ataque complexo.

Ao final desta exposição, o ponto estrutural sobre o qual achamos menos análises foi a coda e a separação interna de vários encontros consonantais, que não se apresenta bem clara. Bavant afirma ainda em seu estudo que, em 4% das palavras estudadas, o local de separação silábica na porção medial das palavras foi duvidoso, segundo os seus postulados. A palavra com maior encontro consonantal encontrada pelo autor foi *angstromo*, com 5 consoantes em seguida. É digno de nota, entretanto, que a entrada lexical para esta palavra no dicionário citado traz a seguinte afirmação: "angstrom/o = anstromo"⁴, ou seja, ao longo da evolução da língua o /g/ da porção medial foi elidido, possivelmente por não se enquadrar na estruturação silábica permitida. A entrada lexical de *anstromo* é: "Anstrom/o. Sveda fizikisto (A.J. Ångström,

⁴ Disponível em: <<http://vortaro.net/#angstromo>>.

1814—1874). anstromo. Unuo de longo, uzata en spektroskopio, k egala al 10^{-10} metroj; simb.: Å."⁵

Outro ponto a ser ainda aprofundado é a possível extrassilabicidade do segmento /s/ em ataques com três elementos, como *sklavo* (escravo), já que a sequência *skl* fere a escala de sonoridade, pois a sonoridade decresce do *s* para o *k* e cresce do *k* para o *l*, para logo em seguida atingir o pico de sonoridade no núcleo silábico *a*. "Kla", portanto, seria uma sílaba bem formada, mas "skla" não. A pesquisa sobre o tema tratado neste artigo, por conseguinte, ainda requer uma maior reflexão teórica (e prática).

Referências Bibliográficas

BAVANT, M. **Silabo kaj Silabado**. Disponível em: <<http://lingvakritiko.com/2006/12/20/silabo-kaj-silabado/>>. Acesso em: 20 nov. 2014.

_____. **Perkomputila silaba analizo de la vortprovizo de PIV 2002**. Disponível em: <<http://perso.orange.fr/kursoj/studoj/silab.htm>>. Acesso em: 10 dez. 2014.

BLEVINS, J. The syllable in Phonological Theory. In: Goldsmith, John A. **The Handbook of Phonological Theory**. Blackwell Publishers, 1995.

ECO, U. **A procura da língua perfeita**. Lisboa: Editorial Presença, 1996.

FRAWLEY, W J. (ed.). Artificial Languages. **Internacional encyclopedia of linguistics**. Oxford: Oxford University Press, 2003.

KALOCSAY, K; WARINGHIEN, G. **Plena Analiza Gramatiko**. 5-a korektita eldono. Rotterdam, 1985. 598 p.

OOSTENDORP, M van. Syllable structure in Esperanto as an instantiation of universal phonology . **Esperantologio – Esperanto Studies**, n. 1, 1999, pp. 59-80. Disponível em: <<http://www2.math.uu.se/esperanto/oostendorp.pdf>>. Acesso em: 11 mar. 2014.

SEPPIK, H. **La tuta Esperanto**. 4-a eldono. HEA, Budapest, 1987. 181 p. Disponível em: <http://www.esperanto.mv.ru/Seppik/lec01.html>. Acesso em: 17 jan. 2015.

WENNERGREN, B. **Plena Manlibro de Esperanta Gramatiko**. Sacramento/California: ELNA, 2005.

WÜSTER, E. Die Silbentrennung der Wörter und die Aussprache des Selbstlaute im Esperanto. **Germana Esperantisto**, junio 1923, 3p.

⁵ Disponível em: <<http://vortaro.net/#anstromo>>. Tradução: "Anstrom/o. Físico sueco. (A.J. Ångström, 1814—1874). anstromo. Unidade de comprimento, usada em espectrografia e igual a 10—10 metros; simb.: Å."

ZAMENHOF, L. **Fundamento de esperanto**. 1905. Disponível em: <http://i-espero.info/files/elibroj/eo%20-%20fundamento%20de%20esperanto.pdf>. Acesso em: 22 jan. 2014.

ZEC, D. The syllable. In: LACY, Paul de. **The Cambridge Handbook of Phonology**. Cambridge: Cambridge University Press, 2007.

Bibliografia

BERGEN, B. K. Nativization processes in L1 Esperanto. **Journal of child language**, v. 8, n. 3, 2001.

COLLING, I. E. **Comparação entre os sistemas fonéticos do esperanto e do português**. 2012. (Cadernos da Semana de Letras – Ano 2012 volume I). Disponível em: <http://www.semanadeletras.ufpr.br/cadernos-da-semana/>. Acesso em: 21 maio 2014.

CORSETTI, R. A mother tongue spoken mainly by fathers. **Language Problems and Language Planning**, 20: 263–273, 1996a. **crossref** <http://dx.doi.org/10.1075/lplp.20.3.05cor>

_____. Esperanto kaj kreolaj lingvoj: komunaj kaj malkomunaj trajtoj en la kreoliĝo. In: KOŠECKÝ, S. (ed.). **Multkulturaj familioj de nuntempa Eŭropo**. Aktoj de la seminario okazinta 01–02.08.1996, pp. 11-36. Bratislavo: Esprima. Includes also an English summary: “Esperanto and Creole languages: common and different characteristics in the process of creolisation.” 1996b.

_____. Regularizing the regular: The phenomenon of overregularization in Esperanto-speaking children. **Language Problems and Language Planning**, 28: 261–282, 2004. **crossref** <http://dx.doi.org/10.1075/lplp.28.3.04cor>

DIAS, A. E. W. **Renovação lexical do Esperanto: mecanismos de formação de neologismos**. Dissertação de Mestrado. Brasília: Universidade de Brasília, 2007.

DOLS SALAS, N. Perspektivoj en fonologia kaj fonetika esplorado de Esperanto: Bjalistoko, Pollando 25 julio-1 aŭgusto 2009. pp. 24-47. VERGARA, J. A. (ed.). **IKU 62: Internacia Kongresa Universitato, 62a sesio**. Rotterdam, Netherlands: Universala Esperanto-Asocio, 2009. 140 pp. Disponível em: <http://www.uea.org/pdf/IKU/IKU2009.pdf>. Acesso em: 21 maio 2014.

_____. **Phonology and morphology and the limits of freedom in an artificial language**. 2012. Disponível em: http://roa.rutgers.edu/content/article/files/1288_dols_salas_1.pdf. Acesso em: 05 out. 2015.

DOMINTE, C.. Fonemstatistikaj rimarkoj pri Esperanto: Festlibro omâge al la 60-jariĝo de Detlev Blanke/Festschrift für Detlev Blanke zum 60. Geburtstag. pp. 337-45. FIEDLER, S.; LIU, H. (ed., introd., and bibliography). **Studoj pri interlingvistiko/Studien zur Interlinguistik**. Prague, Czech Republic: Kava-Pech, 2001. 736 pp. Disponível em: http://www.lingviko.net/db/23_Dominte.htm. Acesso em: 05 out. 2015.

HERRING, J.. **Syntactic and Lexical Changes in Esperanto: a Quantitative and Corpus-Based Survey**. Disponível em: http://cllt.osu.edu/mclc/paper/syntactic_herring.pdf. Acesso em: 24 mar. 2014.

LAROCA, M. N. de C.. **O caráter verbo-nominal do aspecto em esperanto**. Tese (Doutorado em Linguística) – Instituto de Letras da Universidade Federal Fluminense, Niterói, 2009.

LINDSTEDT, J. **Kreoliĝo kaj spontanea ŝanĝigo de Esperanto**. Manuscript, University of Helsinki., 1997.

_____. **Native Esperanto as a Test Case for Natural Language**. Disponível em: http://www.linguistics.fi/julkaisut/SKY2006_1/1FK60.1.5.LINDSTEDT.pdf. Acesso em: 15 abr. 2014.

MINER, K. **Limigoj al esperanta elizio**. Disponível em: http://www.sunflower.com/~miner/LIMIGOJ_package/limigoj.html. Acesso em: 15 abr. 2014.

NAGATA, H. Esperantaj parolsonoĵoj perceptitaj de japanlingvanoĵoj. **Scienca Revuo**: 55.1 [196] (2004), pp. 13-24. Disponível em: <https://scienca-revuo.info//article/view/867/846>. Acesso em: 24 mar. 2014.

OKHUIJSEN, G. **Comparison between English Loan Words in Dutch and their Esperanto Equivalents**. Amsterdam: Dissertação de mestrado. Universiteit van Amsterdam – Faculteit der Geesteswetenschappen. Disponível em: <http://dare.uva.nl/cgi/arno/show.cgi?fid=345847>. Acesso em: 24 mar. 2014.

OOSTENDORP, M. van. A note on exceptional syllable structure in esperanto. In: BRUYN, A. ; ARENDS, J. (eds.). **Mengelwerk voor Muysken; Voor Pieter Muysken bij zijn afscheid van de Universiteit van Amsterdam**, pp. 98–101. Amsterdam: Institut voor Algemeine Taalwetenschap, University of Amsterdam, 1998. 268 pp. Disponível em: <http://www.vanoostendorp.nl/interlinguistiek/esse.html>. Acesso em: 24 mar. 2014.

POKROVSKIJ, S. Duonvokaloĵoj kaj diftongoĵoj. **Lingva Kritiko**: Studoj kaj notoj pri la Internacia Lingvo, Esperantologio Interreta. Disponível em: <http://lingvakritiko.com/2014/01/21/duonvokaloj-kaj-diftongoj/>. Acesso em: 15 abr. 2014.

VITALI, D. La nazaloĵoj de la Internacia Lingvo: Festlibro omâĝe al la 60-jariĝo de Detlev Blanke/Festschrift für Detlev Blanke zum 60. Geburtstag". pp. 326-36. FIEDLER, S.; LIU, H. (ed., introd., and bibliography). **Studoj pri interlingvistiko/Studien zur Interlinguistik**. Prague, Czech Republic: Kava-Pech, 2001. 736 pp. Disponível em: http://www.lingviko.net/db/22_Vitali.pdf. Acesso em: 15 abr. 2014.

WELLS, J. **Lingvistikaj aspektoj de esperanto**. Roterdã: UEA, 1978.

Artigo recebido em: 05.10.2015

Artigo aprovado em: 04.04.2016

Variações de F0 e configurações de frase entoacional: análise de estruturas contrastivas

F0 variations and intonational phrase configurations: analysis of contrastive structures

Geovana Soncin*
Luciani Tenani**

RESUMO: Segundo métodos da fonologia laboratorial, o presente trabalho tem como objetivo descrever variações de F0 em sentenças iguais quanto à cadeia segmental, mas distintas entre si quanto à configuração prosódica em frases entoacionais. Nessa descrição, as variações de F0 são interpretadas de modo a identificar eventos tonais que caracterizam o contorno entoacional de sentenças em contraste. No decorrer da análise, são discutidas decisões metodológicas adotadas no procedimento experimental que norteiam a pesquisa. Os resultados encontrados mostraram que as diferenças tonais nas configurações da frase entoacional em comparação são definidas, privilegiadamente, em função (i) do número de palavras prosódicas que compõem as frases entoacionais; (ii) da posição que as palavras prosódicas ocupam nos diferentes fraseamentos desse constituinte e (iii) do número de sílabas das palavras prosódicas. Por meio desses resultados, mostramos que a organização dos eventos tonais é fator de importância primária para o estabelecimento de diferenças estruturais no domínio da frase entoacional em Português Brasileiro.

PALAVRAS-CHAVE: Entoação. Fonologia Prosódica. Interface sintaxe-fonologia. Língua Portuguesa. Vírgula.

ABSTRACT: Through methods of laboratory phonology, this article describes F0 variations in sentences which share identical segmental strings, but present differences related to the prosodic configurations, especially in the domain of intonational phrase. In this description, F0 variations are interpreted to identify tonal events which characterize the intonation contour of contrastive sentences. During the analysis, this work presents methodological decisions adopted in the experimental procedure that guides the research. The results showed that the tonal differences in the configurations of intonational phrase in comparison are defined according to: (i) the number of prosodic words that form the intonational phrases; (ii) the position of the prosodic words within the different intonational phrases compared and (iii) the number of syllables of the prosodic words. Based on these results, we show that the tonal events organization is a factor of primary importance for contrasting structural differences in the intonational phrase domain in Brazilian Portuguese.

KEYWORDS: Intonation. Prosodic Phonology. Syntax-phonology interface. Brazilian Portuguese. Comma.

* Pesquisadora (pós-doutorado FAPESP - Proc: 2014/24778-3) junto ao Departamento de Estudos Linguísticos e Literários do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista – UNESP, São José do Rio Preto.

** Docente do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista – UNESP, São José do Rio Preto.

1. Introdução

Assim como afirmam D’Imperio *et al* (2005), uma das principais funções da prosódia é estruturar a mensagem em “blocos” de determinados tamanhos dotados de estrutura interna. No Português Brasileiro, assim como em Português Europeu (cf. a propósito: FROTA & VIGÁRIO, 2000; TENANI, 2002; FERNANDES, 2007; SERRA, 2009) e em outras línguas românicas (cf. sobre o italiano, D’IMPERIO, 2002; sobre o espanhol, ELORDIETA, FROTA e VIGÁRIO, 2000), a frase entoacional é um dos principais constituintes prosódicos responsáveis por essa estruturação interna. A organização que se configura no fraseamento de sentenças permite que, no plano fonológico representacional, se reconheçam fronteiras às quais, no plano fonético, estão associados eventos tonais que compõem padrões entoacionais de diferentes tipos de sentenças nessas línguas.

No Português Brasileiro, as frases entoacionais mapeadas a partir de sentenças declarativas neutras são reconhecidas por apresentarem, predominantemente, um evento tonal L + H* em sua fronteira esquerda e um evento tonal H + L* em sua fronteira direita (cf. TENANI, 2002; FERNANDES, 2007). Esses tons marcam, respectivamente, o início e o fim de uma sentença declarativa neutra nessa variedade do Português.

Ocorre, no entanto, que o fraseamento de uma dada cadeia segmental pode variar a depender de relações semântico-pragmáticas, por exemplo. Ou seja, uma mesma cadeia segmental pode ser organizada em uma ou mais frases entoacionais ao se considerar a possibilidade de alteração de suas fronteiras no interior de uma mesma sequência sintática em função de relações semânticas, como se exemplifica em (1) e (2):

- (1)
a. não espere
b. [não espere]
c. [não] [espere]

- (2)
a. isso só ele resolve
b. [isso] [só ele resolve]
c. [isso só] [ele resolve]

Em ambos os exemplos, (a) apresenta a cadeia segmental em questão, enquanto (b) e (c) apresentam diferentes fraseamentos (delimitados por colchetes). Qualitativamente, no exemplo (1), os fraseamentos se diferenciam pelo número de frases entoacionais formadas: uma

em (b) e duas em (c); por sua vez, no exemplo (2), a diferença entre (b) e (c) reside na posição onde a cadeia segmental é dividida, já que nos dois casos o fraseamento formou duas frases entoacionais.

Os diferentes fraseamentos prosódicos em (1) e (2) apontam para uma diferença estrutural no modo como a cadeia segmental pode ser dividida no eixo sintagmático. Como consequência desses fraseamentos, criam-se, no eixo paradigmático, dois tipos distintos de segmentação da cadeia, implicando em diferentes possibilidades de interpretação semântica, sobre as quais trataremos mais à frente.

Do ponto de vista prosódico, para a ocorrência desses diferentes fraseamentos, o contorno entoacional exerce função primordial, já que reorganiza as relações de proeminência tonal no interior da cadeia segmental conforme a organização sintática do enunciado. Desse modo, no domínio da frase entoacional, a alteração dos limites da estrutura sintático-prosódica é *refletida na* (e pode ser também entendida como *reflexo da*) realização do contorno entoacional.

Levando em consideração essas questões relativas ao fraseamento prosódico de sentenças e ao modo como diferentes fraseamentos alteram os contornos da entoação, o presente estudo centra-se na descrição das variações de F0 em sentenças de mesma sequência segmental, mas que se diferem estruturalmente no domínio da frase entoacional. Particularmente, o presente estudo objetiva verificar de que modo os parâmetros acústicos de F0 contribuem para a configuração tonal de sentenças como as exemplificadas em (1) e (2).

Antes de passar à discussão dos resultados obtidos, apresentamos, na sequência, a abordagem teórica que fundamenta o presente estudo, os aspectos metodológicos e o material analisado.

2. Quadro teórico

Neste trabalho, a entoação é discutida sob o prisma da relação entre fonética e fonologia de modo a descrever parâmetros acústicos de F0 que, associados à estrutura fonológica, promovem configurações diferentes da frase entoacional.

Para tanto, a descrição da entoação é desenvolvida considerando de modo conjunto os princípios da Fonologia Prosódica (segundo o modelo proposto por Nespor & Vogel, 1986) e da Fonologia Entoacional (conforme a organização de Ladd, 1996, a partir das proposições da

Teoria Autossegmental e Métrica da Entoação, propostas por Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988).

A assunção desses dois modelos teóricos de fonologia se fundamenta na perspectiva da visão integrada entre entoação e domínios prosódicos (defendida para variedades do Português por Frota, 1998; 2000; Frota e Vigário, 2000; Tenani, 2002; Fernandes, 2007; Serra, 2009; entre outros). Segundo tal perspectiva, há interação entre unidades definidas pela entoação e domínios prosódicos formados pelos algoritmos de formação do modelo de Nespor & Vogel (1986). Em outras palavras, assumir a visão integrada implica considerar que fenômenos entoacionais evidenciam a relevância da estrutura prosódica, sendo a direção inversa também verdadeira.

Nas subseções seguintes, apresentamos os princípios teóricos da Fonologia Prosódica e da Fonologia Entoacional que mobilizamos no estudo.

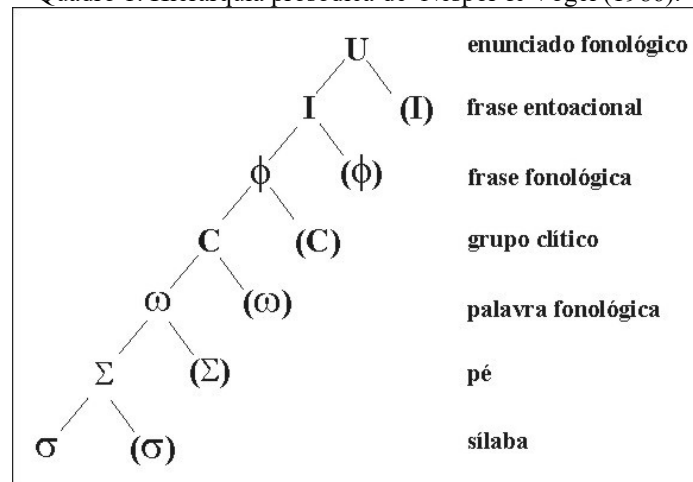
2.1. Fonologia Prosódica

A Fonologia Prosódica é uma teoria não linear de base gerativista caracterizada pela interface com a sintaxe. A teoria postula que os sons de uma língua se organizam hierarquicamente em constituintes prosódicos, cujos algoritmos de formação são embasados em informações de constituintes da sintaxe, sem haver necessidade de isomorfia entre ambos. Em alguns casos, constituinte prosódico e constituinte sintático podem ser equivalentes, mas a coincidência não é uma regra.

Tendo em vista a formação de constituintes prosódicos, a Fonologia Prosódica pressupõe a existência de uma estrutura abstrata subjacente à realização fonética dos fenômenos prosódicos na cadeia da fala. Essa estrutura abstrata confere às línguas regularidades e previsibilidades na organização dos sons, como a existência de regras fonológicas que podem ser aplicadas ou bloqueadas no interior e entre limites de certos constituintes.

Segundo o modelo de Nespor & Vogel (1986), assumido neste trabalho, são sete os constituintes que compõem a hierarquia prosódica, apresentada no Quadro 1 (abaixo).

Quadro 1: Hierarquia prosódica de Nespor & Vogel (1986).



Neste trabalho, a frase entoacional (*I*) é o constituinte prosódico primordialmente contemplado, tendo em vista que a entoação caracterizadora dos diferentes mapeamentos desse constituinte é o alvo da análise apresentada. A *I* define-se pela existência de uma ou mais frases fonológicas e se configura, em geral, no nível da sentença¹. No que se refere à entoação, de acordo com Nespor & Vogel (1986), o contorno entoacional e a produção de pausas são pistas de fundamental importância para a delimitação de *I*, uma vez que a variação desses elementos pode alterar os seus limites no interior de uma mesma cadeia segmental e sintática.

Conforme mostraram Frota e Vigário (2000), Tenani (2002), Fernandes (2007) e Serra (2009), a *I* é um domínio prosódico de fundamental importância para a realização da entoação do Português Brasileiro, uma vez que os eventos tonais distribuídos ao longo dele definem os contornos de entoação dos diferentes padrões de sentenças dessa variedade do Português. Desse modo, embora a *I* seja o domínio privilegiado e discutido no presente trabalho, a frase fonológica (ϕ) e a palavra fonológica (ω) são mencionadas na análise, pois são domínios essenciais para a formação e para a configuração da entoação que caracteriza uma *I*. A combinação dos eventos tonais associados a esses constituintes compõe o que se caracteriza como padrão entoacional de sentenças declarativas neutras ou focalizadas bem como de sentenças interrogativas no Português Brasileiro.

Desse modo, esses três constituintes são relevantes para a descrição da entoação dos dados em análise, uma vez que diferentes eventos tonais (a saber, acentos tonais, acentos frasais

¹ De acordo com Nespor & Vogel (1986) são bases sintáticas para a formação de *I*: (i) quaisquer sentenças raiz, (ii) elementos não anexáveis à estrutura da sentença raiz e (iii) elementos remanescentes de uma sentença raiz interrompida por elemento(s) anexado(s) a ela (cf. NESPOR & VOGEL, 1986, p. 189).

e tons de fronteira) ocorrem distintamente no domínio de cada um deles para a caracterização da entoação do Português Brasileiro. Trabalhos anteriores que descreveram a configuração prosódica e entoacional do Português Brasileiro apresentam resultados que permitem sistematizar a relação entre evento tonal e domínio prosódico nos seguintes termos: (i) acentos tonais se associam a ϕ s (cf. FROTA & VIGÁRIO, 2000; TENANI, 2002) e a ω s (cf. FERNANDES, 2007); (ii) acentos frasais se associam a ϕ s (cf. FERNANDES, 2007; TENANI & FERNANDES-SVARTMAN, 2008); (iii) tons de fronteira se associam a I s (cf. FROTA & VIGÁRIO, 2000; TENANI, 2002; FERNANDES, 2007; SERRA, 2009).

A definição de cada um desses eventos tonais e o modo como são concebidos no interior da teoria fonológica da entoação são apresentados na subseção 2.2, a seguir.

2.2. Fonologia Entoacional

A Fonologia Entoacional é uma teoria da estrutura entoacional, cuja abordagem considera que a entoação tem uma organização fonológica (cf. Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988; Ladd, 1996). À luz dessa abordagem, um contorno entoacional consiste, fonologicamente, em uma sequência de unidades discretas, chamadas de eventos tonais.

Mais especificamente, a teoria da Fonologia Entoacional (LADD, 1996) define que a entoação tem uma organização fonológica ao considerar que o contorno de frequência fundamental (F0), ao invés de um contínuo variável, consiste em uma sequência de eventos tonais discretos, localmente definidos em pontos específicos na cadeia segmental. Desse modo, F0 é o correlato fonético-acústico de uma sequência abstrata de eventos tonais.

Conforme apresenta Frota (2000), uma consequência importante da assunção de que a entoação possui organização fonológica é considerar que

contornos entoacionais intuitivamente idênticos deverão ser descritos através da mesma representação fonológica e contornos intuitivamente diferentes deverão ser representados por alguma diferença na sequência estruturada de categorias tonais (cf. Prieto 1995 e Ladd 1996, entre outros). Exemplificando, o contorno que tem sido apresentado como o contorno declarativo neutro do PE deverá ser definido por uma representação fonológica específica, independentemente das categorias sintáticas das palavras/sintagmas envolvidos na sequência segmental em causa ou do tamanho diverso das frases. Também as diferenças sistemáticas na forma entoacional devem ser tomadas como refletindo a presença de categorias tonais distintas, do mesmo modo que as diferenças sistemáticas de significado entoacional correspondem

a distinções categóricas (ver, por exemplo, Gussenhoven 1983, Pierrehumbert e Steele 1989, Pierrehumbert e Hirschberg 1990). (FROTA, 2000, p. 514)

Nessa abordagem da entoação, o conjunto de eventos tonais possíveis é restrito a duas categorias: (i) os acentos tonais (*pitch accents*), eventos que se associam a elementos proeminentes na cadeia segmental (isto é, cabeças prosódicas) e (ii) os tons de fronteira (*boundary tones*), eventos associados a fronteiras de constituintes prosódicos. Tais eventos são formados apenas por dois tipos de tons, são eles H (tom alto) e L (tom baixo), e podem constituir eventos tonais simples ou complexos, respectivamente, monotonais ou bitonais.

No Português Brasileiro, como mostram Frota & Vigário (2000), Tenani (2002), Fernandes (2007) e Serra (2009), os acentos tonais podem ser monotonais (L* ou H*) e bitonais (H*+L, H+L*, L*+H ou L+H*)². Por sua vez, os tons associados a fronteiras de domínios prosódicos podem ser classificados em duas categorias: tons de fronteira (L% ou H%) ou acentos frasais (L* ou H*).

A respeito da relação entre evento tonal e pontos específicos da cadeia segmental, Ladd (1996) faz uma distinção entre associação e alinhamento. Segundo o autor, na interpretação dos eventos tonais mostrados pelo contorno de F0, enquanto a associação desses elementos a pontos específicos da cadeia segmental se realiza levando em consideração sua estrutura fonológica (por exemplo, as fronteiras de constituintes prosódicos), ao contrário, o alinhamento dos eventos do contorno de F0 a elementos da sequência segmental ocorre no âmbito fonético. Nesse quadro teórico da fonologia entoacional, portanto, como salienta Frota (2000), a associação tonal condiciona a realização tonal, pois a ancoragem do tom é feita tendo como base a posição prosódica proeminente. A fim de esclarecer essa premissa da teoria, a autora exemplifica: “factores de variação fonética à parte, prediz-se que um T* deverá apresentar correlatos fonéticos que o distingam de um T, designadamente o seu alinhamento relativamente a posição proeminente e sua estabilidade tonal (cf. Pierrehumbert e Beckman, 1998, Pierrehumbert e Steele, 1989, Hayes e Lahiri 1991, Grice 1995a e Arvantini *et alii* 1998)”. (FROTA, 2000, p. 520)

Na análise dos dados, abordamos e detalhamos as questões relacionadas à associação e ao alinhamento, uma vez que são importantes para a descrição da entoação das sentenças que

² Na notação, o sinal asterisco (*) indica a sílaba mais proeminente à qual determinado evento tonal se associa.

compõem os dados de análise: sentenças de mesma cadeia segmental, mas com diferentes configurações prosódicas.

3. Metodologia e descrição do material de análise

O material analisado é formado por sentenças gravadas em laboratório que constituem o *input* para teste de percepção auditiva a ser feito em um segundo momento da pesquisa³. O referido teste terá como objetivo principal avaliar se variações no contorno de F0 são auditivamente percebidas como pausas em fronteiras de *Is* não finais.⁴ É, portanto, crucial a caracterização fonético-fonológica dos dados de *input*, a qual fazemos neste artigo. De modo mais amplo, a descrição e a discussão que propomos têm o mérito de focalizar elementos distintivos do contorno de frases entoacionais. Em outras palavras, os eventos tonais que descrevemos a partir de F0 se definem pela relação que estabelecem com esse domínio prosódico de modo a constituir representação fonológica dos contornos entoacionais em questão. Tendo em vista a abordagem da entoação que assumimos, não são apontadas variações fonéticas que, apesar de serem relevantes no que diz respeito à caracterização do sinal acústico, não são distintivas no plano fonológico para a representação dos contornos.

As sentenças de cada um dos pares que constituem o material de pesquisa se caracterizam por terem a mesma cadeia segmental, mas por serem diferentes estruturalmente quanto à organização prosódica em *Is*. Os quatro pares de sentenças são apresentados no Quadro (2):

³ Pesquisa financiada pela FAPESP (Proc. 2014/24778-3).

⁴ O procedimento experimental consistirá de: (i) apresentação do estímulo auditivo formado por sentenças com configurações tonais que seguem os padrões prosódicos de sentenças neutras do Português Brasileiro, nas quais, comprovadamente do ponto de vista acústico, não há produção de pausas nas fronteiras relevantes, (ii) para que, imediatamente após a apresentação de cada sentença que forma o *input*, os participantes, por meio da tarefa de escolha forçada, respondam se ouvem ou não pausas nas posições indicadas no estímulo visual igualmente apresentado durante o procedimento experimental.

Quadro 2: Pares de sentenças que compõem o material de análise.

Identificação do Par	Sentenças (correlato gráfico)	Segmentação prosódica em <i>I</i>
Par A	1: Não, quero ler.	[não]I [quero ler]I
	2: Não quero ler.	[não quero ler]I
Par B	1: Aceito, obrigado.	[aceito]I [obrigado]I
	2: Aceito obrigado.	[aceito obrigado]I
Par C	1: Isso só, ele resolve.	[isso só]I [ele resolve]I
	2: Isso, só ele resolve.	[isso]I [só ele resolve]I
Par D	1: Vamos perder, nada foi resolvido.	[vamos perder]I [nada foi resolvido.]I
	2: Vamos perder nada, foi resolvido.	[vamos perder nada]I [foi resolvido.]I

Esses pares de sentenças integraram a propaganda comercial de comemoração aos 100 anos da Associação Brasileira de Imprensa (ABI), intitulada “A vírgula”, veiculada em formato impresso e em formato audiovisual (o vídeo é de domínio público e pode ser acessado por meio do link: <http://www.youtube.com/watch?v=uWKpx5Ls1zg>). Para a análise, no entanto, as sentenças foram gravadas em laboratório, uma vez que, em primeiro lugar, o áudio do vídeo conta com fundo musical simultâneo à realização das sentenças e, em segundo lugar, a gravação em ambiente de laboratório permite controle de parâmetros acústicos para o desenvolvimento do procedimento experimental. A gravação foi realizada por um locutor de gênero feminino, de 21 anos, graduando em Letras, nascido e residente em cidade do noroeste paulista.

No que se refere ao modo como os pares de sentenças constituem o anúncio publicitário, a propaganda explora o jogo de possibilidades de realizações fonéticas e de interpretações possíveis dos referidos pares de sentenças. Esses pares de sentenças, apesar de contar com a mesma sequência de segmentos são estruturalmente diferentes pelo uso da vírgula. Considerando os diferentes modos de segmentar a sentença por meio da vírgula, interpretamos que o que está em jogo para as diferentes sentenças – que têm os sentidos contrastados para cada par – é a configuração prosódica das sentenças em *Is*, o que implica o modo como a entoação se realiza para cada sentença em contraste.

No presente artigo, apresentamos a descrição entoacional de dois dos quatro pares de sentenças que compõem o *input* do experimento: são o par A e o par B (cf. Quadro 2). A seleção desses pares se justifica pela estrutura comum das sentenças que os constituem: neles, a diferença entre a sentença 1 e a sentença 2 consiste na presença ou na ausência de fronteira de *I* não final (enquanto a sentença 1 de ambos os pares apresenta fronteira de *I* não final, a sentença 2 não a apresenta). Por sua vez, o que diferencia as sentenças 1 e 2 dos pares C e D é

a mudança de posição da fronteira de *I* não final, já que ambas as sentenças apresentam essa fronteira.

A descrição que fazemos dos pares selecionados na seção de análise tem como objetivo mostrar, do ponto de vista da entoação, como essas sentenças se diferenciam no que se refere à distribuição dos eventos tonais em relação às diferentes estruturas prosódicas comparadas no estudo. Com os resultados obtidos nessa descrição, fazemos dois movimentos: (i) discutimos as decisões metodológicas necessárias para garantir boas condições de realização de experimento de percepção, explorando a detecção de contextos fonético-fonológicos que poderiam pôr à prova os resultados a serem encontrados no experimento que realizaremos em etapa posterior; e (ii) comprovamos o que se realiza em termos acústicos nas sentenças comparadas que as difere e as assemelha a fim de obter uma representação fonológica da realização de F0 nos pares de sentenças em contraste. Essa sistematização nos permitirá, ainda, fazer correlações com os resultados obtidos no teste de percepção que faremos em futuro próximo. Fica, pois, registrado o percurso da investigação da relação entre produção e percepção de F0 na configuração de frases entoacionais.

A gravação das sentenças foi realizada por meio do software *Sound Forge* versão 8.0 em cabine de isolamento acústico nas dependências do Laboratório de Fonética, da UNESP, campus de São José do Rio Preto⁵. Por sua vez, a análise acústica dos padrões entoacionais das sentenças em cada par foi realizada por meio do programa PRAAT.

4. Discussão dos resultados

Na descrição do contorno entoacional de sentenças do Português Brasileiro, a distribuição dos tons associados às *ws* da cadeia segmental em questão depende (i) da posição ocupada por elas no interior de *I* e (ii) do número de sílabas pretônicas que as constitui (cf. TENANI, 2002; FERNANDES, 2007). Nos pares de sentenças analisados neste trabalho, esses fatores são de fundamental importância, uma vez que, considerando os diferentes fraseamentos de *I* para a mesma cadeia segmental, a posição das *ws* é reorganizada no interior das *Is* de acordo

⁵ O Laboratório de Fonética (LabFon) está vinculado ao Departamento de Estudos Linguísticos e Literários do Instituto de Biociências, Letras e Ciências Exatas (IBILCE) da UNESP. O LabFon compreende uma sala de aula e um estúdio de gravação e análise de fala. No estúdio, há isolamento acústico, microcomputador com mesa de som de 10 canais, um decibímetro e microcomputadores para realização de análises de áudio de fala. Mais detalhes sobre o LabFon se encontram em: <http://www.ibilce.unesp.br/#!/departamentos/estudos-ling-literarios/laboratorio-de-fonetica/>

com suas configurações distintas e, por consequência, observamos que as sílabas pretônicas assumem papel de relevância para o estabelecimento de contrastes nos padrões entoacionais comparados relativos aos diferentes fraseamentos de *I*.

Essas observações gerais são detalhadas na descrição acústica dos pares selecionados. São eles: [aceito]I [obrigado] *versus* [aceito obrigado]I e [não]I [quero ler] *versus* [não quero ler]I, descritos nessa ordem. Iniciamos pelo par cuja descrição acústica é menos complexa em termos de atribuição de eventos tonais e, em seguida, descrevemos o par de maior complexidade para discutir de modo mais detalhado sua configuração entoacional.

Consideremos o par apresentado em (3):

- (3)
 a. [aceito]I [obrigado]I
 b. [aceito obrigado]I

Na realização fonética de (3.a), formado por duas *I*s, encontramos acento tonal H+L* associado à sílaba tônica da primeira e da segunda *I* e tom de fronteira L% associado à fronteira final de cada *I*. De acordo com Frota & Vigário (2000), Tenani (2002), Fernandes (2007) e Serra (2009), o padrão entoacional de uma sentença neutra se caracteriza pela associação de um tom L* ou L+H* à fronteira esquerda de *I* e um tom H+L* à fronteira direita de *I*, marcando, assim, o início e o fim da sentença, respectivamente, pela subida e pela queda de F0. Os referidos trabalhos, realizados à luz da Fonologia Entoacional, mesmo quadro teórico assumido neste trabalho, encontraram esse resultado comum por meio de investigações tanto de amostras compostas por dados de fala controlada (é o caso dos trabalhos de FROTA & VIGÁRIO, 2000; TENANI, 2002 e FERNANDES, 2007) quanto por amostras compostas por dados de fala espontânea e de leitura (é o caso do trabalho de SERRA, 2009).

No caso da realização fonética de (3.a), o acento tonal H+L* encontrado nas duas *I*s em questão corresponde ao evento tonal que caracteriza o fim de sentenças neutras no Português Brasileiro conforme os resultados dos referidos trabalhos. Tanto na primeira *I* quanto na segunda *I*, o acento tonal H+L* está alinhado à sílaba tônica da única ω pertencente à ϕ a que se associa, que, por extensão, é a única ω da *I* a que pertence. Dessa hierarquização de categorias prosódicas numa única palavra, resulta que todo o contorno entoacional se restringe a um único evento tonal de sequência HL por dois motivos centrais: (i) tal sequência representa a principal informação prosódica de uma sentença neutra do Português Brasileiro: justamente

a queda que marca seu fim, tendo em vista que o principal acento de uma *I* em português é aquele que se encontra mais à direita no interior do domínio desse constituinte; (ii) outros eventos tonais não ocorrem à esquerda desse evento tonal principal devido à falta de sílabas às quais poderiam se associar.

A associação tonal descrita para a estrutura (3.a), em análise, é exemplificada na representação apresentada em (4) e na respectiva figura (1).

(4)
 [[[aceito) ω] ϕ]*I* [[(obrigado) ω] ϕ]*I*]U
 | | | |
 H+L* L% H+L* L%

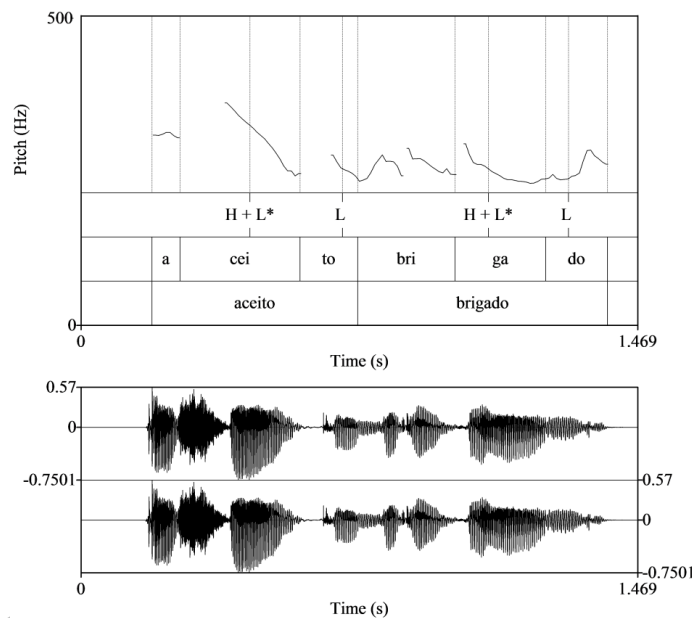


Figura 1: F0 de [aceito]I [obrigado]I.

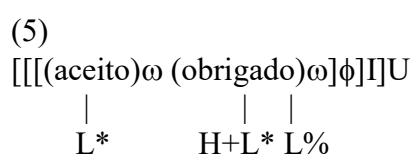
É possível observar, na Figura 1, o alinhamento do o acento tonal H+L* nas duas *I*s. O movimento de queda de F0 que caracteriza o contorno descendente de enunciado afirmativo ocorre no interior de uma única sílaba: a sílaba tônica da única ω que constitui, respectivamente ϕ e *I*. No que se refere ao tom de fronteira, o tom L% é associado à sílaba pós-tônica após a sílaba que recebe o acento tonal principal de cada uma das *I*s: são os casos da sílaba “to” e “do”, na primeira e na segunda *I*, respectivamente. Segundo o quadro teórico-descritivo da Fonologia Entoacional, quando uma sílaba pós-tônica é realizada após o acento tonal mais à direita de uma *I*, a essa sílaba se associa um tom de fronteira: tom semelhante àquele que caracteriza o

último movimento do acento anterior, ou seja, se o movimento é descendente (HL), como é caso das *Is* analisadas em (3.a), atribui-se um tom de fronteira L%; por sua vez, se o movimento é ascendente (LH), atribui-se um tom de fronteira H%. Desse modo, como apresentado em (4), as sílabas “to” e “do” são representadas pelo tom de fronteira L.

Observa-se na Figura 1, no entanto, que, particularmente no intervalo de realização da sílaba “do”, encontra-se uma leve subida de F0. Como já apresentado, a gama de categorias tonais utilizadas pela Fonologia Entoacional é restrita a acento tonal, acento frasal e tom de fronteira, não havendo outras possibilidades de tons senão H e L, simples ou complexos. Não sendo possível caracterizar a subida de F0 na sílaba “do” por meio das categorias tonais disponíveis, o que se conclui é que a referida subida de F0, sob a perspectiva do quadro teórico assumido para análise, não corresponde a um evento que caracterizaria a configuração tonal de *I*. A subida de F0 na sílaba “do” liga-se à articulação fonética da vogal alta [u] que acaba por refletir em F0: um processo no âmbito da microprosódia, porque diz respeito à variação de F0 relacionada à característica dos segmentos.

Consideremos, pois, (3.b), formado por uma única *I*: [aceito obrigado]. Encontramos acento tonal L* associado à primeira ω de *I*, acento tonal H+L* associado à segunda ω de *I* e tom de fronteira L% associado à fronteira final de *I*. Essa distribuição tonal atende ao padrão entoacional de uma sentença neutra do Português Brasileiro, caracterizado por Frota & Vigário (2000), Tenani (2002), Fernandes (2007) e Serra (2009), conforme apresentado na descrição de (3.a). Também em (3.b), o início de *I* é marcado por um tom baixo e o seu término é marcado por um movimento de descendência HL, caracterizando, assim, o início e o fim da sentença, respectivamente, pela subida e pela queda de F0. A diferença de (3.b) em relação a (3.a), anteriormente descrito, reside no fato de que, em (3.b), ocorrem eventos tonais que caracterizam uma única *I* que corresponde a uma sentença neutra do PB (tanto em seu início quanto em seu fim), essa prosodização é diferente em (3.a), já que a mesma sequência segmental da sentença é organizada prosodicamente em duas *Is*.

A representação em (5) e a respectiva figura (2) exemplificam o tipo de associação tonal encontrado para (3.b).



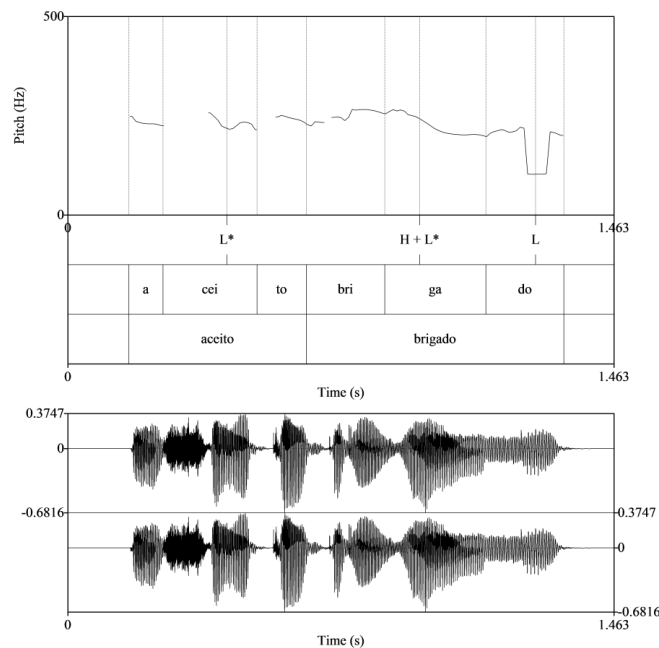


Figura 2: F0 de [aceito obrigado]L.

No que se refere ao tom de fronteira L%, sua presença ocorre devido à realização da sílaba pós-tônica “do”, igualmente ao que ocorre em (3.a). Particularmente, o vale de F0 que se observa na Figura 2, justamente no intervalo de realização da sílaba pós-tônica “do”, não corresponde a uma alteração na configuração tonal, mas decorre da articulação do fonema /d/, cuja oclusão seguida da soltura pode provocar vales, característicos da realização fonética dos segmentos que envolvem obstrução total da passagem de ar.

Na comparação das variações de F0 entre (3.a) e (3.b), observamos que a distinção entre eles reside na alteração de tons associados a “aceito”, primeira ω das I s analisadas. Em (3.a), “aceito”, além de ω , é também uma ϕ , que forma por si só uma I . Em (3.b), “aceito” não forma uma I por si só, pois estabelece relação sintática com a palavra que a segue “obrigado”, formando com ela uma ϕ .

A configuração de “aceito” nas diferentes I s leva em conta os fraseamentos que se estabelecem no eixo sintagmático. Considerando que o Português Brasileiro é uma língua de recursividade à direita, concluímos que: (i) em (3.a), “aceito”, ao ser a única palavra que compõe a I à qual pertence, é caracterizada pela posição que recebe o tom característico da fronteira final de I (H+L); (ii) por sua vez, “aceito”, em (3.b), não ocupa a posição mais à direita da I à qual pertence, pois essa posição é ocupada pelo seu complemento “obrigado” e se alinha

à fronteira esquerda de *I*, recebendo, assim, o tom L* que caracteriza o início de uma sentença neutra (e não o seu fim).

Como mostra a figura (1), a ausência de pausa é evidenciada pelo traço contínuo de F0 na fronteira não final, na qual ocorre sândi vocálico na junção da palavra “aceito” com a palavra “obrigado”. No caso, o processo de sândi em questão se trata de degeminação, caracterizado por Tenani (2002) pela queda da segunda vogal, obrigatoriamente átona, de uma sequência de vogais idênticas. A realização de sândi corrobora a observação de que não há produção de pausa no contexto, pois a realização de pausa bloqueia a aplicação de sândi. Portanto, a diferença entre as sentenças do par ora descrito está ancorada na diferença de configuração de eventos tonais.

Passemos à descrição do par apresentado em (6).

- (6)
 a. [não]I [quero ler]I
 b. [não quero ler]I

No sinal acústico referente à realização de (6.a), observamos acento tonal H+L* associado à sílaba tônica da primeira *I*, acento tonal H* associado à sílaba tônica da palavra cabeça da ϕ da segunda *I* e acento tonal H+L* associado à palavra não cabeça da ϕ da segunda *I*. Desse modo, igualmente ao que ocorre em (3.a), a distribuição dos tons nas duas *Is* em sequência atende à realização tonal das sentenças neutras do Português Brasileiro, sendo o tom H+L* associado à ω mais à direita de cada *I*, característico de final de sentenças neutras.

A representação em (7) e a respectiva figura (3) exemplificam o tipo de associação tonal encontrado para o dado sob análise.

- (7)
 [[[(não) ω] ϕ]I [(quero) ω (ler) ω] ϕ]I]U
 | | |
 H+L* H* H+L*

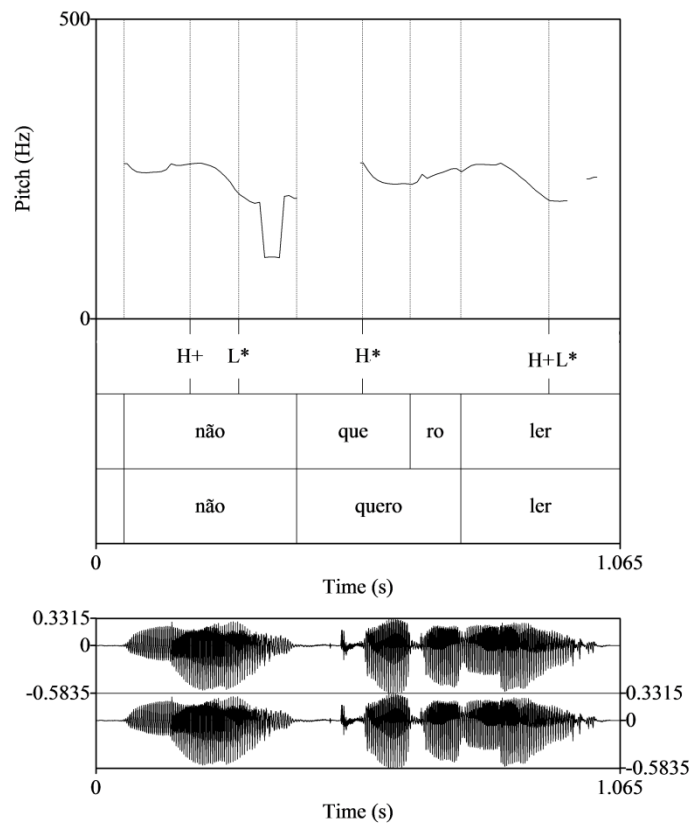


Figura 3: F0 de [não]I [quero ler]I.

Ocorre, no entanto, como é possível verificar na Figura 3, que o contorno de F0 é descontínuo no limite entre as duas *Is*, gerando dúvida sobre a existência de uma pausa. A descontinuidade do contorno de F0 pode, muito provavelmente, estar associada a características acústicas de [k], uma consoante oclusiva, em “quero”, palavra que inicia a segunda *I*. Apesar dessa possibilidade de interpretação para a descontinuidade do contorno de F0, não é possível garantir, pelo sinal acústico obtido, que uma pausa não tenha sido produzida na fronteira entre as duas *Is* em (6.a).

Nosso questionamento sobre a produção ou não de pausa se justifica em função da posição estrutural em que ela pode ter ocorrido. A descontinuidade do contorno de F0 mostrada na Figura 3 ocorre, sintaticamente, no limite entre termo e oração e, fonologicamente, no limite entre duas *Is*. Nessas posições, do ponto de vista linguístico, a produção de pausa é prevista, conforme nos mostra os trabalhos Cagliari (1992) e Chacon e Fraga (2014), com base em arcabouços teóricos distintos.

Para Cagliari, a função da pausa não se limita ao domínio aerodinâmico, uma vez que ela funciona como recurso prosódico ao sinalizar deslocamentos sintáticos e mudanças

semânticas bruscas, seja nos limites de frases, sintagmas e termos, seja no limite de sílabas quando se faz uma silabação. Desse modo, com base em uma perspectiva fonética, a pausa atua como um recurso prosódico que, na sentença em análise, marcaria limite sintático entre “não” e “quero” e sinalizaria que não há relação semântica entre os dois.

Em estudo sobre a realização de pausas na interpretação de atores, Chacon e Fraga (2014) concluíram que a variabilidade de duração das pausas não se caracteriza pela aleatoriedade, mas está condicionada à organização dos constituintes prosódicos – particularmente *I* e *U*⁶. Os autores demonstram haver correlação entre duração de pausa e tipo de fronteira prosódica: maior duração de pausa em limite de *U*; menor duração de pausa em limite de *I*⁷.

Esse último resultado de Chacon e Fraga (2014) interessa de modo particular, uma vez que salienta a possibilidade de produção de pausas na delimitação de *Is*, motivo que nos leva a tomar cuidados metodológicos capazes de garantir que a produção de pausa nos limites a serem testados no procedimento experimental não seja realizada.

Para tanto, uma possibilidade é substituir a sequência segmental “não quero ler”, que tem palavra com o fonema [k], por outra sequência que, tendo a mesma estruturação sintático-prosódica, não contenha uma consoante obstruinte, garantindo, assim, a não interrupção do contorno de F0 na fronteira de *I* não final. No contexto da estruturação da sentença, o verbo “quero” poderia ser substituído por outro verbo iniciado com consoante nasal, como o verbo “mereço”.

Tendo feitas considerações sobre (6.a), tratamos de (6.b) cuja representação em domínios prosódicos é dada em (8), seguida da figura (4), em que se visualiza a associação tonal.

⁶ A relação da produção de pausas com as fronteiras de *I* e *U* não é um resultado aleatório, uma vez que, segundo o modelo de Fonologia Prosódica, a produção de pausas é prevista nos limites desses constituintes pela própria organização do componente fonológico da gramática. Fora dos limites dessas estruturas, a pausa pode ser interpretada como marca de hesitação por interromper constituintes em posições não previstas.

⁷ Os autores mediram a duração de unidades VVs em que ocorreram pausas com os limites de *I* e *U* e verificaram que a maior duração de pausas ocorreu em limites de enunciados fonológicos com média e mediana respectivamente iguais a 1,16s e 1,21s; e a menor duração de pausas ocorreu em limites de frases entoacionais com média e mediana respectivamente iguais a 0,86s e 0,80s. Com base nesses resultados, Chacon e Fraga (2014) interpretam: (i) “a maior duração de pausas em limites de *U* como marca da menor flexibilidade desse constituinte, na medida em que seus limites costumam coincidir com os limites sintáticos de uma sentença” (CHACON e FRAGA, 2014, p. 141) e (ii) “a menor duração de pausas para limites de *I* como marca de sua própria variabilidade/flexibilidade estrutural, que decorre do fato de que *I* integra, em sua constituição, informações de natureza fonológica, sintática e semântica da gramática” (CHACON e FRAGA, 2014, p. 141).

(8)
 [[[(não)ω]φ [(quero)ω (ler)ω]φ]I]U
 | | |
 L* H* H+ L*

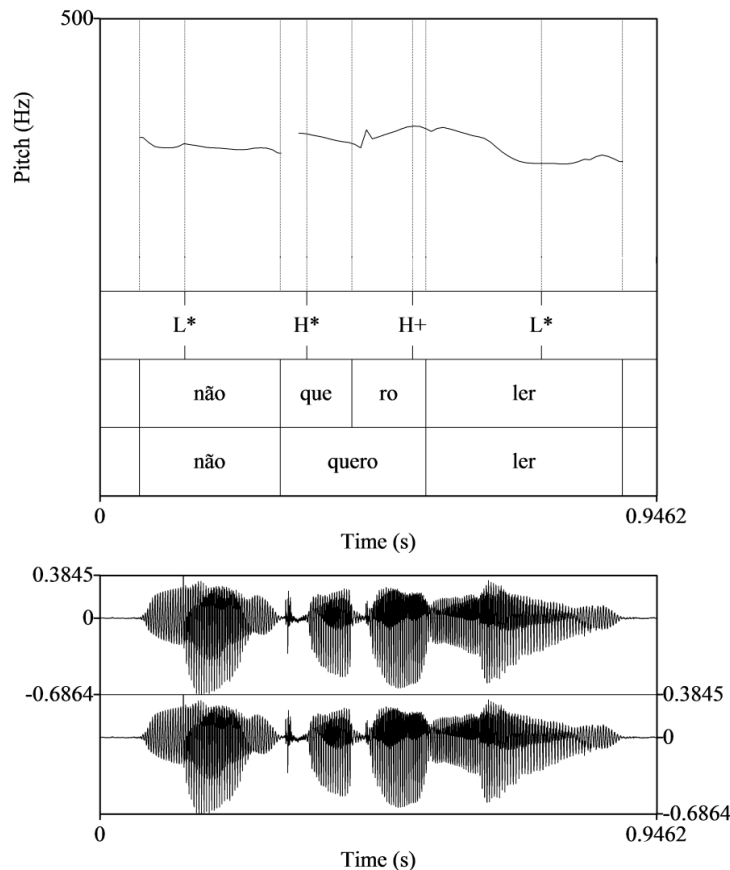


Figura 4: F0 de [não quero ler]I.

A notação (8) representa a associação de um tom L* à sílaba tônica da primeira φ de I, um tom H* alinhado à sílaba tônica da ω cabeça da última φ de I e um tom H + L* que se inicia na sílaba postônica da ω cabeça da última φ de I e se encerra na tônica da ω não cabeça da última φ de I. Segundo Tenani (2002), no que se refere à distribuição tonal, o Português Brasileiro se distingue do Português Europeu por apresentar como característica principal a alternância entre tons baixos e altos. Verificamos, porém, que a distribuição tonal apresentada em (8) não atende a essa característica do Português Brasileiro, uma vez que dois tons H ocorrem lado a lado.

Essa realização dos tons parece relacionada à organização rítmica da sentença em questão, pois detectamos choque acentual entre as sílabas “não” e “que”. Trata-se, pois, de uma

estrutura rítmica que atua de modo decisivo na configuração da altura dos eventos tonais associados às sílabas em contexto de choque acentual. Para o dado em análise, as sílabas acentuadas recebem tons de modo a ficarem em alturas diferentes (cf. discussão sobre tons associados a sílabas em contexto de choque acentual em Tenani, 2002).

Considerando o choque acentual em (6.b) e o contorno de F0 em (6.a), nossa decisão metodológica é, então, a de substituir o par de sentenças [não]I [quero ler] *versus* [não quero ler]. Essa decisão consiste em substituir o referido par de sentenças por outro que atenda à mesma estruturação prosódico-sintática em questão e que apresente (i) uma sequência segmental que garanta continuidade de F0 na fronteira de *I* não final e (ii) uma sequência rítmica sem choque acentual. Propomos, assim, o par apresentado em (9). Passamos a descrevê-lo.

- (9)
 a. [não]I [mereço saber]I
 b. [não mereço saber]

Em (9.a), a segunda *I* se inicia com uma consoante nasal, evitando, assim, a presença de uma oclusiva que causaria dúvida sobre a realização ou não de pausa na fronteira de *I* não final. Em (9.b), as ω s que constituem a cadeia segmental contam com sílabas átonas entre as sílabas tônicas, evitando, assim, o choque acentual.

Na realização fonética de (9.a), formado por duas *I*s, encontramos o tom H+L* associado à sílaba tônica da primeira *I*, o tom L* associado à sílaba tônica do verbo que constitui a ω cabeça da ϕ da segunda *I* e o tom H+L* associado à sílaba tônica da ω não cabeça da ϕ da segunda *I*. Tal distribuição tonal segue o padrão das sentenças neutras do Português Brasileiro por dois motivos que aqui retomamos: a primeira *I*, por ser formada por apenas uma ω , é caracterizada pelo evento tonal que caracteriza o fim de uma sentença neutra, que se realiza pela queda no contorno de F0; a segunda *I*, formada por duas ω s constituídas de, ao menos uma sílaba pretônica, é marcada pela realização tonal padrão de *I* em sentenças neutras, tanto em sua fronteira esquerda quanto em sua fronteira direita, por meio da realização dos tons L* e H+L*, respectivamente (cf. TENANI, 2002). O tom de fronteira L%, também considerado característico de sentenças neutras, não é realizado por não existir sílaba pós-tônica na ω que recebe o acento tonal mais à direita com a qual ele poderia se associar.

A representação em (10) e a respectiva figura (5) exemplificam o tipo de associação tonal para (9.a).

(10)
 [[[(não)ω]φ]I [[(mereço)ω (saber)ω]φ]I]U
 | | |
 H+L* L* H+L*

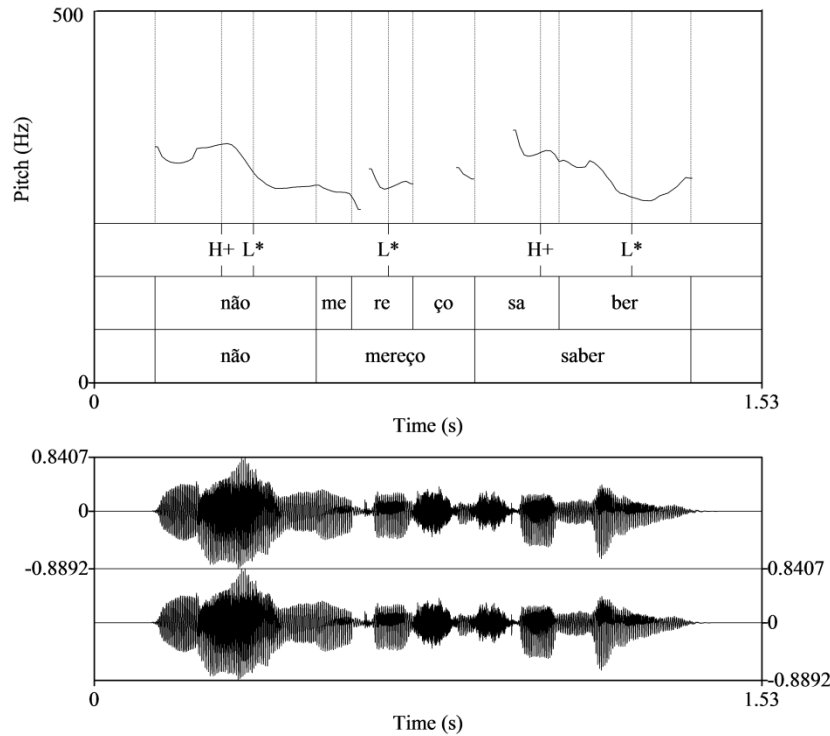


Figura 5: F0 de [não]I [mereço saber]I

Vejamos, pois, a distribuição tonal da cadeia segmental “não mereço saber” de (9.b). Na realização fonética dessa estrutura, formado por uma única *I*, encontramos: acento tonal L* associado à sílaba tônica da primeira φ de *I*; evento tonal H associado à sílaba pretônica da ω cabeça da segunda φ de *I* e acento tonal H+L* associado à sílaba tônica dessa mesma ω cabeça da segunda φ; e acento tonal H+L* associado à sílaba tônica da ω não cabeça da segunda φ de *I*.

A representação em (11) e a respectiva figura (6) ilustram a associação tonal descrita.

(11)
 [[[(não)ω]φ] [(mereço)ω (saber)ω]φ]I]U
 | | |
 L* H H+L* H+L*

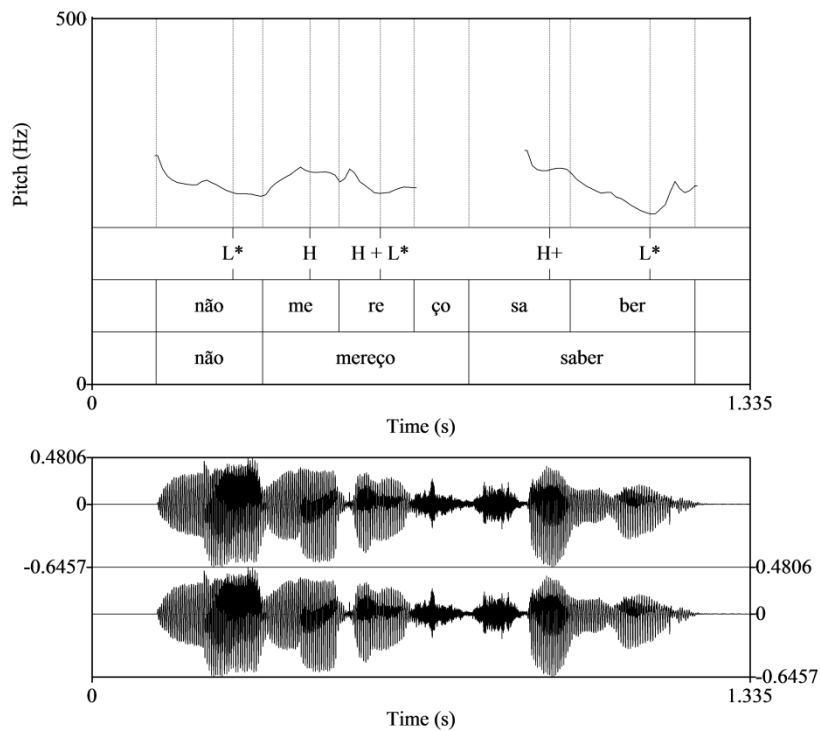


Figura 6: F0 de [não mereço saber]I.

A distribuição tonal de (9.b), além de seguir o padrão de sentenças neutras do Português Brasileiro ao apresentar tom L* associado à sílaba tônica da ω que fica mais à esquerda de I e H+L* associado à sílaba tônica da ω que fica mais à direita de I , mostra como a porção central de I , ocupada pelo verbo “mereço”, adquire eventos tonais por apresentar sílabas pretônicas, diferentemente do que foi observado para “quero”, em (6), (7) e (8), que não têm sílaba pretônica.

Da análise da figura (6), duas possibilidades se abrem para a distribuição de tons entre as sílabas pretônica e tônica desse verbo. A primeira delas é identificar um único evento tonal associado ao verbo: o acento tonal bitonal H+L*, estando H alinhado à sílaba pretônica “me” e L* alinhado à sílaba tônica “re”. Desse modo, interpretaríamos que o movimento de queda de F0 se inicia na sílaba pretônica para se completar na sílaba tônica do verbo. Por sua vez, a segunda possibilidade é identificar dois eventos tonais associados ao verbo: o evento tonal monotonal H associado à sílaba pretônica e o acento bitonal H+L* associado à sílaba tônica. Nessa interpretação, o tom se mantém alto em toda a realização da sílaba pretônica, enquanto o movimento de queda ocorre por completo no interior da sílaba tônica.

Como se pode observar na notação apresentada em (11) e na respectiva figura (6), assumimos a segunda interpretação em eventos tonais da configuração de F0. Essa opção se

justifica pela comparação do contorno de F0 entre (9.a) e (9.b). Enquanto em (9.b) “mereço” é caracterizado por tom relativamente alto quando comparado ao tom associado a “não” no interior da mesma *I*, em (9.a), o verbo é caracterizado por um tom baixo, de início de enunciado neutro, em relação ao acento H+L* que finaliza a *I* anterior. Desse modo, para representar a distinção que se mostra nos tons associados à sílaba tônica e à sílaba pretônica de “mereço” em cada estrutura prosódica analisada, assumimos que: em (9.b), o evento tonal H se alinha à sílaba pretônica do verbo e o acento tonal H+L* se associa à sílaba tônica, em contraste com (9.a), em que um acento tonal L* se associa à sílaba tônica do verbo, havendo, assim, oposição entre os tons associados ao verbo em cada uma das estruturas. Caso interpretássemos que, em (9.b), houvesse acento tonal H+L* associado ao verbo, estando H alinhado à sílaba pretônica e L* alinhado à sílaba tônica, perderíamos a distinção entre as estruturas em contraste (9.a) e (9.b) no que se refere ao tom alinhado à sílaba tônica do verbo (já que, em ambos, haveria L* alinhado a essa sílaba).

Concluimos, portanto, que, na comparação de F0 entre ambas as estruturas prosódicas de (9), a distinção entre elas reside:

- (i) na oposição dos eventos tonais associados à primeira palavra de *I* a depender da configuração dessa *I*: “não” recebe tom prototípico de **fim** de enunciado neutro (H+L*) em (9.a) e recebe tom prototípico de **início** de enunciado neutro (L*) em (9.b)⁸;
- (ii) na oposição dos eventos tonais associados às sílabas pretônica e tônica do verbo a depender da posição que ocupa nas diferentes configurações de *I*: “mereço” recebe acento tonal L* alinhado à sílaba tônica de (9.a), caracterizando o início do enunciado neutro [mereço saber]*I*, enquanto, em (9.b), recebe evento tonal H alinhado à sílaba pretônica e acento tonal H+L* alinhado à sílaba tônica, caracterizando, assim, as alternâncias H L H L do enunciado, já que as fronteiras esquerda e direita são marcadas pelo tons característicos de início e de fim de enunciados neutros (L* associado a “não”, H+L* associado a “saber”).

⁸ Como apresentado ao longo da análise, o padrão entoacional de sentenças do Português Brasileiro a que nos referimos é descrito em trabalhos anteriores que descreveram essa variedade do português com base nas formulações da Fonologia Prosódica, de Nespor e Vogel (1986) e da Fonologia Entoacional, de Ladd (1996). Para mais detalhes sobre as descrições realizadas, bem como sobre as especificidades de cada trabalho, conferir Frota & Vigário (2000), Tenani (2002), Fernandes (2007) e Serra (2009).

5. Considerações finais

A análise dos dados, levando em conta a descrição acústica de F0 e o fraseamento das sentenças em *I*s, permite concluir que, de modo geral, as sentenças que formam os pares analisados têm diferenças significativas na atribuição de tons.

Essas diferenças se apresentam nos tipos de eventos tonais associados às ω s que iniciam cada sequência segmental analisada, a depender do número de ω s que as segue e da posição que elas ocupam nos diferentes fraseamentos de *I*. Se a *I* é formada por uma única ω , à sua sílaba tônica será associado o acento tonal H+L*, característico do fim de uma sentença neutra (como é o caso de “aceito” e “obrigado” no fraseamento (3.a) – [aceito]I [obrigado]I – e, também, como é o caso de “não” no fraseamento (9.a) – [não]I [mereço saber]I). Se a *I* é formada por mais de uma ω , o acento tonal L* ou L+H*, característico do início de uma sentença neutra, é associado à primeira ω da sequência, já que a posição mais à direita de *I* é ocupada por outra ω a cuja sílaba tônica se associará o acento tonal H+L* que caracteriza o fim da sentença (como são os casos de “aceito” no fraseamento (3.b) – [aceito obrigado]I – e de “não” no fraseamento (9.b) – [não mereço saber]I.). Por meio da observação dessas regularidades, nossos resultados fazem valer o esclarecimento dado por Frota (2000) de que a associação tonal condiciona a realização tonal, uma vez que a ancoragem do tom é feita tendo como base a posição prosódica proeminente.

Ainda foram observadas as interferências do número de sílabas de ω s na ocorrência e na configuração dos eventos tonais. Por meio do par apresentado em (9.a) e (9.b), mostramos que o contraste se manifesta por meio dos eventos tonais diferentes que estão associados às sílabas pretônica e tônica de ω s a depender da posição que ocupam no interior de *I*. No caso analisado, a forma verbal “mereço” recebeu acento tonal L* em (9.a), característico de contorno inicial de *I*, por ser a primeira ω da *I* à qual pertence, mas recebeu H* em sua sílaba pretônica e H+L* em sua sílaba tônica devido ao fato de ocupar a posição medial da *I* [não mereço saber], na qual os acentos iniciais e finais estão associados, respectivamente à primeira e à última ω . Um estudo controlado da variável relativa ao número de sílabas não é objeto desta pesquisa, mas se faz relevante observar esse fator na constituição de experimentos sobre entoação.

Dos resultados apresentados, concentramos nossa atenção ao que, de mais geral, eles nos levam a concluir: aos diferentes significados dos enunciados estão associados diferentes fraseamentos de *I*, os quais, por sua vez, levam a diferentes configurações entoacionais

identificadas por meio da descrição acústica realizada. Desse modo, a pausa – fenômeno prosódico alvo dos testes de percepção que serão desenvolvidos nos encaminhamentos futuros da pesquisa – não é, portanto, o recurso prosódico único a que estão associados os diferentes fraseamentos de *I*, conforme apresentam alguns trabalhos como Gelamo (2006)⁹ e Chacon e Fraga (2014).¹⁰ Os resultados do presente trabalho levam à conclusão de que a pausa apenas “realça” uma distinção que se mostra na própria configuração tonal, não sendo, portanto, uma realização fonética necessária nas fronteiras de *I*.

Feitas essas considerações, acreditamos ter apresentado as bases para discutir, em etapa futura de nossa pesquisa quando desenvolvermos experimento de percepção, como se dá a percepção pelo ouvido humano das diferenças entoacionais encontradas nas estruturas comparadas. Haveria pausas que são apenas percebidas sem haver necessariamente input fonético?...

Referências Bibliográficas

BECKMAN, M.; PIERREHUMBERT, J. Intonational Structure in Japanese and English. *Phonology Yearbook*, 1986, n. 3, p. 255-310.

CAGLIARI, L. C. Prosódia: Algumas funções dos supra-segmentos. *Caderno de Estudos Linguísticos*, 1992, n. 23, p. 137-151.

CHACON, L.; FRAGA, M. Pausas na interpretação teatral: delimitação de constituintes prosódicos. *Filologia e Linguística Portuguesa*, 2014, v. 16, n. 1, p. 121-146. **crossref** <http://dx.doi.org/10.11606/issn.2176-9419.v16i1p121-146>

D'IMPERIO, M.; ELORDIETA, G.; FROTA, S.; PRIETO, P.; VIGÁRIO, M. Intonational Phrasing in Romance: The role of prosodic and syntactic structure. In: FROTA, S., VIGÁRIO, M.; FREITAS, M. J. (Eds.). *Prosodies: with special reference to Iberian Languages*. Phonetics & Phonology Series. Berlin: Mouton de Gruyter, 2005, p. 59-97.

⁹ Gelamo (2006) analisa como se dá a organização prosódica de um texto cantado em quatro diferentes interpretações. Analisando particularmente a frase entoacional, o trabalho conclui que a diferente organização desse constituinte, nas interpretações de canções, pode provocar diferenças de atribuição de sentido para o texto das canções – daí, a aproximação do trabalho da autora com o que ora apresentamos.

¹⁰ No caso de Gelamo (2006), a análise dos diferentes fraseamentos de *I* em quatro interpretações de uma mesma canção é explicada em termos de produção de pausas na delimitação desse constituinte, o que leva a autora a concluir que as pausas funcionam como “diferentes modos de organização prosódica da formulação linguística de uma mesma canção” (Gelamo, 2006, p. 48). No caso de Chacon e Fraga (2014), a importância da produção de pausa é tratada de modo análogo na interpretação de atores, com a diferença de que, nesse último trabalho, as pausas são explicadas mais por seu caráter de regularidades da língua, do que por questões ligadas à expressividade no trabalho de interpretação.

FERNANDES, F. R. **Ordem, focalização e preenchimento em português: sintaxe e prosódia**. Campinas: 2007. Tese (Doutorado em Linguística). Instituto de Estudos da linguagem, Universidade Estadual de Campinas, Campinas.

FROTA, S. **Prosody and focus in European Portuguese**. Lisboa: 1998. Tese (Doutorado em Linguística). Universidade de Lisboa, Lisboa. Publicado por Garland Publishing (Serie Outstanding Dissertations on Linguistics). NewYork/London, 2000.

FROTA, S. Questões de associação e alinhamento tonal: implicações para uma teoria da entonação. In: CASTRO, R. V. & BARBOSA, P. (Eds.). **Actas do XV Encontro Nacional da Associação Portuguesa de Linguística**, v. 1. Coimbra: APL, 2000, p. 513-532.

FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. In: CASTRO, R. V.; BARBOSA, P. (Eds.). **Actas do XV Encontro Nacional da Associação Portuguesa de Linguística**, v. 1. Coimbra: APL, 2000, p. 533-555.

GELAMO, R. P. **Organização prosódica e interpretação de canções: A frase entonacional em quatro diferentes interpretações de Na batucada da vida**. São José do Rio Preto: 2006. Dissertação (Mestrado em Estudos Linguísticos). Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto.

LADD, D. R. **Intonational Phonology**. Cambridge: CUP, 1996.

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

PIERREHUMBERT, J. **The phonology and phonetics of English intonation**. Tese de Doutorado. Massachusetts: M.I.T. Press, 1980.

_____; BECKMAN, M. **Japanese tone structure**. Cambridge, Massachusetts : M.I.T Press, 1988.

SERRA, C. R. **Realização e percepção de fronteiras prosódicas no português do Brasil: fala espontânea e leitura**. Rio de Janeiro: 2009. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade Federal do Rio Janeiro, Rio de Janeiro.

TENANI, L. E. **Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos**. Campinas: 2002. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.

_____; FERNANDES-SVARTMAN, F. R. 2008. Prosodic phrasing and intonation in neutral and subject-narrow-focus sentences of Brazilian Portuguese. **Proceedings of Fourth Conference on Speech Prosody**, 2008. Campinas: RG/CNPq.

Artigo recebido em: 27.10.2015

Artigo aprovado em: 13.02.2016

Análise perceptiva e acústica em fonética forense: uma pesquisa em disfarce de voz

Perceptive and acoustic analysis in forensic phonetics: research in voice disguise

Maria Lúcia de Castro Gomes*

Denise de Oliveira Carneiro**

Andrea Alves Guimarães Dresch***

RESUMO: O objetivo maior deste trabalho de pesquisa foi introduzir um grupo de alunos de graduação em Letras nas atividades de análise perceptiva e acústica em contexto forense. Para isso um texto de 75 palavras foi criado para simular uma situação de sequestro. Cinquenta participantes, divididos em grupos de dez de acordo com gênero e faixa etária, foram gravados lendo o texto em voz normal e, em seguida, disfarçando a sua voz. Para a análise perceptiva, foi utilizado o protocolo VPAS (LAVER, 1980; CAMARGO E MADUREIRA, 2008), realizada por seis pessoas (juízes). Para a análise acústica, usou-se o software PRAAT para medição de duração, F0, F1 e F2. Vozes femininas e masculinas foram comparadas e alguns resultados confirmam algumas tendências universais, como: a duração intrínseca das vogais, na qual as vogais mais baixas são mais longas que as vogais mais altas; informantes do sexo feminino apresentam valores de frequência de F0, F1 e F2 mais altas que os do sexo masculino; as mulheres apresentam espaço vocálico maior que os homens. Os ajustes mais frequentes para o disfarce de voz foram: a mandíbula aberta e a protrusão labial, como ajustes supralaríngeos, o falsete e a voz áspera, como ajustes de fonação.

PALAVRAS-CHAVE: Análise perceptiva. Análise acústica. Fonética forense. Disfarce de voz. Análise de vogais.

ABSTRACT: The main objective of this research was to introduce a group of undergraduate students in activities of perceptual and acoustic analysis in a forensic context. A 75-word text was created to simulate a kidnapping situation. Fifty participants, in groups of ten according to gender and age, were recorded reading the text in a normal voice and then disguising their voice. For the perceptive analysis, we used the VPAS protocol (LAVER 1980; CAMARGO AND MADUREIRA, 2008), and six researchers (judges) analyzed the voices. For the acoustic analysis, we used PRAAT for measuring duration, F0, F1 and F2. Male and female voices were compared and some results confirm some universal trends: vowel intrinsic duration, in which the lower vowels are longer than the higher vowels; female informants have F0, F1 and F2 higher frequency values than males; women have larger vowel spaces than men. The most common settings for voice disguise were: open jaw and lip protrusion, as supralaryngeal settings, falsetto and harsh voice, as phonation settings.

KEYWORDS: Perceptive analysis. Acoustic analysis. Forensic phonetics. Voice disguise. Vowel analysis.

* Doutora em Letras, Linguista, Professora no Departamento de línguas Estrangeiras Modernas – DALEM da UTFPR Câmpus Curitiba.

** Mestranda em Engenharia Biomédica, Fonoaudióloga, Perita Criminal no Instituto de Criminalística do Paraná – IC-PR.

*** Mestre em Engenharia Biomédica, Engenheira Eletricista, Perita Criminal no Instituto de Criminalística do Paraná – IC-PR.

1. Introdução

À medida em que as atividades e relações humanas adquirem maior complexidade, os estudos interdisciplinares tornam-se mais necessários. A linguística vem estabelecendo parcerias com diversas áreas de conhecimento para melhor compreensão dos variados processos que envolvem a linguagem. Para possibilitar melhor entendimento sobre a aquisição da linguagem, a linguística e a psicologia se associam e assim nasce a psicolinguística. Aliando-se à sociologia, a linguística consegue explicar fatos de variação e mudança na língua, e dessa forma surge a sociolinguística. Os pesquisadores do discurso vão buscar na psicanálise e no marxismo noções importantes sobre a constituição do sujeito e da produção de sentidos, e mais uma área importante na linguística é criada – a análise do discurso.

Com os avanços tecnológicos e o desenvolvimento da tecnologia da fala para, por exemplo, reconhecimento de voz e comparação de locutor, novas parcerias têm sido formadas. Os estudiosos da linguagem veem a necessidade, mais do que nunca, de ultrapassar suas áreas de pesquisa e trabalhar com estudiosos de outras áreas, como as engenharias, as ciências da computação, a fonoaudiologia, a psicologia.

Diversos ramos da linguística aplicada, a partir de tais parcerias, têm se dedicado ao estudo das particularidades da fala com objetivos diversos. Dentre essas áreas está a ciência forense, que tem por objetivo atribuir autoria a falas por meio de comparação. Segundo Braid (2003), “a verificação de locutor é a atividade pericial dentro da Fonética forense capaz de determinar se duas falas foram produzidas por um mesmo falante” (p.6). O exame consiste “na comparação de uma gravação de voz suspeita com um padrão de voz coletado do provável autor da voz suspeita” (ESPINDULA, 2005, p. 338).

Sob a ótica forense, a fala é analisada como um sinal biológico eletronicamente processado, armazenado em mídia. Isso significa que os estudos não podem ser restritos à produção da fala, mas devem se expandir ao processamento do sinal. Por isso, diferentes métodos são empregados para a análise, sem ainda haver um consenso sobre qual seria o mais eficaz, conforme pesquisa de Gold e French (2011), que agruparam os métodos em quatro possibilidades: análise perceptiva somente, análise acústica somente, combinação entre análise perceptiva e acústica, e análise automática para reconhecimento de locutor. Tem sido crescente o uso de sistemas automáticos e muita discussão sobre a utilização de tais sistemas tem acontecido, principalmente em eventos como a Conferência Anual da Associação para Fonética e Acústica Forense (IAFPA, 2015).

Este trabalho, porém, concentra-se nos métodos de análise perceptiva e acústica, escolhidos por um grupo de pesquisa organizado com o intuito de introduzir alunos de um curso de graduação em Letras nos mistérios da fonética forense. O objetivo da pesquisa foi comparar vozes de 50 falantes, distribuídos em grupos de acordo com gênero e faixa etária, em simulação de um telefonema para pedido de resgate durante uma situação de sequestro, em suas vozes normais e usando estratégia de disfarce.

O objetivo deste texto é, primeiramente, compor uma revisão dos conceitos básicos e fundamentais sobre qualidade de voz e sobre alguns dos principais parâmetros acústicos para análise de vogais: duração, primeiro e segundo formantes e frequência fundamental. Em segundo lugar, descreve-se a metodologia da pesquisa, os principais resultados da análise acústica, e a relação entre a análise perceptiva dos disfarces com a análise acústica das vogais.

2. Pressupostos Teóricos

Este trabalho teve três focos de análise: a qualidade vocal, os parâmetros acústicos das vogais do português brasileiro e o disfarce de voz.

2.1 Qualidade Vocal

Muito pode se dizer de uma pessoa a partir de sua voz. A voz pode demonstrar características físicas e psicológicas, estados emocionais e de humor. Dada a diversidade de perspectivas de análise, definir voz ou qualidade de voz tem sido objeto de discussão. A definição de qualidade de voz de Abercrombie (1967) é citada por Laver (1980, p.1) e por Kreiman e Sidtis (2013, p.8) e indica “as características que estão mais ou menos presentes o tempo todo em que a pessoa está falando; trata-se da qualidade quase permanente que perpassa todo som que sai de sua boca”¹.

Kreiman e Sidtis (2013) apontam para a dificuldade de atribuir uma definição de voz que sirva a todos os propósitos, e que possa servir a uma variedade de disciplinas. Segundo as autoras, a definição de voz pode ter um sentido estreito ou amplo. Um som produzido pela vibração das pregas vocais é uma definição estreita do termo voz, pois exclui os efeitos acústicos do trato vocal. Definir voz dessa forma aproxima-se ao que os linguistas definem como traço

¹ “those characteristics which are present more or less all the time that a person is talking; it is a quasi-permanent quality running through all the sound that issues from his mouth” (ABERCROMBIE, 1967, *apud* LAVER, 1980 e KREIMAN; SIDTIS, 2013). Todas as traduções neste artigo foram realizadas pela primeira autora.

de vozeamento, que faz com que /z/ seja diferente /s/, por exemplo. Nesse caso, “voz seria sinônimo do termo ‘fonte laríngea’, que enfatiza o fato de que as vibrações das pregas vocais são a fonte de energia acústica para grande parte da fala”². Num sentido amplo, voz seria sinônimo de ‘fala’, pois incluiria “os resultados acústicos da ação coordenada entre o sistema respiratório, língua, mandíbula, lábios e palato mole”³.

A *qualidade de voz*, para Kreiman e Sidtis (2013), também pode ser definida de forma estreita ou ampla, o que seriam na verdade dois lados de uma mesma moeda – de um lado, características da percepção e, de outro, características da produção. No sentido estreito, a qualidade de voz é vista como um aspecto único do processo de fonação, como a percepção da corrente de ar no sinal vocal. Num sentido não tão estreito, é tida como o resultado percebido do processo de fonação. E num sentido mais amplo, é a resposta do ouvinte ao som da fala, ou seja, conforme Denes e Pinson (1993, *apud* KREIMAN; SIDTIS, 2013), a qualidade de voz é o resultado final de uma sequência complexa de eventos acústicos, psicológicos e cognitivos, a conhecida ‘cadeia da fala’. O foco que se dá a essa cadeia de eventos depende do interesse do pesquisador ou profissional envolvido, se um cirurgião, um engenheiro, um psicólogo, cada um vai adotar uma definição diferente e ter uma preocupação específica.

Kreiman e Sidtis (2013) também discutem questões sobre a distinção entre produção e percepção da voz, que têm sido mantidas separadas nos trabalhos de pesquisa. No entanto, defendem uma perspectiva adotada em trabalhos mais recentes no dialogismo dentro da linguística, na qual percepção e produção são inseparáveis. A voz é um evento concreto produzido por uma pessoa num contexto comunicativo que, necessariamente, inclui o ouvinte, mas também manifesta a consciência abstrata e não observável do falante. As autoras concluem, afirmando que “a voz expressa quem nós somos, tanto de forma isolada como em relação a outros indivíduos”⁴.

De acordo com Behlau (2004), que também afirma que a terminologia é bastante imprecisa e confusa, a qualidade vocal é um termo empregado para denominar a impressão total criada por uma voz, é “nossa avaliação perceptiva principal”. Behlau e Pontes (*apud* BEHLAU,

² “Voice in this sense is also synonymous with the term ‘laryngeal source’, which emphasizes the fact that vocal fold vibrations are the acoustic energy source for much of speech” (KREIMAN; SIDTIS, 2013, p.5).

³ “the acoustic results of the coordinated action of the respiratory system, tongue, jaw, lips, and soft palate” (KREIMAN; SIDTIS, 2013, p. 6).

⁴ “voice expresses who we are, both in isolation and with respect to other individuals” (KREIMAN; SIDTIS, 2013, p.10)

2004) estipulam 23 tipos de vozes, das quais podem ser extraídas, perceptivamente, as características de fenômenos biológicos, psicológicos e socioeducacionais.

Rejeitando o sentido estreito de qualidade de voz, apenas como uma atividade laríngea, Laver (1980) adota um sentido amplo, como a ‘coloração característica auditiva’ da voz de uma pessoa. Para o autor, não somente as características laríngeas, mas também as supralaríngeas são importantes para a qualidade da voz. Inspirando-se em Abercrombie (1967, *apud* LAYER, 1980) e o conceito de ajuste muscular (*muscular adjustment*), e em Honikman (1964, *apud* LAYER, 1980), que traz o termo ajuste articulatório (*articulatory setting*), Laver propõe uma taxonomia descritiva baseada em componentes fonéticos da qualidade de voz. O sistema descritivo proposto apresenta-se com uma fundamentação de base auditiva, mas que se correlaciona com outros níveis de análise, todos passíveis de verificação instrumental, quais sejam, os níveis articulatório, fisiológico e acústico.

A perspectiva fonética da proposta de Laver tem como resultado o roteiro VPAS - *Vocal Profile Analysis Scheme* que, segundo Camargo e Madureira (2008), oferece a possibilidade de estudos a partir de uma noção básica: a plasticidade do aparelho vocal. As autoras adaptaram o protocolo de Laver para o contexto brasileiro, o BP-VPAS (*Brazilian Portuguese Vocal Profile Analysis Scheme*).

Esse protocolo oferece uma perspectiva a partir de unidades analíticas que, segundo Camargo e Madureira (2008), são tendências musculares de longo termo, os chamados ajustes. Tais ajustes podem ser de ordem fonatória, definidos pela vibração das pregas vocais (como o falsete e a voz áspera), ou de ordem supraglotal ou articulatória (como a protrusão labial ou mandíbula aberta) ou, ainda, de tensão, seja laríngea ou do trato vocal. Segundo as autoras, os ajustes são definidos a partir de variações de um ponto neutro de referência, em termos de forças de adução ou tensão longitudinal.

O uso do protocolo BP-VPAS para avaliação de voz permite também a classificação por níveis, escalados de 1 a 6, de moderado a extremo. Como no âmbito forense, a avaliação não é ensejada por queixas vocais ou alterações trazidas pelos locutores, nesta pesquisa, o modelo foi utilizado apenas como roteiro para identificação dos ajustes utilizados nos disfarces de voz, sem considerar-se os níveis escalares, conforme será descrito na seção destinada à metodologia.

Defende-se que a avaliação da qualidade vocal deva ser predominantemente perceptiva. Kreiman e Gerratt (1993) apontam que a qualidade de voz é, por natureza, perceptual, por isso a área clínica valoriza muito mais as medidas perceptivas do que as medidas instrumentais. No

entanto, essa avaliação é cunhada em subjetividade, gerando discordância entre os ouvintes e dificuldade em assumir um consenso em torno do uso desta ou daquela terminologia. Também se defende que a avaliação forense da qualidade vocal, assim como a avaliação clínica, deva ser predominantemente perceptivo-auditiva e, em caso de dúvidas, proceda-se à busca de fenômenos que possam ancorar ou refutar os achados por meio da avaliação acústica. A soberania da avaliação perceptivo-auditiva é citada por pelo menos dois autores, pesquisadores de parâmetros acústicos, indicando a necessidade de corroborar os achados de modo mais objetivo. (DAJER, 2010; MASTER, 2005).

Partimos, então, para a descrição de alguns parâmetros acústicos importantes em análises de comparação de locutor, que serão correlacionados com a análise perceptiva da qualidade de voz.

2.2 Teoria da produção da fala

A teoria acústica da produção da fala descrita por Fant (1960) é conhecida como teoria fonte-filtro, segundo a qual o sinal da fala é resultante da combinação de uma fonte de energia sonora e de filtros resultantes da ressonância da energia da fonte nas cavidades supraglóticas. Segundo tal modelo linear (KENT; READ, 2002), a fonte de energia acústica (dos sons vozeados) é proveniente dos pulsos glóticos, gerados a partir da vibração das pregas vocais situadas na laringe, em movimentos sucessivos de abertura e fechamento. A frequência de tal vibração é medida em Hertz (Hz), constituindo o valor da frequência fundamental (F0) e tem relação direta com a massa, tamanho e geometria das pregas vocais (GILLIER, 2011). As pregas vocais em vibração produzem um espectro, que de acordo com Kent e Read (2002), poderia ser idealizado por frequências discretas, múltiplas da frequência vibratória fundamental, conforme demonstrando na Figura 1.

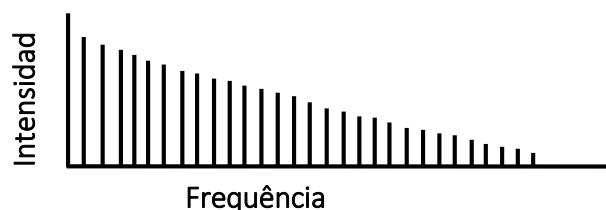


Figura 1: espectro idealizado, no qual a energia está distribuída em frequências discretas em múltiplos inteiros da frequência fundamental (KENT; READ, 2002, p. 23).

Conforme já mencionado acima, as cavidades do trato vocal supralaríngeo atuam como ressoadores e atenuadores (filtros) da energia proveniente da fonte. As frequências realçadas são os chamados formantes. Os “formantes consistem em realces de energia acústica concentrada que representa a ressonância em um tubo acústico (modelo proposto pela teoria fonte-filtro). Sendo relacionados à anatomofisiologia do trato vocal, suas medidas constituem dados relevantes para a individualização de um sinal” (GILLIER, 2011). A partir da fonte produtora da voz e da filtragem pelos articuladores, produzem-se os segmentos que formam as palavras: as vogais e as consoantes. Neste trabalho, o foco se dá nos segmentos vocálicos.

2.3 Vogais e parâmetros acústicos

Segundo Camara Jr. (2002), a divisão mínima na fonologia, a chamada segunda articulação⁵ da língua, “é a dos sons vocais elementares, que podem ser vogais ou consoantes” (p. 33). O autor afirma que há dois critérios para diferenciar vogais de consoantes, um de caráter mais fonético, que considera a vogal como um som produzido pela ressonância do trato vocal, com passagem livre do ar pela boca. Na mesma linha, Marchal e Reis (2012) explicam que “a vogal é produzida com um canal aerífero aberto, sem constricção maior e na ausência de geração de ruído de fricção” (p. 142). Já o segundo critério refere-se ao fato de a vogal constituir-se como núcleo de uma sílaba. Embora as consoantes nasais e líquidas também possam figurar como centro de sílaba, no português apenas vogais podem estar no centro ou ápice, ficando as consoantes como elementos marginais (CAMARA JR., 2002; MARCHAL; REIS, 2012).

As vogais, tidas como os sons mais simples de se analisar e descrever (KENT; READ, 2002), têm como parâmetros principais para descrição acústica a duração, a frequência fundamental, o padrão formântico e o espectro (este não considerado nesta pesquisa).

2.3.1 Duração

Um dos importantes parâmetros acústicos para análise de vogais é a duração, que pode figurar como elemento distintivo nas línguas. No entanto, mesmo na inexistência desse tipo de distinção, as vogais variam substancialmente em sua duração. Lindblom (1967) aborda a variação sistemática da duração da vogal como um problema da fonética clássica. O problema

⁵ André Martinet, linguista do Círculo Linguístico de Praga propôs a teoria da dupla articulação da linguagem, segundo a qual a morfologia é a primeira articulação e a fonologia, a segunda articulação (PIETROFORTE, 2005).

consiste na composição de fatores que exercem influência na duração, quais sejam, a característica da vogal em si e o ambiente consonantal adjacente. No primeiro caso, fatores aerodinâmicos são responsáveis pela diferença – vogais mais abertas tendem a ser mais longas que vogais mais fechadas (duração intrínseca). Quanto ao ambiente fonológico, diz-se que uma vogal é geralmente mais longa antes de consoante vozeada do que precedendo uma consoante não vozeada.

Kent e Read (2002) ilustram esses dois fatores com as palavras *bet*, *bed* e *bad*, da língua inglesa, conforme Figura 2. A vogal em *bed* é mais longa que a vogal em *bet*, pelo vozeamento de /d/, e a vogal de *bad* é mais longa que a de *bed* porque /æ/ é mais baixa, ou seja, mais aberta que /ε/.

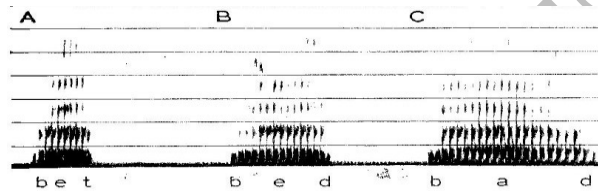


Figura 2: ilustração em espectrogramas de variação na duração da vogal. A – [bet]; B – [bed]; C – [bæd] (KENT; READ, 2002, p. 127).

Estudos de Keating (1985), com o uso de eletromiografia, atestam que línguas como o chinês, o tcheco e o árabe não apresentam essa característica, ou seja, as vogais não se alongam diante de consoantes vozeadas. A autora defende, por conta disso, que essa é uma característica específica de determinadas línguas.

Quanto à duração intrínseca das vogais, Escudero *et al* (2009) analisaram as sete vogais orais do português brasileiro e europeu e concluíram com confiança que as vogais mais baixas são mais longas que as vogais mais altas na língua portuguesa. Afirmam também que o ouvinte de língua portuguesa usa a duração como pista para a identidade da vogal em maior grau do que ouvintes de outras línguas.

Se a duração intrínseca da vogal tem relação com abertura que, segundo Marchal e Reis (2012), para as vogais “descreve normalmente a distância que separa o ponto mais elevado da cúpula da língua até o palato” (p. 137) e que com esse parâmetro é possível distinguir vogais mais altas de vogais mais baixas, como /i/ de /e/, por exemplo, conclui-se que a duração tem estreita relação com o primeiro formante.

2.3.2 Frequência fundamental

A frequência fundamental ou F0 corresponde à taxa de vibração das pregas vocais e é, como visto acima, a fonte de energia para a produção da voz. A F0 é o correlato acústico da vibração das pregas vocais, enquanto o *pitch* é o seu correlato perceptivo (KREIMAN; SIDTIS, 2012). É muito comum, no entanto, o uso do termo *pitch* para se referir à frequência fundamental (NOOTBOOM, 1997), ou o tratamento dos dois termos como sinônimos, mas na verdade, a frequência fundamental é a pista mais importante para a percepção do *pitch* (KREIMAN; SIDTIS, 2012).

Assim como a frequência de formantes, a frequência fundamental também é determinada por questões anatômicas, constituindo em fator importante na caracterização de locutores. Como a F0 de cada indivíduo depende do tamanho e massa das pregas vocais, os fatores sexo e idade também são cruciais. Nootboom (1997) reporta que, para os homens, os valores ficam entre 80 e 200 Hz, para as mulheres, entre 180 e 400 Hz, e para crianças os valores são consideravelmente mais altos. Em pesquisa com o português brasileiro, Madazio (2009) reporta que “para os homens, a faixa da frequência fundamental varia de 80 a 150Hz; e, para as mulheres, de 150 a 250Hz. Para falantes do português de São Paulo, a frequência fundamental média encontrada foi de 113Hz para homens e 205Hz para mulheres” (p. 37). Essa autora ainda argumenta que a F0 é um parâmetro robusto por ser resistente aos diferentes sistemas de análise acústica e menos sensível aos meios de gravação.

Rose (2002) afirma que a frequência fundamental é uma medida extremamente importante para a fonética em geral e para a fonética forense, em particular. Kreiman e Sidtis (2012) também argumentam sobre a importância da F0, afirmando que “a F0 apresenta correlações com as características físicas e o estado interno de um indivíduo, e ainda pode ser muito bem controlada e bastante saliente aos ouvintes”⁶. Essa combinação de fatores, segundo as autoras, é ideal para fornecer informações sobre um locutor, além de ser um parâmetro robusto mesmo com ruído no ambiente. Rose (2002) afirma ainda que a F0 pode ser extraída mesmo de gravações de má qualidade. Além de Rose, outros autores defendem a F0 como um parâmetro robusto para o trabalho de reconhecimento de voz e comparação de locutor (NOLAN, 1983; SAMBUR, 1975; JIANG, 1996, *apud* KINHOSHITA, 2009). Mas também,

⁶ “F0 displays correlations with an individual’s physical characteristics and internal state, but is also finely controllable and highly salient to listeners” (KREIMAN; SIDTIS, 2012, p. 139)

há argumentos em contrário. Kinoshita (2009), por exemplo, não considera a F0 um parâmetro eficaz, por apresentar razão de variância bastante fraca.

Em fonética forense, além da F0, os formantes das vogais também se destacam como parâmetros de grande importância.

2.3.3 Padrão formântico das vogais

Stevens (1997) afirma que “as frequências dos formantes, em particular os dois primeiros, F1 e F2, são dependentes do formato do espaço entre a glote e os lábios, e esse formato, por sua vez, é determinado pela posição do corpo da língua e dos lábios”⁷. O valor de F1 varia com a elevação e abaixamento da língua, ou seja, com o movimento vertical, enquanto F2 varia com o movimento horizontal, para frente e para trás. Kent e Read (2002) advertem para o uso dessa regra porque há exceções, no entanto, descrevem vários experimentos que confirmam a precisão dessa escala multidimensional (RAKERD; VERBRUGGE, 1995; PETERSON-BARNEY, 1952; HILLENBR *et al*, 1995, *apud* KENT e READ, 2002).

Essa relação das medidas acústicas com movimentos articulatorios tem sido simbolizada, a partir de trabalhos de Bell (1879) e Jones (1862), em um espaço vocálico, representado por um trapézio que imita o espaço do trato bucal. A delimitação do espaço é feita pelas chamadas vogais cardeais, que indicam os limites das pronúncias vocálicas, conforme Figura 3. Segundo Marchal e Reis (2012), “o sistema das vogais cardeais representa um instrumento para descrição das vogais com base articulatória e auditiva” (p. 145).

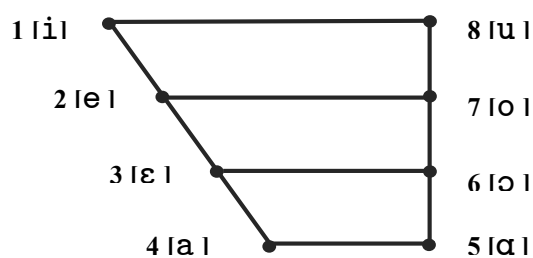
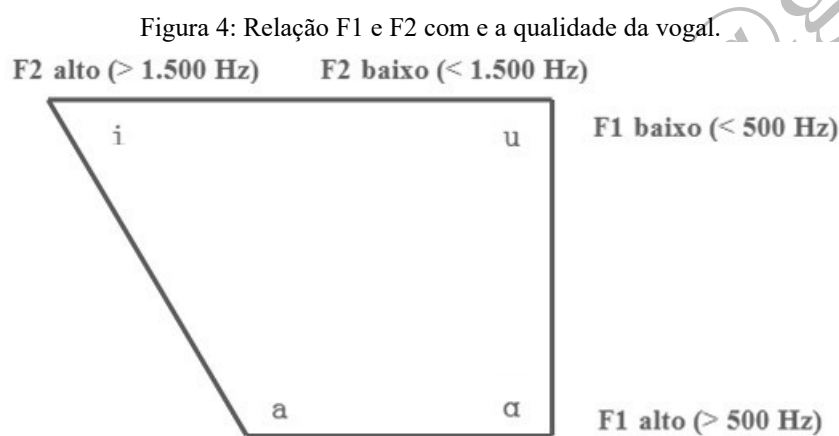


Figura 3: Vogais cardeais primárias (MARCHAL; REIS, 2012).

⁷ “The frequency of the formants, particularly the first two formants F1 and F2, are dependent on the shape of the airway between the glottis and the lips, and this shape in turn is determined by the position of the tongue body and the lips” (STEVENS, 1997, p. 472)

A posição das vogais no trapézio indica a posição da língua durante pronúncia da vogal e, acusticamente, as constrictões definem as frequências dos formantes (medidas em Hz). Conforme já mencionado, o primeiro formante está relacionado com o levantamento e abaixamento da língua, concomitante à abertura e fechamento de mandíbula. A elevação do corpo da língua diminui o valor de F1, enquanto que o abaixamento da língua aumenta o valor de F1. O segundo formante, por sua vez, tem relação com a anterioridade ou posterioridade da língua. Se o corpo da língua estiver anteriorizado, o valor de F2 será alto; se a língua estiver recuada, o valor será baixo. A Figura 4 ilustra essa relação.



Fonte: as autoras.

Os valores dos formantes dependem da geometria do trato vocal e o cálculo se baseia em modelos de tubo de ressonância (BRAID, 2003). O comprimento do tubo determina a frequência da ressonância, ou seja, dos formantes. Quanto mais longo o tubo de ressonância, menor os valores de frequência de formantes. Isso significa que os valores das frequências de formantes variam de acordo com as características dos falantes, e os dois fatores principais são sexo e idade (KENT; READ, 2002). Como exemplo, podemos citar valores médios encontrados por Escudero *et al* (2008) para a vogal /ε/ do português brasileiro: nas vozes femininas, F1: 646 Hz, F2: 2.271 Hz, F3: 2.897 Hz; nas vozes masculina, F1: 518 Hz, F2: 1.831 Hz, F3: 2.572 Hz. Vale mencionar, ainda, que existe um número infinito de formantes. Duckworth *et al* (2011) reportam sobre diversos autores que demonstraram que os formantes mais altos são mais relacionados a características dos falantes, o que em tese seria muito interessante para análise em contexto forense. No entanto, o sinal analisado em situações forenses é comumente originado de interceptações telefônicas e a largura da banda desse canal deve ser considerada.

Embora os formantes mais altos sejam ricos em informações que levem a particularização, essas frequências não são transmitidas pelo canal telefônico. Kunzel (2001) cita que a faixa a ser considerada encontra-se entre 350-3400 Hz. Assim, serão considerados, nesta pesquisa, apenas o primeiro e o segundo formantes.

2.4 Disfarce de voz

O uso de disfarce de voz em fonética forense tem chamado a atenção de alguns pesquisadores (MASTHOFF, 1996; KÜNZEL, 2000; GILLIER, 2001) embora, segundo Eriksson (2010), não seja muito comum sua utilização na prática de delitos. De acordo com esse autor, um disfarce pode causar sérios problemas para a identificação de locutor, especialmente se for realizado com o uso de equipamentos eletrônicos.

É importante considerar que o estudo do disfarce não é exaurido nele próprio: pode trazer informações importantes sobre o comportamento dos parâmetros acústicos quando submetidos a condições fisiologicamente distintas. Por exemplo, se um sujeito realiza como disfarce a voz soprosa, a comparação de sua voz normal com sua voz disfarçada exibirá quais medidas são alteradas e de que forma, possibilitando correções em casos específicos. Essa observação carece, ainda, de maiores estudos.

Para reconhecer/detectar um disfarce, é necessário ter conhecimento da voz natural do falante, ou não será possível afirmar se uma característica especial do dado de fala é disfarçada ou natural (KÜNZEL, 2000). Esse autor analisou a preferência de homens e mulheres no uso de disfarce. Os resultados de sua pesquisa demonstram que pessoas, ao disfarçar a voz, variam bastante a frequência fundamental. Os sujeitos com F0 mais alta que a média tendem a elevá-la ainda mais. O contrário também é verdadeiro, ou seja, pessoas com F0 mais baixa que a média, no disfarce, são percebidas como apresentando *pitch* mais baixo. Sua pesquisa demonstra também que as mulheres são mais relutantes que os homens em alterar drasticamente a F0. A preferência de tipos de disfarce já tinha sido abordada na pesquisa realizada por Masthoff (1996), a qual revelou a preferência por alterações na fonação. Revelou também que as alterações realizadas no disfarce afetavam no máximo dois parâmetros fonéticos, deixando muitos outros aspectos do comportamento vocal disponíveis para análise forense.

Hollien (2002) analisou a efetividade do disfarce no contexto de identificação de locutor. Segundo o autor, por exemplo, o sussurro pode eliminar ou reduzir informações sobre

a frequência fundamental e, conseqüentemente, a percepção do *pitch*, dificultando a identificação.

Gillier (2011) também pesquisou sobre disfarce de voz em fonética forense, analisando não somente a F0, mas também as frequências de formantes. Alguns dos pontos importantes no relato da autora são: os dois parâmetros são eficientes para discriminação de indivíduos; nem todos os disfarces são eficazes, pois não alteram as frequências dos parâmetros; o efeito do disfarce não é homogêneo entre as várias vogais, sendo possível recuperar marcas específicas de cada falante através dos triângulos vocálicos⁸.

Esses trabalhos, em suma, revelam que o estudo de disfarce é relevante em pesquisa de comparação de locutor. No entanto, o objetivo maior da decisão do nosso grupo de pesquisa no uso de disfarce foi obter parâmetros de comparação de voz para treinamento em análise perceptiva e acústica.

3. Metodologia

Conforme já exposto, o objetivo inicial deste projeto de pesquisa foi a realização de um trabalho multidisciplinar para inserir os alunos de um curso de Letras em uma pesquisa em fonética forense. Os objetivos específicos deste recorte do trabalho foram (i) analisar a qualidade vocal e a estratégia de disfarce de acordo com o protocolo VPAS (LAVÉ, 1980) – VPAS PB (CAMARGO; MADUREIRA, 2008); (ii) examinar quatro correlatos acústicos das sete vogais do português: duração; frequência fundamental; primeiro formante – F1, segundo formante – F2 (ESCUDEIRO *et al*, 2009) em voz normal e disfarçada; (iii) comparar vozes femininas com vozes masculinas; (iv) relacionar os espaços vocálicos verificados nos trapézios com as estratégias de disfarce percebidas pelos juízes.

3.1 Informantes, texto e gravação

Para o projeto de pesquisa como um todo, os informantes foram reunidos em cinco grupos de 10 pessoas cada, de acordo com gênero e faixa etária, com as seguintes características: mulheres pesquisadoras com idade entre 25 e 55 anos (média de 37), chamado Grupo Controle (GC); homens de 30 a 55 anos (média de 44,4), chamado Grupo de Homens (GH); mulheres de 30 a 55 anos (média de 42,3), chamado Grupo de Mulheres (GM); rapazes

⁸ Usamos aqui o termo “triângulo” porque assim foi utilizado por Gillier (2011). Ao longo do texto, porém, utilizamos o termo “trapézio vocálico”.

de 19 a 25 anos (média de 22,4), chamado Grupo de Rapazes (GR); garotas de 19 a 25 anos (média de 22,2), chamado Grupo de Garotas (GG). Neste recorte, estamos comparando as vozes femininas com as vozes masculinas, sem considerar o grupo controle como tal, nem as diferenças de idades. Todos os informantes assinaram termo de consentimento para uso dos dados exclusivamente para pesquisa.

A coleta de dados foi realizada em laboratório com tratamento acústico com o seguinte ferramental: Computador Pentium Dual Core 5.300 2.60 GHz, 1.99 Gb RAM Processador XP 2002 Service Pack 3, placa de som externa M-Audio Fast Track Pro 4x4, Microfone AKG C 3000 B. Para as gravações o programa utilizado foi o Audacity⁹. Os locutores forneceram suas vozes no interior de uma cabine acústica, na qual encontrava-se o microfone, recebendo a orientação de não variar sua posição corpórea, evitando tanto o distanciamento do microfone, como também uma inclinação cervical que pudesse modificar o movimento laríngeo.

O texto elaborado simulava um pedido de resgate durante um telefonema e era lido pelo informante por duas vezes em sua voz normal e duas vezes disfarçando a voz. A instrução era para que, no disfarce, buscasse não ser reconhecido. O texto continha para análise 14 palavras paroxítonas, contemplando as sete vogais orais do português /i, ε, a, o, u/, presentes em suas sílabas tônicas e localizadas entre consoantes plosivas não vozeadas /p, t, k/. As palavras inseridas no texto para análise eram:

/a/	/e/	/ε/	/i/	/o/	/o/	/u/
/'patu/	/ka'petε/	/'tεkε/	/'pike/	/'toku/	/'køkε/	/'putu/
/'kækε/	/ka'tetu/	/es'petu/	/'pipe/	/'koku/	/'tøkε/	/ku'tuku/

Em cada coleta, o registro de áudio foi capturado com as seguintes configurações: formato *wave* monocal, frequência de amostragem de 44,1 KHz e 16 bits. Os arquivos foram nominados de acordo com o grupo e o número sequencial do informante, para facilitar os procedimentos seguintes.

⁹Software livre, versão 2.0.3, disponível em www.audacity.sourceforge.net/.

3.2 A análise perceptiva das vozes disfarçadas

A análise dos disfarces foi feita por um grupo de seis pessoas¹⁰ (juízes), com orientação prévia sobre a aplicação do BP-VPAS (ZULEICA; MADUREIRA, 2008). Faz-se necessário reforçar que o protocolo não foi aplicado em toda extensão, nem foram utilizados os níveis de graduação. O objetivo era apoiar na detecção de qual teria sido o ajuste escolhido pelo participante para disfarçar sua voz, para posterior comparação com alterações de F1 e F2, ou seja, analisar a relação entre a percepção dos juízes e a variação no trato vocal visualizada por meio da análise acústica. Os juízes ouviam as vozes quantas vezes considerassem suficiente para preencher de seus formulários individualmente. Depois, as respostas eram comparadas e, voltando a ouvir novamente quantas vezes fossem necessárias, uma resposta final era negociada para que houvesse um consenso de grupo. Também é importante relatar que a diferença que aqui se faz entre vozes normais e vozes disfarçadas se refere à voz natural do informante em relação à sua voz com a estratégia de disfarce. Não foram considerados os possíveis ajustes que o locutor pudesse realizar naturalmente em sua emissão.

3.2.1 Resultados das análises

Como se observa na Tabela 1, a escolha dos disfarces se deu tanto em ajustes supralaríngeos, tal como a protrusão labial e a mandíbula aberta, como em ajustes fonatórios, incluindo o falsete e a voz crepitante. Nos dados de variação de *pitch*, o segundo valor refere-se à composição desse ajuste com outro, isto é, o informante protruiu os lábios e baixou o *pitch*. Algumas das estratégias foram usadas em combinação, por exemplo, mandíbula aberta + voz nasal. A voz nasal, na verdade, esteve na maioria das vezes combinada com outro ajuste.

Tabela 1: Estratégias de disfarce utilizadas pelos informantes.

Estratégia de disfarce	Vozes femininas (30)	Vozes masculinas (20)
Mandíbula aberta	2	7
Lábios protruídos	6	1
Língua avançada	2	1
Língua recuada	2	0
Ajustes supralaríngeos		
Língua abaixada	1	0
Língua elevada	0	1
Voz nasal	4	5
Voz denasal	1	0

¹⁰ Cinco alunos e a professora de fonética de um curso de Letras.

Ajustes fonatórios	Hiperfunção laríngea	4	1
	Falsete	4	0
	Voz crepitante	1	3
	Voz áspera	2	4
	Pitch elevado	5 + 2	1 + 7
	Pitch abaixado	2 + 6	1

Verifica-se, pelos resultados, que a protrusão labial e a mandíbula aberta foram os ajustes mais frequentes no trato vocal, enquanto que a voz áspera foi um ajuste fonatório bastante utilizado. Diferente de resultados anteriores (KÜNZEL, 2000), muitas mulheres baixaram o *pitch*, enquanto vários homens elevaram o *pitch*. Foi interessante notar que houve uma relação dessa variação de *pitch* com duas estratégias específicas de disfarce, a protrusão labial por informantes do sexo feminino e a voz áspera por informantes do sexo masculino. As seis informantes que usaram a estratégia de protrusão labial foram também percebidas com o *pitch* abaixado. Segundo Laver (1980), a protrusão labial aumenta o eixo longitudinal do trato vocal e, conforme o modelo do tubo de ressonância, o comprimento maior do trato diminui as frequências (KENT; READ, 2002). Levantamos a hipótese de que, na tentativa de parecerem mais agressivas, pelo contexto de um pedido de resgate em sequestro, essas informantes escolheram baixar o *pitch* da voz, conseqüentemente, baixando os valores da frequência fundamental.

Também, como será apresentado mais adiante, as frequências de formantes tiveram valores mais baixos quando no ajuste de protrusão labial, fato também exposto por Laver (1980). Alguns informantes masculinos que elevaram o *pitch* combinaram essa estratégia com o ajuste de voz áspera (quatro informantes). As duas informantes do sexo feminino que usaram a voz áspera como estratégia de disfarce, por outro lado, baixaram o *pitch*. Aqui também foi levantada a hipótese de que a escolha da voz áspera teve relação com o contexto e deve ter sido utilizada para causar a impressão de agressividade. Laver (1980) afirma que a voz áspera é usada no inglês como sinal paralinguístico de raiva, e Moisk (2012) constata que a emoção exagerada na voz áspera pode gerar atributos agressivos e transgressivos.

Quatro dos ajustes usados pelos informantes foram escolhidos para uma relação com a análise acústica, dois ajustes de fonação - a voz áspera e o falsete, e dois do trato vocal - a protrusão labial e a abertura de mandíbula. Essa análise será apresentada mais adiante.

3.3 Os procedimentos de medição para análise acústica

A partir dos arquivos de áudio, e com o recurso de etiquetagem (*textgrids*) do PRAAT¹¹, realizou-se a segmentação manual de cada vogal alvo da análise, bem como de sua respectiva palavra, conforme pode ser observado na Figura 5. A delimitação dos trechos das vogais foi realizada conforme Escudero *et al* (2009), sendo os pontos inicial e final definidos pelo primeiro e último períodos que continham amplitude considerável e com formatos compatíveis com os dos períodos mais centrais.

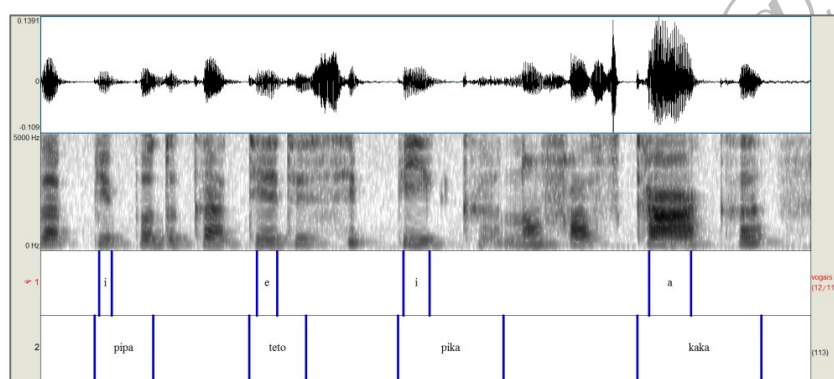


Figura 5: Tela do software do PRAAT, com visualização da etiquetagem do trecho da amostra GG9: “*pipa do Cateto e se pica, não faz caca*”. Na primeira camada está a forma da onda (oscilograma), na segunda o espectrograma, na terceira etiquetagem dos trechos correspondentes às vogais alvo etiquetadas e, na quarta e última, as palavras etiquetadas.

As transcrições das vogais (‘a’, ‘e’, ‘\ef’, ‘i’, ‘o’, ‘\ct’ ‘u’), serviram de referência para extração das medidas de duração, F0, F1 e F2 das vogais. A extração dos valores foi realizada com aplicação de script, desenvolvido pela terceira autora com os recursos disponibilizados no próprio Praat, baseando-se em tutoriais para aplicações semelhantes^{12,13}. A seguir são detalhadas as rotinas utilizadas para as medições, sendo importante ressaltar que após a extração os valores foram confirmados, sendo as medidas realizadas manualmente quando necessário.

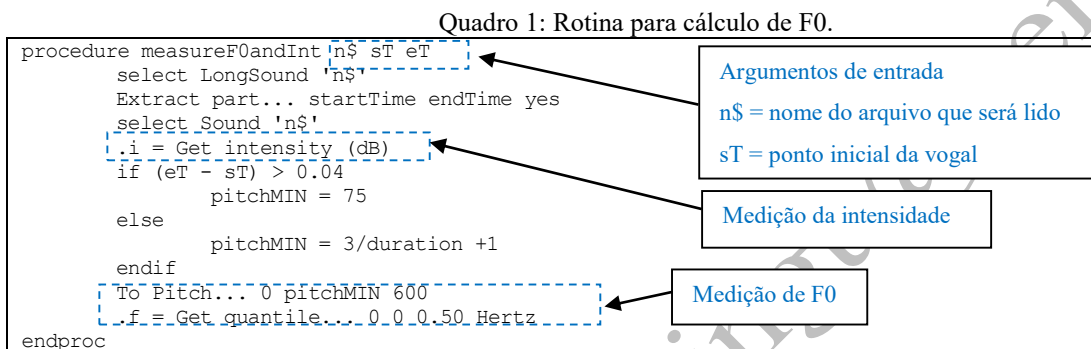
- a) Duração: determinada através de cálculo simples, pela subtração do ponto final e inicial de cada vogal.

¹¹ Software livre criado por Boersma e Weenick, 2008, disponível em <http://www.fon.hum.uva.nl/praat/>.

¹² Tutorial para Análise automática de formantes e plotagem de vogais no Praat, disponibilizado por Andréia Rauber em http://www.nupffale.ufsc.br/rauber/Praat_tutorial.pdf.

¹³ Scripts para extração de valores de F0 e de formantes, disponibilizados por Mietta Lennes (2003) em <http://www.helsinki.fi/~lennes/praat-scripts/>.

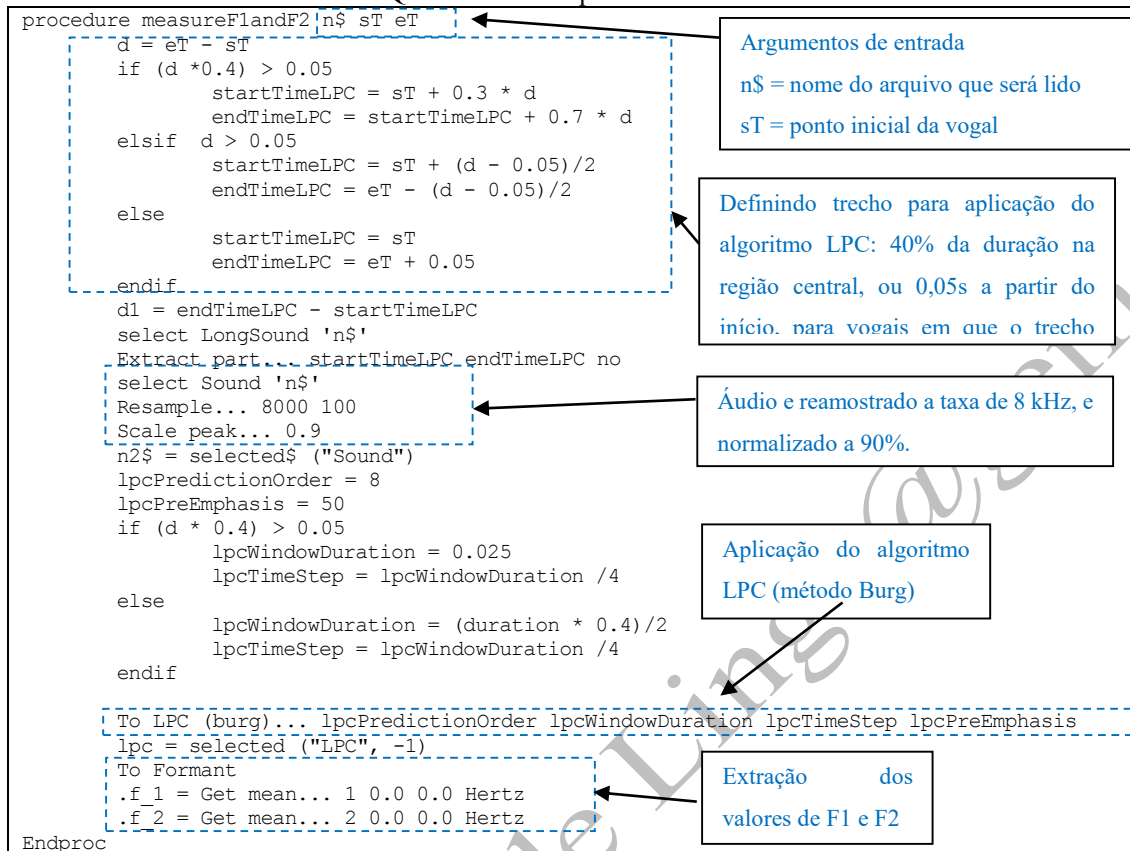
- b) F0: procedeu-se à extração do *pitch*¹⁴ do intervalo em análise (duração da vogal), sendo consideradas duas casas decimais no resultado. Foram mantidos os valores *default* do Praat para *ceiling* (valor máximo) de 600 Hz e o *floor* (valor mínimo) de 75 Hz. Para durações menores que 40 ms o algoritmo utilizado exige que o valor de *floor* seja recalculado, para atender a condição de resultar em no mínimo $(3/dur + 1)$. A seguir rotina utilizada para efetuar tal medição:



- c) F1 e F2: selecionou-se a porção central da vogal, correspondente a 40% da duração, com reamostragem do sinal para 8 kHz, precisão de 100 amostras e escala ajustada para 90% do valor de pico (normalização). Utilizou-se o algoritmo LPC, método *Burg*, sendo estabelecida ordem de predição igual a 8 (número de coeficientes utilizados no algoritmo, que deve corresponder a pelo menos o dobro do número de formantes a serem detectados), frequência pré-ênfase de 50 Hz, largura de janela de 25 ms para trechos maiores que 50 ms (em caso contrário, foi considerada a metade da duração do trecho) e um *timestep* de 25% da duração da janela. Daí, então, mediu-se a média do primeiro e do segundo formante. A seguir rotina utilizada para tal medição:

¹⁴ Vale lembrar que o *pitch* é o correlato perceptivo da frequência fundamental.

Quadro 2: Rotina para cálculo de F1 e F2.



4. Hipóteses e Resultados

4.1 Duração

Levando em conta os estudos de Escudero *et al.* (2009) sobre as vogais do português, levantaram-se as seguintes hipóteses em relação à duração: a) as mulheres produzem vogais mais longas que os homens; b) as vogais baixas são mais longas que as vogais altas (duração intrínseca); c) vogais posteriores tendem a ser mais longas que suas correspondentes anteriores. Também foi levantada a hipótese de que d) as vogais seriam mais longas no disfarce do que na voz normal, considerando que em situação de instrução a fala pode ser mais articulada.

A partir de testes não paramétricos Wilcoxon, foram encontrados os seguintes resultados:

- a) Hipótese não confirmada. Embora as mulheres tenham produzido sistematicamente vogais mais longas que os homens, tanto na voz normal como disfarçada, os resultados não foram estatisticamente significativos;

- b) Hipótese parcialmente confirmada. As diferenças foram significativas para as seguintes vogais: [e] > [i], [a] > [ɛ] e [ɔ] > [o], com $p < 0,05$ tanto na voz normal, como na voz disfarçada. Embora [ɛ] tenha sido mais longa que [e], as diferenças não foram significativas, com $p > 0,05$, nas duas modalidades. A vogal [o] foi mais longa que [u] na voz normal e mais curta no disfarce, mas também sem diferença significativa.
- c) Hipótese confirmada, salvo pela diferença entre [e] e [o] no disfarce, que teve a vogal anterior mais longa que a posterior, mas com diferença não significativa.
- d) Hipótese parcialmente confirmada. Os resultados apontam para uma tendência real de maior duração das vogais no disfarce, mas as diferenças não são significativas em todas as vogais nos diferentes grupos. A diferença de duração entre as vogais na fala normal e disfarçada parece ser mais relevante nos grupos masculinos, que tiveram diferença significativa ($p < 0,05$) nas vogais [e], [ɛ] e [o] no GH, e nas vogais [i], [e], [ɛ], [ɔ] e [o] no GR. Nos grupos femininos, GC teve resultados significativos em [i] e [o], GG em [o] e GM, em [i] e [e]. [a] e [u] não tiveram diferenças significativas de duração entre voz normal e disfarçada em nenhum dos grupos.

Conforme afirmam Escudero *et al* (2009), as vogais do português não apresentam a duração como traço fonológico, mas isso não exclui a possibilidade de haver diferenças fonéticas na duração das vogais, entre dialetos e entre gêneros. A pesquisa de Escudero *et al* (2009) comparou português brasileiro e português europeu, e vozes femininas e masculinas. Embora, na pesquisa que aqui se descreve, nem todas as hipóteses tenham se confirmado, tanto na diferença entre gêneros como no efeito da altura, duração intrínseca, a tendência se confirma. Quanto à diferença de duração na relação anterioridade-posterioridade da vogal, na pesquisa de Escudero *et al* (2009) para as vogais do português brasileiro, apenas a diferença entre [ɛ] e [o] não foi significativa. Nesta pesquisa, o que não foi significativa foi a diferença entre [e] e [o].

4.2 Frequência fundamental

Para a frequência fundamental, reportamos aqui as seguintes hipóteses em relação à frequência fundamental: a) as mulheres falantes de português têm valores de F0 mais altos que

dos homens¹⁵. A mudança do *pitch* é considerada uma das estratégias preferidas em disfarce de voz, conforme relatado inicialmente nos estudos de Eriksson (2010). Mais uma hipótese que aqui se levanta, então, é que consequentemente: b) haverá diferenças significativas nos valores de F0 entre voz normal e disfarçada, nos diversos grupos.

Extraídas as medidas de F0, testes estatísticos não paramétricos foram aplicados e os resultados foram:

- a) Hipótese confirmada. O teste de Mann Whitney revela que para todas as vogais, tanto na voz normal quanto disfarçada, as mulheres tiveram F0 mais altas que os homens, $p < 0,05$.
- b) Os resultados apontam para significativa diferença entre F0 das sete vogais na voz normal e disfarçada, quando se comparam os cinco grupos, conforme se observa nos resultados do teste Kruskal Wallis apresentados na Tabela 2.

Tabela 2 - Teste Kruskal Wallis – Comparação de F0 entre os cinco grupos.

		/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
Voz Normal	Chi square	31,862	28,840	27,443	28,587	22,418	26,873	32,184
	Asymp Sig.	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Voz Disfarçada	Chi square	23,824	14,640	16,473	22,704	10,087	20,676	20,411
	Asymp Sig.	0,000	0,007	0,002	0,000	0,039	0,000	0,000

O efeito do gênero é realmente significativo no tocante à F0, já que todas as vogais tiveram valores mais altos nas vozes femininas do que nas vozes masculinas. Além disso também comprova-se, pelos resultados das diferenças entre a voz normal e disfarçada que, além das questões anatômicas que determinam os valores de F0, esse é um parâmetro acústico que pode ser muito bem controlado pelo falante, conforme afirmam Kreiman e Sidtis (2012). As autoras também afirmam que esse é um parâmetro saliente ao ouvinte, fato comprovado pelos resultados da análise perceptiva.

Outros trabalhos de análise de F0 foram realizados a partir da mesma análise perceptiva (KREMER; GOMES, 2014; KREMER, 2015), nos quais foram comparadas as vozes femininas

¹⁵ Embora essa diferença universal entre vozes de homens e mulheres seja um dado confirmado em várias pesquisas (vide ESCUDERO ET AL 2009), levantou-se a hipótese pelo caráter formativo da pesquisa.

e masculinas a partir da curva de F0 e da estratégia de disfarce. Resultados dessas pesquisas mostraram que a maioria dos participantes elevaram o *pitch*, uns mais significativamente que outros, sendo que a maior elevação de *pitch* aconteceu com um participante masculino e o maior abaixamento do *pitch* foi percebido em duas participantes do sexo feminino. A escolha de elevação do *pitch* foi mais frequente entre os homens, e o abaixamento mais frequente entre as mulheres. Esses resultados são consistentes com os de Masthoff (1996), mas contrários aos resultados de Künzel (2000).

4.3 F1 e F2

Para os valores de F1 e F2, a análise foi feita na diferença entre as vozes femininas e masculinas, e duas hipóteses foram levantadas: a) mulheres tendem a ter valores de F1 e F2 mais altos que de homens (ESCUDERO et al, 2009); b) mulheres tendem a ter espaço vocálico maior que dos homens (GILLIER, 2011).

Para os valores de F1, o teste não paramétrico de Mann Whitney revelou que as diferenças são significativas para todas as vogais na voz normal e para quase todas na voz disfarçada ($p < 0,05$), consistente com o trabalho de Escudero *et al* (2009), que afirmam que essa observação já foi reportada de forma abundante na literatura. Nesta pesquisa, as vogais altas [i] e [u] tiveram diferenças não significativas no disfarce ($p > 0,05$).

Os valores de F2, por outro lado, apresentaram diferenças significativas apenas nas vogais anteriores nas vozes normais, e na vogal [ɔ] nas vozes normais e disfarçadas. Diferente do trabalho de Escudero *et al* (2009), que afirmam que as mulheres têm valores de F2 mais altos que dos homens, neste trabalho apenas a vogal média baixa apresentou tal resultado. As demais vogais posteriores tiveram diferenças estatisticamente não significativas nas vozes normais. Nas vozes disfarçadas, apenas a vogal [ɔ] apresentou valores com diferenças significativas entre vozes femininas e masculinas. Isso demonstra que os informantes, ao escolherem ajustes para disfarçar sua voz, realizaram movimentação no trato vocal, especificamente no corpo da língua, que resultaram em aproximação dos valores de F2 nas duas modalidades de voz.

O uso de padrão formântico das vogais em contexto forense, como um parâmetro distintivo robusto, tem sido defendido por vários autores (por exemplo, NOLAN *et al*, 2009, 2011; NOLAN; GRIGORAS, 2005). Duckworth *et al* (2011) alertam, no entanto, para a

importância na consistência no processo de medição de formantes. Recomendam que se estabeleça um protocolo para a análise forense para facilitar a validação das evidências.

Quanto à hipótese de que as mulheres apresentam espaço vocálico mais amplo que os homens, utilizou-se um gráfico com a média dos valores de F1 e F2. Nesse caso, para equilíbrio dos valores, apenas valores de 40 informantes foram utilizados. Não foram considerados os valores do chamado grupo controle, com vozes femininas.

Gillier (2011) defende que o triângulo (ou trapézio) vocálico pode ser uma importante ferramenta para análise, uma vez que revelam tendências articulatórias e características anatômicas do trato vocal dos indivíduos. É interessante verificar algumas questões que se apresentam nos gráficos das Figura 6 e 7, que representam os dados de voz normal e disfarçada, respectivamente, dos quatro grupos (GM, GG, GH e GR).

No gráfico da Figura 6, pode-se compreender o que representam as diferenças significativas dos valores de F1 entre a produção das vozes femininas (GM e GG) com a das masculinas (GH e GR), sempre mais altos para as mulheres, trazendo para baixo as linhas em tons de rosa. Para F2, conforme reportado acima, as vogais anteriores e a vogal posterior média baixa tiveram diferenças significativas entre as vozes das mulheres e dos homens, mas as vogais posteriores média alta e alta não apresentaram diferenças significativa. Nota-se que as regiões de [u] e [o] estão muito próximas nos quatro grupos.

No gráfico da Figura 7, observa-se que os valores de F1 nas vozes disfarçadas apresentaram diferenças significativas para quase todas as vogais, salvo as altas. Isso se constata pela proximidade, no gráfico, dos quatro grupos nas regiões de [i] e [u]. Como F2 se refere à anterioridade-posterioridade e, no disfarce, as diferenças não foram significativas, as regiões de produção ficaram muito próximas em quase todas as vogais.

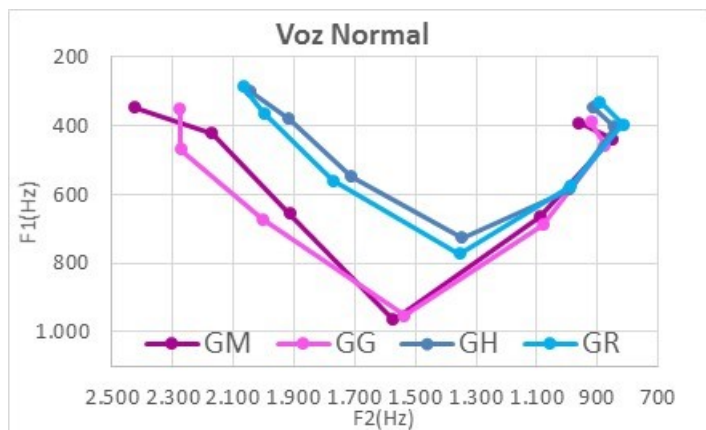


Figura 6: Espaço vocálico da média dos valores de F1 e F2 – voz normal.



Figura 7: Espaço vocálico da média dos valores de F1 e F2 – voz disfarçada.

5. Relação entre análise perceptiva e análise acústica

Para a melhor compreensão da relação entre a articulação empreendida pelo locutor, a percepção do ouvinte e o efeito acústico da voz, alguns gráficos de F1-F2 individuais agrupados por alguns ajustes na produção do disfarce serão analisados.

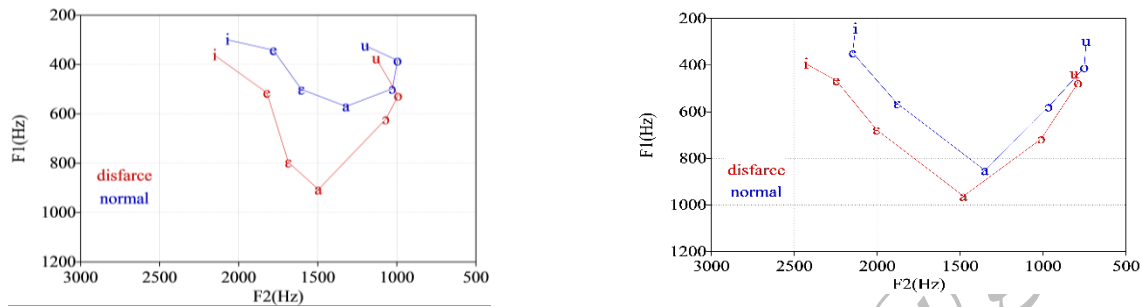
5.1 Estratégia de disfarce e efeitos acústicos

Após a análise auditiva dos disfarces, quatro ajustes foram selecionados para comparação com a análise acústica nas vozes normal e disfarçada. Por vezes, esses ajustes foram combinados com outros, por exemplo, o arredondamento dos lábios combinando com o abaixamento do *pitch*. Alguns dos ajustes mais comuns na análise dos 50 participantes, ao disfarçar a voz foram: a) mandíbula aberta – nove participantes (duas vozes femininas; sete vozes masculinas); b) protrusão labial – sete participantes (seis vozes femininas; uma voz masculina); c) voz áspera – seis participantes (duas vozes femininas; quatro vozes masculinas); d) falsete – quatro participantes (todas vozes femininas). Esses resultados reforçam a afirmação de Künzel (2000) de que há diferenças entre os gêneros na preferência por modos de disfarce, assim como de suas estratégias articulatórias para implementá-los. Nos participantes desta pesquisa, observa-se que a protrusão labial e o falsete são mais usados pelas participantes do sexo feminino, ao passo que a mandíbula aberta e a voz áspera são mais usadas por participantes do sexo masculino.

Na comparação entre os ajustes percebidos auditivamente e as análises dos gráficos das vogais formados pelos valores de F1 e F2, foi possível compreender a relação entre os movimentos realizados no trato vocal com a estratégia do disfarce utilizada. Os gráficos abaixo

mostram alguns exemplos. A cor azul indica as médias para as vogais produzidas na voz normal, e a cor vermelha, as médias das vogais produzidas na voz disfarçada.

5.1.1 Mandíbula aberta



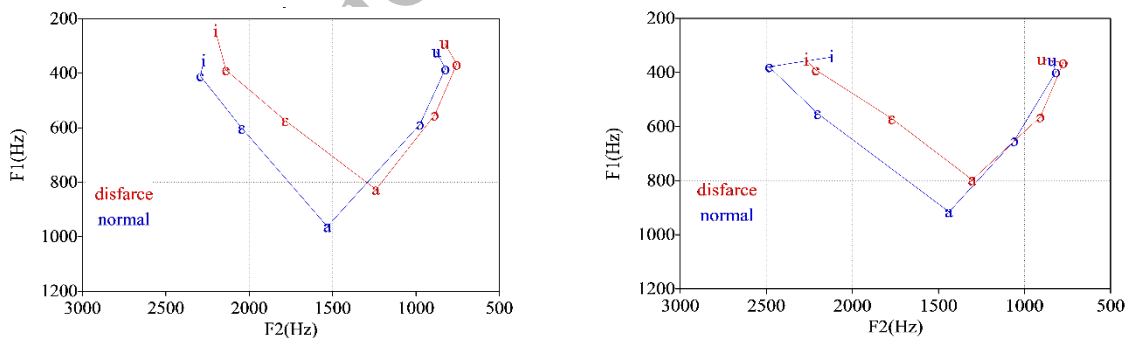
(a) Informante GH2.

(b) Informante GR9.

Figura 8: Gráficos de F1/F2 com as médias para voz normal e voz com disfarce de dois dos informantes que usaram o ajuste de mandíbula aberta.

Os gráficos demonstram como a percepção do ajuste de mandíbula aberta, escolhida pelos juízes, coincide com aumento do valor de F1, comprovando o abaixamento da língua pela abertura da mandíbula. Laver (1980) afirma que o efeito mais importante na posição da mandíbula é verificado nas mudanças de F1, com aumento de frequência na medida em que a mandíbula se torna mais aberta.

Protrusão labial



(a) Informante GM4.

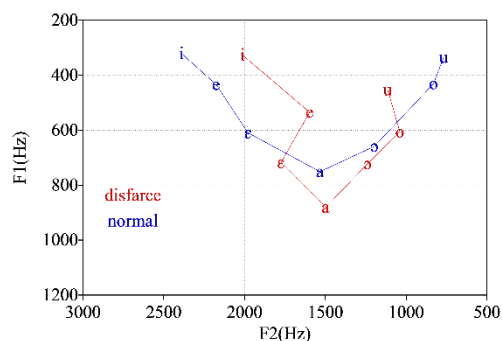
(b) Informante GG10.

Figura 9: Gráficos de F1/F2 com as médias para voz normal e voz com disfarce de dois dos informantes que usaram o ajuste de protrusão labial.

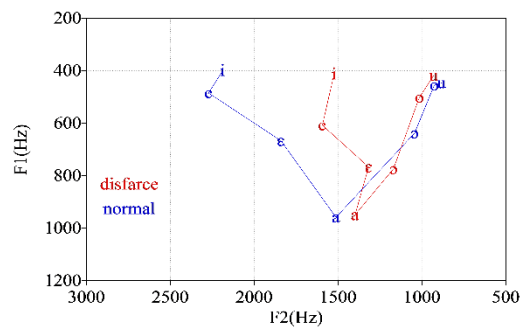
No ajuste de arredondamento dos lábios ou protrusão labial, percebe-se o recuo da língua pela diminuição de F2, principalmente nas vogais anteriores. A protrusão labial aumenta o eixo longitudinal do trato vocal, adicionando mais uma seção de comprimento variado ao

trato, o que resulta no efeito acústico de frequência mais baixa em todos os formantes (LAVÉ, 1980).

Falsete



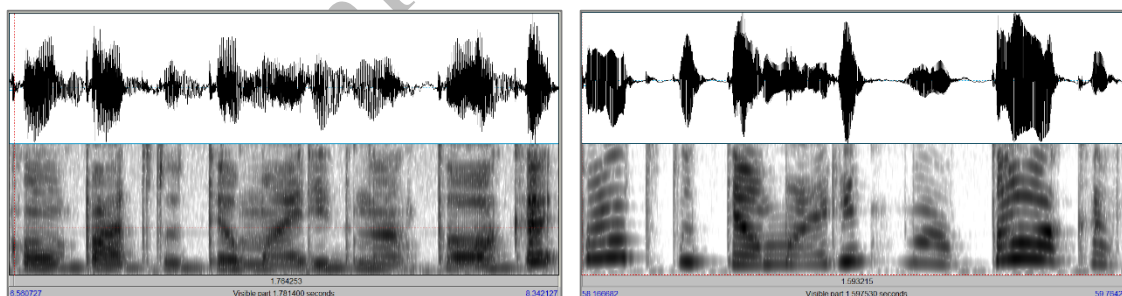
(a) Informante GC5



(b) Informante GG1.

Figura 10: Gráficos de F1/F2 com as médias para voz normal e voz com disfarce de duas das informantes que usaram o ajuste de falsete.

Pelos gráficos, observa-se que o efeito do falsete parece operar com um estreitamento do espaço vocálico. Segundo Laver (1980), o falsete é caracterizado acusticamente por vários fatores: F0 mais alta do que em voz modal; interação da F0 alta e o modo de vibração das pregas vocais; o espectro apresenta queda mais íngreme, caindo aproximadamente -20 dB por oitava. Enquanto a voz modal tem uma porção fechada da forma da onda laríngea como um componente abrupto, no falsete a porção aberta é mais inclinada.



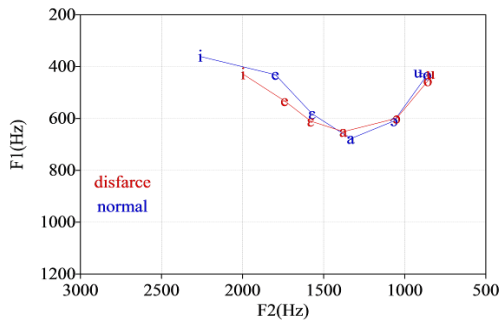
(a) Voz normal.

(b) Voz disfarçada com falsete.

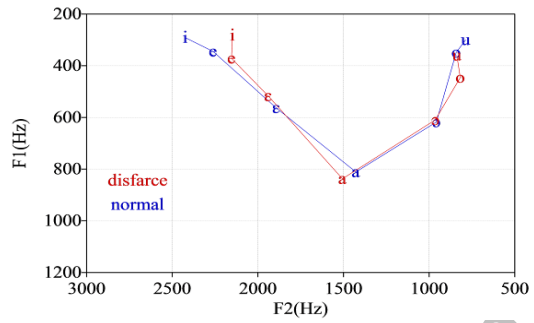
Figura 11: Oscilogramas e espectrogramas de trechos da informante GC3.

Verifica-se pelas formas da onda nas vozes normal e disfarçada da informante GC3, na Figura 11, as inclinações mais abruptas na forma da onda. O espectrograma, na Figura 11 (b), do falsete reflete a tensão das pregas vocais.

Voz áspera



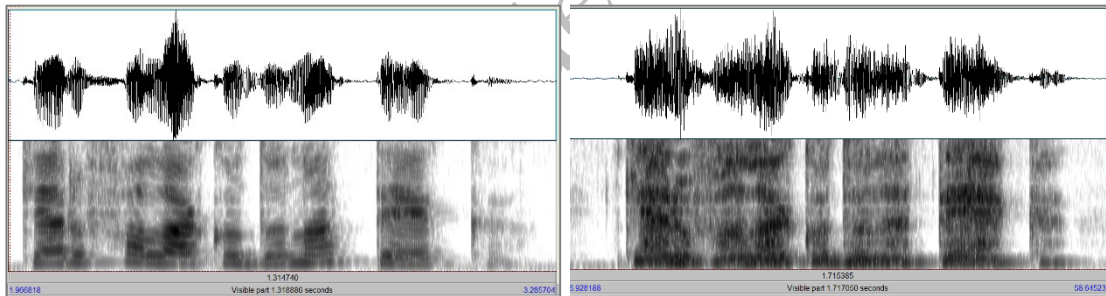
(a) Informante GH1.



(b) Informante GR6.

Figura 12: Gráficos de F1/F2 com as médias para voz normal e voz com disfarce de dois dos informantes que usaram o ajuste de voz áspera.

Na voz áspera praticamente não há alteração na configuração das vogais. Esse parece ser um ajuste que, realizado no nível laríngeo, não requer movimentos importantes na região supralaríngeo. As características acústicas da voz áspera estão relacionadas principalmente com a irregularidade da forma da onda glotal e ruído espectral (LAVER, 1980), o que pode ser observado na Figura 13.



(a) Voz normal.

(b) Disfarce com voz áspera.

Figura 13: Oscilogramas e espectrogramas de trechos da informante GC2.

O ruído característico da voz áspera pode ser observado tanto na forma da onda como no espectrograma da Figura 13(b). Segundo Moisk (2012) a correlação fisiológica da voz áspera é o aumento de tensão na região laríngeo e faríngeo, e atividade muscular intensa nas pregas vocais.

Podemos, assim, resumir as estratégias dos participantes na realização dos seus disfarces: a) a maioria buscou alterar o *pitch* na voz disfarçada, utilizando diferentes estratégias, ou de elementos do trato vocal ou de elementos fonatórios (ERIKSSON, 2010); b) elementos do trato vocal são mais efetivos para a variação de F1 e F2; c) dos elementos fonatórios, o

falsete foi o mais efetivo para a variação de F1 e F2, conforme pesquisa anterior (GILLIER, 2011); d) as vogais foram mais longas no disfarce, possivelmente por ser uma situação de instrução; e) a variação de F0 foi mais significativa nos grupos masculinos; f) a variação de F1 foi mais significativa nos grupos masculinos (vários usaram abertura de mandíbula), e de F2 nos grupos femininos (várias usaram o arredondamento dos lábios e avanço de língua).

6. Considerações finais

Este trabalho ressalta a importância da abordagem multidisciplinar da análise de fala em fonética forense. O cunho instrucional proposto foi levado a termo, tendo sido os membros do grupo instruídos sobre os procedimentos para realização de gravações e análises acústicas relacionadas à pesquisa. Foram capazes de atuar como juízes para a realização da análise perceptiva da qualidade vocal normal e disfarçada, identificando ajustes. Estiveram aptos também a analisar os parâmetros extraídos, estabelecendo relações entre a avaliação perceptiva e os achados acústicos, relacionando-os.

As hipóteses levantadas em relação à duração não foram em sua totalidade confirmadas, mas os resultados revelaram tendências que confirmam resultados de trabalhos anteriores, como o de Escudero *et al* (2009) para as vogais do português brasileiro, principalmente na diferença entre os gêneros e no efeito de altura.

Em relação à frequência fundamental, os resultados das diferenças entre a voz normal e disfarçada revelam a facilidade de controle desse parâmetro acústico pelo falante. As diferentes preferências por elevação ou abaixamento do *pitch* por informantes do sexo masculino e feminino também foram relevantes nesta pesquisa, pois foram condizentes com o contexto estabelecido para coleta de dados. O abaixamento do pitch, preferido por informantes do sexo feminino, pode ter relação com uma vontade de parecerem agressivas.

A análise do padrão formântico, que neste trabalho focou apenas o primeiro e segundo formantes, foi importante para que os participantes em instrução pudessem refletir sobre a relação entre a análise perceptiva que fizeram dos disfarces realizados e os movimentos articulatórios necessários para implementá-los. A partir da análise dos gráficos, esses participantes foram capazes de visualizar a abertura da mandíbula causando um aumento do valor de F1, o recuo da língua causando a diminuição do valor de F2, e o estreitamento do espaço vocálico causado pelo falsete.

Já defendemos em outros textos (GOMES *et al*, 2012; GOMES; CARNEIRO, 2014) a carência das pesquisas em fonética forense e a importância da abordagem interdisciplinar. Reforçamos aqui essa defesa, conclamando mais pesquisadores, profissionais e estudantes a realizar pesquisas para o desenvolvimento da área.

Referências

- BEHLAU, M. **Voz: O livro do especialista**. Rio de Janeiro: Revinter, 2004. Vol. 1.
- BELL, A. G. Vowel theories. In: **American Journal of Otology**, Vol. 1, p. 163-180, 1879.
- BRAID, A. C. M. **Fonética Forense**. Campinas: Millennium, 2003, 2. ed.
- CAMARA JR., J. M. **Estrutura da língua portuguesa**. Petrópolis: Editora Vozes, 2002. 35. ed.
- CAMARGO, Z.; MADUREIRA, S. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In: **Proceedings Of The Fourth Conference On Speech Prosody**. Campinas, 2008.
- DAJER, M. E. **Análise de Sinais de Voz por padrões visuais da dinâmica vocal**. Tese de Doutorado. Programa de Engenharia Elétrica da Escola de Engenharia da USP, 2010.
- DUCKWORTH, M.; MCDUGAL, K.; DE JONG, G.; SHOCKEY, L. Improving the consistency of formant measurement. In: **The International Journal of Speech, Language and the Law**, Vol. 18.1, 2011, p. 35-51. **crossref** <http://dx.doi.org/10.1558/ijsl.v18i1.35>
- ERIKSSON, A. The Disguised Voice: Imitating Accents or Speech Styles and Impersonating Individuals. In: **Language and Identities**, 8. Edinburg: Kapitell, 2010. p.86 – 96.
- ESCUADERO, P.; BOERSMA, P.; RAUBER, A. S.; BION, R. A. H. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. **Journal of Acoustic Society of America**, 126 (3), 2009, 1379–1393. **crossref** <http://dx.doi.org/10.1121/1.3180321>
- ESPINDULA, A. **Perícia Criminal e Cível: Uma visão geral para peritos e usuários da perícia**. Campinas: Millennium Editora Ltda, 2005, 2. ed.
- FANT, G. **The Acoustic Theory of Speech Production**. The Hague: Mouton, 1960.
- GILLIER, R. **O disfarce de voz em fonética forense**. Master Thesis, Departamento de Linguística Geral e Românica, Faculdade de Letras, Universidade de Lisboa, 2011.
- GOMES, M. L. C.; CARNEIRO, D. O. **A fonética forense no Brasil: cenários e atores**. *Language and Law/Linguagem e Direito*. Vol. 1(1), 2014. p. 751-754.

GOMES, M. L. C.; RICHERT, L. C.; MALAKOSKI, J. Identificação de locutor na área forense: a importância da pesquisa interdisciplinar. **Anais do X Encontro do CELSUL – Círculo de Estudos Linguísticos do Sul UNIOESTE - Universidade Estadual do Oeste do Paraná Cascavel-PR**, 2012.

GOLD, E.; FRENCH, P. An International Investigation of Forensic Speaker Comparison. In: **ICPhS XVII**. Hong Kong, 2011. p. 751-754.

HOLLIEN, H. **Forensic Voice Identification**. London, UK: Academic Press, 2002.

IAFPA. **Proceedings of IAFPA 2015**. 24th Annual Conference of the International Association for Forensic Phonetics and Acoustics. Leiden: Leiden University, 2015.

JONES, D. **An outline of English phonetics**. Cambridge: Heffer, 1962.

KEATING, P. Universal phonetics and the organization of grammars. In: FROMKIN, V. **Phonetic Linguistics: essays in honor of Peter Ladefoged**. New York: Academic Press, 1995. p. 115-132.

KENT, R. D.; READ, C. **Acoustic Analysis of Speech**. Albany: Singular, 2002. 2. ed.

KINOSHITA, Y., ISHIHARA, S., ROSE, P. **Exploring the discriminatory potential of F0 distribution parameters in traditional forensic speaker recognition**. Equinox Publishing Online, p. 92, 2009.

KREIMAN, J.; GERRATT, B. R. Perceptual Evaluation of Voice Quality: Review, Tutorial, and a Framework for Future Research. **Journal of Speech and Hearing Research**. Vol. 36, p. 21-40, Fev. 1993. **crossref** <http://dx.doi.org/10.1044/jshr.3601.21>

KREIMAN, J.; SIDTIS, D. **Foundations of Voice Studies**. Malden: Wiley-Blackwell, 2013.

KREMER, R. L.; GOMES, M. L. C. A eficiência do disfarce em vozes femininas: uma análise da frequência fundamental. **ReVEL**, vol. 12, n. 23, 2014.

KREMER, R. L. **Fundamental frequency: an analysis of age and gender**. Trabalho de Conclusão de Curso apresentado para conclusão de grau de licenciado no curso de Licenciatura em Letras Português/Inglês da UTFPR Curitiba, 2015.

KÜNZEL, H. J. Effects on voice disguise on speaking fundamental frequency. In: **Forensic Linguistics**, 7 (2). Birmingham: University of Birmingham Press, 2000. p. 149-179.

_____. Beware of the ‘telephone effect’: the influence of telephone transmission on the measurement of formant frequencies. **Forensic Linguistics** 8(1) 2001. p. 1350-1771.

LAVER, J. **The Phonetic Description of Voice Quality**. Cambridge: Cambridge University Press, 1980.

LINDBLOM, B. Vowel duration and a model of lip-mandible coordination. **Speech Transmission Laboratory – Quarterly Progress and Status Report**. Stockholm, 1, 29.

MADAZIO, G. **Diagrama de desvio fonatório na clínica vocal**. Tese apresentada à Universidade Federal de São Paulo - Escola Paulista de Medicina, para a obtenção do Título de Doutor em Ciências, 2009.

MARCHAL, A.; REIS, C. **Produção da fala**. Belo Horizonte: Editora UFMG, 2012.

MASTER, S. **Análise acústica e perceptivo-auditiva da voz de atores e não atores masculinos**: Long term average spectrum e o “formante do ator”. Tese de Doutorado no Programa de Pós-Graduação em Distúrbios da Comunicação Humana da Escola Paulista de Medicina da USP, 2005.

MASTHOFF, H. A report on voice disguise experiment. In: **Forensic Linguistics**, Vol.3 (1), Equinox Publishing Online, UK, 1996, p. 160-167.

MOISIK, S. R. Harsh Voice Quality and Its Association with Blackness in Popular American Media. In: **Phonetica**, 69, 2012. p. 193-215. **crossref** <http://dx.doi.org/10.1159/000351059>

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

NOLAN, F.; MCDUGALL, K.; DE JONG, G.; HUDSON, T. The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research. In: **International Journal of Speech Language and the Law**, 16(1), 2009. p. 31-57.

NOLAN, F.; MCDUGALL, K.; HUDSON, T. Some acoustic correlates of perceived (dis)similarity between same-accent voices. In LEE, W. S.; ZEE, E. **Proceedings of the 17th International Congress of Phonetic Sciences**, Hong Kong, 2011, p. 1506-1509.

NOOTEBOOM, S. The Prosody of Speech: Melody and Rhythm. In: HARDCLATLE, W. J.; LAVER, J. **The Handbook of Phonetic Sciences**. Oxford: Blackwell, 1997. p. 640-673.

PIETROFORTE, A. V. A língua como objeto da linguística. In: MUSSALIN, F.; BENTES, A. C. **Introdução à Linguística 3: fundamentos epistemológicos**. São Paulo: Cortez, 2005. p. 75-93.

ROSE, P. **Forensic Speaker Identification**. Canberra: Taylor & Francis, 1997.

STEVENS, K. N. Articulatory-Acoustic-Auditory Relationships. In: HARDCLATLE, W. J.; LAVER, J. **The Handbook of Phonetic Sciences**. Oxford: Blackwell, 1997. p. 462-506.

Artigo recebido em: 20.09.2015

Artigo aprovado em: 20.03.2016

Salvador, Vitória da Conquista e Teófilo Otoni: cidades e falares diferentes? Uma análise discriminante da F₀

Salvador, Vitória da Conquista e Teófilo Otoni: different cities and dialects? A discriminant analysis of the F₀

Vera Pacheco*

Marian Oliveira**

Tássia da Silva Coelho***

RESUMO: Neste artigo, propomo-nos a avaliar a frequência fundamental (F₀) das vogais /a/, /i/ e /u/ em sílabas tônicas e pretônicas em falares de Vitória da Conquista, BA, Salvador, BA e Teófilo Otoni, MG. A escolha dessas cidades se deve ao contato próximo dos sujeitos dessas três cidade em decorrência da estreita relação sócio-econômica entre elas, apesar de estarem geograficamente distantes. Nossa hipótese é de que o falar de Vitória da Conquista pode sofrer influência dos outros falares. Para a realização deste trabalho, procedemos a mensuração da F₀ tanto em sílaba tônica, quanto na sílaba pretônica, em três pontos da vogal: início, meio e fim. Em seguida, para a análise do agrupamento desses falares, realizamos a análise multivariada discriminante. Os resultados mostram que os falares de Vitória da Conquista Teófilo Otoni formam um grupo que se opõe ao falar de Salvador.

PALAVRAS-CHAVE: Salvador. Vitória da Conquista. Teófilo Otoni. Análise acústica. Frequência fundamental.

ABSTRACT: In this article, we propose to evaluate the fundamental frequency (F₀) of the vowels /a/, /i/ and /u/ of the speakers of Vitória da Conquista, Bahia, Salvador, BA and Teofilo Otoni, MG. We choose of these three cities because the contact between the natives due to narrow social and economic relationship between them, even though they are geographically distant. Our hypothesis is that Vitória Conquista dialect will suffer influence from the others dialects. For this work, we measure the vowel F₀ of the stressed syllable and pre stressed syllable, in three points: first, middle and end. To analysis of grouping the dialects, we conducted discriminant multivariate analysis. The results show that Vitória da Conquista and Teofilo Otoni orm a group opposed to Salvador.

KEYWORDS: Salvador. Vitória da Conquista. Teófilo Otoni. Acoustic Analysis. Fundamental Frequency.

* Doutora em Linguística. Docente do Departamento de Estudos Linguísticos e Literários da Universidade Estadual do Sudoeste da Bahia. Professora orientadora do Programa de Pós-Graduação em Linguística e do mestrado Profissional em Letras.

** Doutora em Linguística. Docente do Departamento de Estudos Linguísticos e Literários da Universidade Estadual do Sudoeste da Bahia. Professora orientadora do Programa de Pós-Graduação em Linguística e do mestrado Profissional em Letras.

*** Mestre em Linguística.

1. Introdução

Sabemos que quando se fala em diferenças dialetais, as diferentes realizações das vogais nas sílabas pretônicas é a que mais chama atenção. A realização das vogais médias abertas em sílaba pretônica é atribuída com frequência ao falar do Norte e do Nordeste. Essa divisão não é tão simples assim. Como lembram Lee e Oliveira (2003), podemos ter diferenças intra e interdialetoal, ou seja, diferentes realizações vocálicas, em contextos específicos, dentro de um mesmo dialeto, e diferentes realizações entre dialetos diferentes.

Não podemos perder de vista, contudo, que a realização das vogais médias abertas em posição pretônica não é a única marca dialetal levantada pelos linguistas. O alteamento ou não dessas vogais nessa posição constitui uma forte marca dialetal como evidencia o trabalho de Bisol (1980).

Todavia, outras evidências fonética, além das vogais médias em sílaba pretônica, podem atuar na demarcação dialetal como as variações melódicas, como mostram vários trabalhos na literatura (CUNHA, 2000; COLOMARCO, 2005; LIRA, 2009; ZENDRON DA CUNHA, 2012; OLIVEIRA, 2014, entre outros). Os contornos melódicos dos diferentes tipos de orações do Português do Brasil (PB), declarativa, interrogativa, assertiva podem ser pistas importantes nas caracterizações dialetais. Aspectos prosódicos podem, portanto, variar de região para região.

Nos estudos prosódicos sobre variações dialetais, dentre os parâmetros acústicos avaliados, a frequência fundamental (F_0) tem-se destacado como um parâmetro importante (CUNHA, 2000; JACEWICZ; FOX, 2015).

É nesse contexto que nos propomos a avaliar o papel da F_0 na delimitação do falar de três cidades importantes para os estados de Minas Gerais e Bahia, a saber: – Teófilo Otoni, localizada a nordeste de Minas Gerais –; Vitória da Conquista, no sudoeste da Bahia —; e Salvador, capital baiana, localizada mais ao norte do estado.

A escolha pela investigação desses três falares se deve ao contato próximo dos sujeitos dessas três cidade em decorrência da estreita relação sócio-econômica entre elas, apesar de essas cidades estarem geograficamente distantes. Vitória da Conquista, uma das cidades mais importantes do Sudoeste da Bahia, mantém traços estreitos tanto com Teófilo Otoni quanto com Salvador. Por ser considerada a “capital” do sudoeste baiano e possuir melhores condições de ensino, de assistência médica e maiores oportunidade de emprego, a cidade atrai cidadãos de outras cidades mineiras e baianas, principalmente de Teófilo Otoni, cujos moradores querem ir

para uma cidade grande atrás de melhores condições de vida; e Salvador, pois os soteropolitanos buscam sair da loucura de uma capital e ir para uma cidade do interior mais tranquila, mas sem perder de vista uma infraestrutura satisfatória. Além disso, muitos conquistenses passam parte de sua vida em Salvador e voltam para a cidade depois de um certo tempo.

É nesse cenário de um contato direto e muito próximo entre os falantes dessas três cidades que nos propomos a avaliar a F_0 , parâmetro acústico importante na caracterização dialetal, em sílabas tônica e pretônica nucleadas por /a/, /i/ e /u/.

Analisaremos a F_0 tanto em sílaba tônica quanto pretônica porque padrões prosódicos dessas sílabas podem diferir entre os dialetos (CUNHA, 2000), bem como analisaremos essas sílabas nucleadas pelas vogais /a/, /i/ e /u/, pelo fato de essas três vogais, nos dialetos investigados, não estarem, a princípio, sujeitas a processos fonológicos que mudem a qualidade vocálica como se observa com as vogais médias que podem sofrer abaixamento, alteamento ou harmonia vocálica.

A pergunta que norteia esse trabalho é: o contato direto entre Teófilo Otoni e Vitória da Conquista e Vitória da Conquista e Salvador pode trazer consequências para o padrão de F_0 do falar dos sujeitos nativos dessas três regiões? Em outras palavras, a F_0 pode revelar, de alguma forma, a relação entre essas cidades?

Nosso objetivo é investigar a F_0 das vogais /a/, /i/ e /u/ em sílabas tônica e pretônica de falantes naturais de Teófilo Otoni, Vitória da Conquista e Salvador. Assim objetivamos avaliar em que medida esses falares estão correlacionados entre si no que se refere à F_0 .

Nossa hipótese é que o padrão de F_0 dos falantes de Vitória da Conquista seja intermediário entre o padrão de F_0 dos falantes de Teófilo Otoni e Salvador, pois Vitória da Conquista recebe cidadão dessas duas cidades.

Não há registro na literatura, até onde sabemos, de uma análise discriminante do padrão da F_0 , como a que se propõe aqui, na compreensão da configuração dialetal do Brasil. Os resultados aqui obtidos se somarão aos já encontrados na literatura e poderão contribuir na constituição do Atlas Prosódico de PB.

2. Para além das vogais médias abertas em sílabas pretônicas: variação entoacional, F_0 e caracterização dialetal a partir das vogais /a,i,u/

A realização das vogais médias abertas em sílaba pretônica é uma caracterização dos diferentes falares do PB há tempos já documentada na literatura. Nascentes, em 1953,

considerou a ocorrência ou ausência dessas vogais em posição pretônica em determinadas regiões do país como critério para a sua proposta de divisão de seis subfalares¹ em dois grandes grupos, o dos falares do Norte e o dos falares do Centro-Sul.

A divisão entre os falares, contudo, não é tão simples assim, pois, podemos ter variações das vogais pretônicas intra e inter dialetos (LEE; OLIVEIRA 2003). Além disso, nessa mesma posição silábica, processos de alteamento e harmonia vocálica podem ocorrer tanto nos falares da região norte (NINA, 1991; CRUZ ET AL. 2008, entre outros) e nordeste (MOTA, 1979; BORTONI, ET AL. 1991, VOGELY; HORA, 2013, entre outros) quanto nas regiões sul (BISOL, 1988; KAILER, 2012 entre outros) e sudeste (VIEGAS, 1987, CELIA, 2004, entre outros).

Apesar de a realização das vogais médias em sílaba tônica ser bastante variável entre os falares do PB, não podemos negar, todavia, que a ocorrência das vogais médias abertas na sílaba pretônica é um marcador dialetal utilizado com frequência pelos brasileiros no reconhecimento de falantes nordestinos, especialmente.

A presença ou não das vogais médias abertas em sílaba tônica não é o único delimitador dialetal. Variações nos padrões acústicos de sílabas pretônica e tônica, nos contornos entoacionais de assertivas e interrogativas, por exemplo, podem ocorrer nos diferentes dialetos do PB, como registram os trabalhos realizados no âmbito do projeto ALiB² e do projeto Amper³, além de outros estudos realizados fora do escopo desses dois projetos. Assim retomando as considerações de Cunha (2000, p. 185), “se os falares se distinguem prosodicamente e essa diferença é perceptível e descritível empiricamente por um leigo é porque, de fato, as diferenças existem e são passíveis de um estudo científico”; temos, portanto, variações dialetais marcadas prosodicamente.

Essas variações dialetais prosodicamente marcadas podem ser atestadas na avaliação da configuração dos padrões acústicos de F₀, intensidade e duração de sílabas tônicas e pretônicas

¹ Os termos “falares” e “dialetos” têm o mesmo significado: designam a realização variável de uma língua em função de fatores geográficos. Neste trabalho, optamos por usar os termos que as fontes consultadas utilizam.

² O Projeto ALiB (Atlas Linguístico do Brasil), coordenado pelas pesquisadoras Jacyra Mota e Suzana Cardoso, “tem como meta a realização de um atlas geral do Brasil no que diz respeito à língua portuguesa” (CARDOSO; MOTA, 2007, p. 146) e agrega pesquisadores de diversas regiões do Brasil.

³ O projeto Amper (Atlas Multimídia Prosódico do Espaço Românico), coordenado por Michel Contini e Jean-Pierré Lai, tem como principal objetivo o estudo da organização prosódica das variedades faladas no espaço dialectal românico. O projeto é sediado na Universidade de Grenoble Alpes, França, e conta com a participação de diversas universidades europeias e sul-americanas, inclusive universidades brasileiras (<http://pfonetica.web.ua.pt/AMPER-POR.htm#2>) <acesso: 10/05/2016>.

na caracterização dos falares de Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre. Assim temos, de um lado, as cidades nordestinas Recife e Salvador que apresentam falas com pouca diferença entre os parâmetros acústicos das sílabas pretônicas e tônicas; e Porto Alegre de outro, cujo falar de seus nativos apresenta maiores valores da F_0 , intensidade e duração na sílaba tônica. Entre esses dois falares podemos encontrar São Paulo e Rio de Janeiro, cidades da região sudeste, cujos falares se caracterizam pelas sílabas tônicas e pretônicas com padrões acústicos que ora se assemelham àqueles dos falares nordestinos; ora aos da cidade gaúcha (CUNHA, 2000).

Ao avaliar o padrão melódico das interrogativas total para o nordeste, Lira (2009) apresenta basicamente dois padrões entoacionais, com marcas diferenciadoras no final do enunciado sobre a tônica ou a pós-tônica: 1) tônica final alta, seguida de pós-tônica (s) baixa (s) predominante em João Pessoa e Fortaleza e; 2) tônica final baixa, seguida de pós-tônica alta, predominante em Recife, Salvador e São Luis.

O padrão encontrado em João Pessoa e Fortaleza o é também para os falantes da ilha de Mosqueiro, Pará, apesar da não proximidade entre essas regiões, como ressalta Guimarães (2013).

O contorno melódico das interrogativas totais pode também variar entre os falares de alguns estados das regiões nordeste, sudeste e sul. De acordo com Silva e Cunha (2011), Rio de Janeiro, cidade carioca, da região sudeste, e Florianópolis, da região sul, têm, nesse tipo de sentença, seus falares caracterizados pelo contorno circunflexo (ascendente-descendente) no final dos enunciados, ao passo que Recife, cidade da região nordeste, apresenta contorno ascendente.

A semelhança no padrão da interrogativa total de Florianópolis e Rio de Janeiro, conforme dados de Silva e Cunha (2011), não é encontrada ao se comparar o alinhamento de F_0 entre os falares de Florianópolis e Petrópolis, outra cidade do estado do Rio de Janeiro. De acordo com Seara e Rebollo-Couto (2011), na fala florianopolitana ocorre preferencialmente um alinhamento de F_0 centralizado ou tardio, ao passo que na fala petropolitana ocorre um alinhamento antecipado.

Diferenças dialetais na realização das declarativas têm-se mostrado menos frequentes do que as diferenças encontradas nas interrogativas. Diferença na realização das declarativas foi encontrada em Seara e Rebollo-Couto (2011) também na comparação entre o falar de Florianópolis e Petrópolis. As declarativas nesses dois falares diferem entre si basicamente pela

duração. De acordo com essas autoras, o dialeto fluminense se caracteriza por apresentar maior duração silábica se comparado ao falar dos catarinenses.

As pesquisas sobre caracterização e comparação de padrões prosódicos de diversas cidades de diferentes estados e regiões brasileiros evidenciam nitidamente que esse aspecto linguístico pode diferir dialeto. Mas esse mesmo aspecto linguístico pode diferenciar falares de cidades pertencentes a um mesmo estado e geograficamente próximas, como mostra o estudo de Sosa, Nunes e Seara (2013), ao investigar a fala de sujeitos naturais de três cidades catarinenses: Florianópolis, Blumenau e Lages. Os falantes de Lages apresentam tom ascendente-descendente na região nuclear e os falantes de Blumenau apresentam diferenças tonais na região pré-nuclear.

Os trabalhos que avaliam as diferenças nos padrões prosódicos entre os dialetos brasileiros deixam evidentes que dentre os parâmetros acústicos comumente envolvidos nas variações melódicas, F_0 , duração, intensidade, a F_0 é aquele que mais se destaca na discriminação dialetal.

A F_0 como um marcador dialetal parece não estar associada às variações melódicas, exclusivamente, mas igualmente associada à característica intrínseca das vogais, a F_0 intrínseca, como demonstram Jacewicz e Fox (2015).

Apesar do consenso entre os pesquisadores de que a F_0 intrínseca, ou *pitch* intrínseco, como chamam Ohala e Eukel (1987), das vogais ser uma consequência automática da articulação (CRANDAL, 1925, PETERSON; BARNEY, 1952; LEHISTE; PETERSON, 1961 apud OHALA; EUKEL, 1987), e, por isso, ser universal entre as línguas do mundo (WALEN; LEVITT, 1995)⁴, o trabalho experimental realizado por Jacewicz e Fox (2015) mostra que a F_0 intrínseca pode, sim, variar em contextos específicos entre os dialetos.

Jacewicz e Fox (2015) mostram experimentalmente que o valor da F_0 intrínseca das vogais /I, e, æ, aI / tende a altera-se em função da região de origem do falante. A pesquisa dos autores considerou três regiões diferentes dos Estados Unidos: Carolina do Norte; Ohio Central e Wisconsin, falantes do sul, da região central e do norte, respectivamente.

De acordo com os autores, a F_0 das vogais realizadas por falantes da Carolina do Norte segue um padrão típico da F_0 das vogais realizadas por falantes de línguas tonais. Segundo os autores, a alteração da F_0 intrínseca é observada em contextos prosódicos proeminentes, tais

⁴ Vogais altas tendem a ter o valor da F_0 intrínseca mais alta do que o das vogais baixas (OHALA; EUKEL, 1987).

como foco da sentença. Para eles esse resultado reflete uma utilização diferenciada de produção da F_0 pelas comunidades de fala. Os autores, contudo, não apresentam uma hipótese que possa explicar essa relação. Afirmam somente que a F_0 intrínseca pode ser controlada pelo falante, não sendo somente uma contingência articulatória. Na visão dos autores, a alteração deliberada da F_0 seria uma forma de aumentar a percepção da identidade social.

De fato o reconhecimento de uma língua ou de uma variedade linguística pode se dar somente a partir de informações prosódicas, especialmente pela variação de F_0 , segundo análise experimental realizada por Peres (2011), com base nos pressupostos de Atkinson (1968), Bonte (1975), Maidment (1976); Ohala e Gilbert (1978) e Bezooijen e Goss (1999)

Em condições experimentais de preservação prosódica dos estímulos, Peres (2011) encontrou aumento no reconhecimento de três variedades do Português do Brasil (Pelotas, (RS); São Paulo (bairro da Mooca) e Senador Pompeu (CE), por sujeitos naturais dessas regiões, em comparação com os estímulos que continham manipulações segmentais e de F_0 . Diante desse resultado, o autor afirma que a F_0 pode ser a responsável pelo reconhecimento, por parte dos sujeitos, de variedades linguísticas, isso porque a análise de produção feita por ele evidenciou pistas relevantes nas características desse parâmetro.

O trabalho de Peres (2011) evidencia que a percepção da fala não opera sem considerar questões de ordens dialetais. Assim, as variações dialetais devem ser contempladas nas investigações sobre a percepção da fala, bem como não pode ser negligenciada nas propostas de modelos para a percepção da fala (PISONI; REMEZ, 2005).

Diante das questões aqui apresentadas, podemos afirmar que a F_0 é um parâmetro acústico robusto na caracterização dialetal, em termos de produção, bem como o é no reconhecimento de um dialeto pelo ouvinte. Nesse cenário da eficácia da F_0 na demarcação dialetal e considerando a relação entre os sujeitos naturais de Vitória da Conquista, Teófilo Otoni e Salvador, buscamos investigar a relação entre esses três falares de três regiões importantes para os estados de Minas Gerais e Salvador.

3. Metodologia

Buscando responder à pergunta posta neste trabalho, a saber: a ligação dos falantes naturais de Vitória da Conquista, Teófilo Otoni e Salvador decorrente da relação sócio-econômica dessas cidades pode trazer consequências para o padrão de F_0 do falar dos sujeitos nativos dessas três regiões? e com vistas a alcançar os objetivos aqui propostos quais sejam: a)

investigar a F₀ das vogais /a/, /i/ e /u/ em sílabas tônica e pretônica dos falares de Teófilo Otoni, Vitória da Conquista e Salvador; e b) avaliar em que medida esses falares estão correlacionados entre si no que se refere à F₀, a metodologia utilizada nesta pesquisa foi como se segue abaixo.

3.1 As cidades

O foco de nossa pesquisa foram três cidades diferentes: Vitória da Conquista, Teófilo Otoni e Salvador.

Vitória da Conquista, terceira maior cidade do interior do estado da Bahia, localiza-se no planalto da Conquista, região Sudoeste do estado, da qual é considerada a capital. Conforme estimativa do Instituto de Geografia e Estatística (IBGE) de 2014, a cidade possui 340 mil habitantes e é a capital regional de uma área que abrange oitenta municípios da Bahia e dezesseis do norte de Minas Gerais.

Salvador é a capital do estado da Bahia, estado localizado na Região Nordeste do país. A população soteropolitana é de mais de 2,9 milhões de habitantes, segundo estimativa do IBGE de 2014, e é o município mais populoso do Nordeste e o terceiro do Brasil. Salvador dista 519 km de Vitória da Conquista.

Teófilo Otoni, município do interior do estado de Minas Gerais, Região Sudeste do país, dista 389 km de Vitória da Conquista. Pertence à Mesorregião do Vale do Mucuri e Microrregião de Teófilo Otoni e localiza-se a Nordeste da capital do estado, Belo Horizonte, distando desta cerca de 450 km. Sua população foi estimada pelo IBGE, em 2014, em 140 mil habitantes, sendo então o 18º mais populoso do estado.

A motivação para a escolha dessas cidades foi estreito contato entre os cidadãos dessas três cidades decorrente da relação sócio-econômica que elas possuem. Vitória da Conquista é ao mesmo tempo desvinculada sócio-economicamente da capital Salvador, mas mantém contato com a mesma, graças ao fluxo de conquistenses indo e voltando para a capital para resolver demandas políticas e burocráticas. Mantém igualmente estreito laço com Teófilo Otoni, apesar de ser uma cidade de outro estado, de outra região brasileira, sendo possível encontrar morando em Vitória da Conquista, muitos teófilo-otonenses, graças a importância sócio-econômica de Vitória da Conquista para as cidades do norte de Minas.

3.2 Sujeitos da pesquisa

Participaram dessas pesquisa seis sujeitos, um homem e uma mulher de cada uma das cidades, todos com idades entre 25 e 40 anos, naturais e residentes ao longo de toda suas vidas desses locais, com nível superior concluído ou em andamento.

3.3 Montagem do *corpus* da pesquisa

O *corpus* desta pesquisa foi montado com palavras reais e logatomas dissílabos que obedeceram à estrutura CV.‘CV, em que C é consoante e V vogal, nas quais a posição C foi ocupada por obstruintes sonoras e soantes, visto que as desvozeadas poderiam espriar o desvozeamento para as vogais e desvozeá-las, e V pelas vogais altas /i/ e /u/ e a vogal baixa /a/. Antes da vogal /i/, foram inseridas consoantes [+ anteriores], por estarem na mesma zona de articulação, com a finalidade de atenuar processos fonológicos decorrentes da coarticulação. Antes da vogal /u/, consoantes [-anteriores] pelo mesmo motivo, exceto com as palavras [‘lulu], [‘luta] e [lu‘ta]. Nós optamos por essas vogais por estarem em regiões quânticas e serem menos suscetíveis a sofrerem processos fonológicos. As palavras estão discriminadas nos quadros 1 e 2.

Quadro 1 - Palavras-alvo com consoantes e vogais iguais em sílabas pretônica e tônica.

C ₁ = C ₂ V ₁ = V ₂		
VOGAL /i/		
Oclusivas	Fricativas	Soantes
B	V	M
Bi‘bi	Vi‘vi	Mi‘mi
VOGAL /u/		
Oclusivas	Fricativas	Soantes
G	R	L
Gu‘gu	Ru‘rru	Lu‘lu
VOGAL /a/		
Oclusivas	Fricativas	Soantes
D	Z	N
Da‘da	Za‘za	Na‘na

Fonte: Elaboração própria.

Quadro 2 - Palavras-alvo com consoantes e vogais diferentes nas sílabas pretônica e tônica.

C ₁ ≠ C ₂ V ₁ ≠ C ₂		
VOGAL /i/		
Oclusivas	Fricativas	Soantes
B	V	M
'bica	'vila	'mico
Bi'ca	Vi'la	Mi'co
VOGAL /u/		
Oclusivas	Fricativas	Soantes
G	R	L
'Gula	'Ruga	'Luta
Gu'la	Ru'ga	Lu'ta
VOGAL /a/		
Oclusivas	Fricativas	Soantes
D	Z	N
Da'li	'Zaga	'Nada
Li'da	Za'ga	Na'da

Fonte: elaboração própria.

As palavras foram inseridas na frase veículo “Digo ___ baixinho”, com o objetivo de padronizar o contexto de produção.

3.4 Gravações

As gravações foram feitas por meio do programa *Audacity*, no Mac book pro, processador 2,4 GHz Intel Core i5, software OS X 10.9.4, com microfone interno com redução de ruído externo, a uma taxa de amostragem de 44.100 Hz, do fabricante Apple Inc, em ambientes silenciosos, a fim de evitar interferências do ambiente externo. As frases-veículo foram impressas e apresentadas aos informantes. Foram feitas cinco repetições de cada frase. As gravações foram realizadas nas cidades de origem dos sujeitos avaliados.

3.5 Mensurações

Foram mensurados, por meio do *Software* de análise acústica *Praat* (BOERSMAN; WEENICK, 2006), os valores de F₀ das vogais pretônicas e tônicas em três pontos obtidos ao longo da realização da vogal: no início no meio e no final da mesma. Consideramos como ponto inicial, o primeiro pico regular da vogal, imediatamente depois da consoante; o ponto medial, o meio da vogal, localizado automaticamente pelo Praat, através do comando “control 0”; e o ponto final, último pico regular imediatamente antes da consoante, quando era o caso. Dessa

forma, os três pontos - ponto inicial, ponto medial e ponto final - de obtenção dos valores da F_0 ao longo da vogal foram obtidos como demonstrado na figura 1, indicados, respectivamente, como P1, P2 e P3.

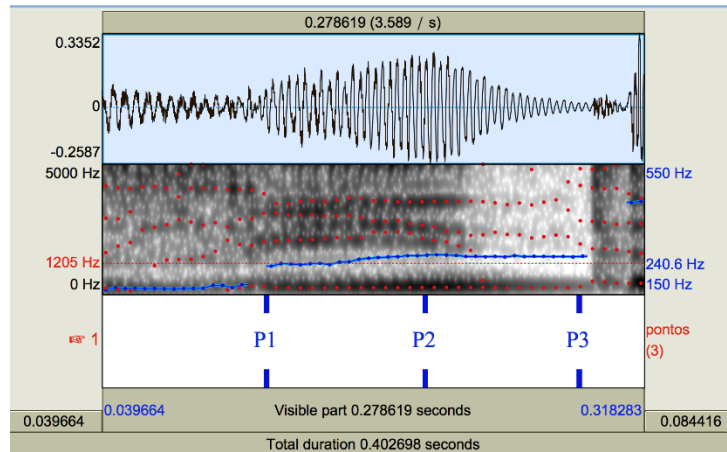


Figura 1- Forma de onda e espectrograma da vogal [i] com indicação dos três pontos de medida da F_0 (P1, início, P2, meio e P3, fim).

Fonte: elaboração própria.

Com esse procedimento metodológico fomos capazes de avaliar a F_0 ao longo da produção da vogal e não somente analisá-la em um ponto específico.

3.6 Análise estatística

Para avaliar a distância e a proximidade dos falares de Salvador, Teófilo Otoni e Vitória da Conquista, no que se refere à F_0 , procedemos a análise multivariada discriminante, que é uma ferramenta estatística que permite ao pesquisador separar e discriminar os grupos de indivíduos ou objetos, a partir da determinação das funções discriminantes de Fischer, Y_1 e Y_2 , que são duas variáveis adimensionais. Essas variáveis permitem melhor separação e discriminação entre os grupos que estão sendo analisados. A análise se dá a partir da geração de um diagrama que identifica os grupos e permite a visualização das separações e aproximações grupais (AYRES et al., 2007).

A análise discriminante permitiu comparar as medidas de F_0 obtidas a partir da extração de valores em três pontos de cada uma das vogais, /a/, /i/ e /u/ (início, meio e fim) na sílaba pretônica e na sílaba tônica de cada uma das cinco repetições. Esse procedimento metodológico atenuou possíveis diferenças nos valores de F_0 decorrentes de características fisiológicas dos sujeitos participantes dessa pesquisa. As medidas para cada uma dessas variáveis para as

produções de Salvador constituíram o **Grupo 1**; **Grupo 2**, as produções de Teófilo Otoni e **Grupo 3**, as produções de Vitória da Conquista.

Os valores das variáveis são distribuídos a partir do cálculo das discriminantes de Fischer em um diagrama constituído por quatro quadrantes, aqui identificados como: **Quadrante I**: superior esquerdo; **Quadrante II**: superior direito; **Quadrante III**: inferior direito e **Quadrante IV**: inferior esquerdo, como esquematizado na figura 2.

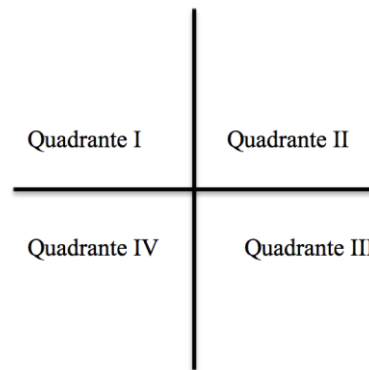


Figura 2 - representação esquemática da distribuição dos quatro quadrantes sobre os quais serão plotados os resultados da análise multivariada discriminante.

Fonte: elaboração própria.

Apresentada a metodologia por nós utilizada, passemos agora a examinar os resultados obtidos e as discussões deles provenientes.

4. Resultados e discussões

A frequência fundamental é a onda de menor valor de uma onda sonora harmônica e a variação de *pitch* é a indicação perceptual da variação contínua de valores altos e baixos da frequência fundamental (KENT; READ, 1992) e é um parâmetro robusto, como aqui discutido, na investigação sobre diferenças dialetais.

A F_0 exerce um papel importante na língua, como a distinção entre consoantes surdas e sonoras e, obviamente na fala, como indicação de sussurro, murmúrio, etc. Aqui buscamos investigar se pessoas de cidades diferentes podem apresentar um padrão de F_0 na sílaba que seja diferente entre cidades.

Nesta seção apresentamos os resultados obtidos na investigação sobre a relação entre os falares de Vitória da Conquista, Salvador e Teófilo Otoni no que se refere a configuração da F_0

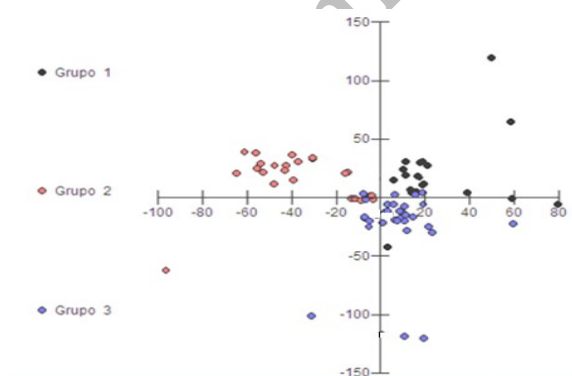
das vogais /a/, /i/ e /u/ tanto em sílaba tônica, quanto pretônica, produzida por sujeitos femininos e masculinos.

4.1 A frequência fundamental das sílabas tônica e pretônica nucleadas pela vogal /a/ como marca dialetal

Os gráficos 1, 2, 3 e 4 referem-se aos resultados da análise multivariada discriminante obtidos para o /a/ na posição tônica de sujeitos femininos; pretônica de sujeitos femininos; tônica de sujeitos masculinos e pretônica de sujeitos masculinos, respectivamente. Nesses e nos demais gráficos desta seção, encontraremos a disposição discriminante da F_0 para as três cidades, sendo possível observar a relação entre elas. Nessa análise foram usados os valores médios da F_0 resultantes da mensuração obtida em três pontos diferentes da vogal: no início, no meio e no fim. Assim, seremos capazes de avaliar o quanto a fala dessas três cidades distam-se ou aproximam-se no que se refere à F_0 média de toda a vogal e seremos igualmente capazes de afirmar se esses lugares guardam entre si semelhanças ou diferenças.

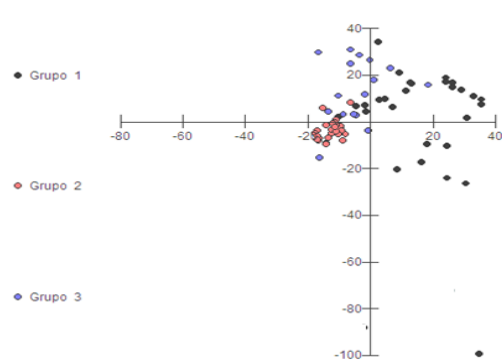
A vogal /a/, para os sujeitos femininos, mostrou-nos ser uma vogal que pode ser realizada de forma diferente nas três cidades, por um lado, e não tão diferente assim por outro. Seu comportamento vai depender do tipo sílaba: tônica e átona, conforme os gráficos 1 e 2.

Gráfico 1 - Análise multifatorial discriminante da F_0 da vogal /a/ **tônica** produzida por sujeitos **femininos** de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Gráfico 2 - Análise multifatorial discriminante da F_0 da vogal /a/ **pretônica** produzida por sujeitos **femininos** de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

A distribuição dos pontos, que representam os valores da F_0 , para os três grupos (Grupo 1, Salvador; Grupo 2, Teófilo Otoni e Grupo 3, Vitória da Conquista) ao longo dos gráficos nos dá uma dimensão de como essas três localidades se encontram em termos acústicos.

Observando a localização dos pontos nos 4 quadrantes (I – superior esquerdo; II- superior direito; III- inferior direito e IV- inferior esquerdo), podemos afirmar, no que se refere à produção da vogal /a/ tônica, que Salvador (Grupo 1, em preto) e Teófilo Otoni (Grupo 2, em vermelho) encontram-se em planos opostos: Teófilo Otoni no quadrante I e Salvador no quadrante II.

A produção dos falantes conquistenses na sílaba tônica encontra-se majoritariamente no quadrante III, com alguns pontos dispersos nos quadrantes I e II, sobrepondo-se à produção das produções de Teófilo Otoni e Salvador, respectivamente.

Um cenário diferente se descortina na produção da vogal /a/ em sílaba pretônica: a produção dos conquistenses se concentra no quadrante I com algumas ocorrências nos quadrantes II, sobrepondo-se à produção dos soteropolitanos e não mais no quadrante III, como nas tônicas. A produção dos teófilo-otonenses, que se concentra no quadrante I nas tônicas, concentra-se, na pretônica, no quadrante IV, estendendo-se timidamente para o quadrante I, sobrepondo-se à produção conquistense. A distribuição da produção dos soteropolitanos atinge os quadrantes II e III, além de alcançar o quadrante I, com leve sobreposição com a produção conquistense.

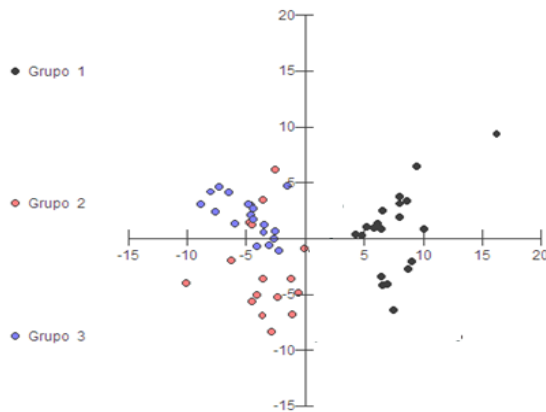
Em suma, podemos afirmar que Salvador, Teófilo Otoni e Vitória da Conquista possuem marcas específicas de F_0 para a vogal /a/, na sílaba tônica. Esse resultado acena para a possibilidade de se investigar, a partir de testes de percepção, se a F_0 pode ser parâmetro importante na distinção dialetal.

Para a sílaba pretônica, podemos sintetizar os dados da vogal /a/, afirmando que há proximidade entre os dados de Teófilo Otoni e Vitória da Conquista e Vitória da Conquista e Salvador. Apesar de não haver sobreposição entre as ocorrências de Teófilo Otoni e Salvador, podemos observar uma aproximação entre essas ocorrências no limite dos quadrantes I e IV (cf. gráfico 2). Diante disso, a vogal /a/ de sílaba pretônica pode não ser parâmetro razoável para distinguir falantes dessas três localidades. Essa separação pouco delimitada pode levar à indistinção dialetal dos falantes dessas cidades.

Enfim, no que se refere à vogal /a/ produzida por sujeitos femininos, podemos afirmar que as três localidades apresentam perfis diferentes especificamente na sílaba tônica, apesar de observarmos alguns pequenos pontos de contato entre Teófilo Otoni e Vitória da Conquista e Salvador e Vitória da Conquista, sendo nítida a distância entre Teófilo Otoni e Salvador.

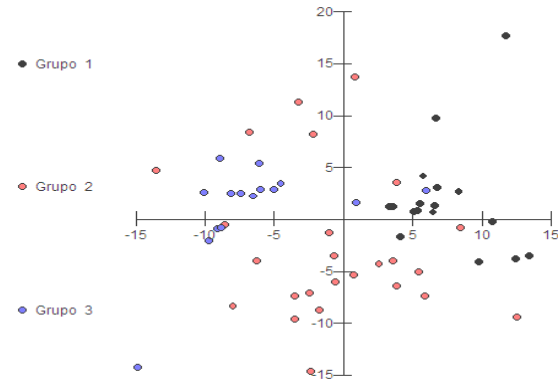
Esse mesmo padrão pode, de forma geral, ser encontrado na produção da vogal /a/ de sujeitos masculinos na sílaba tônica (cf. gráfico 3), uma vez que Teófilo Otoni e Salvador possuem suas ocorrências em quadrantes opostos.

Gráfico 3 - Análise multifatorial discriminante da F_0 da vogal /a/ **tônica** produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Gráfico 4 - Análise multifatorial discriminante da F_0 da vogal /a/ **pretônica** produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Podemos encontrar, contudo, nos dados dos sujeitos masculinos, uma distribuição das ocorrências nos quadrantes com algumas diferenças se comparamos com os dados femininos. Diferentemente do que se observa para os sujeitos femininos, os sujeitos masculinos conquistenses só possuem produção no quadrante I, com algumas poucas ocorrências no quadrante IV, sobrepondo-se às ocorrências de Teófilo Otoni. As produções de Teófilo Otoni concentram-se no quadrante IV com algumas poucas ocorrências no quadrante I, onde se encontra a maioria das produções conquistenses.

Embora os dados femininos e masculinos apresentem algumas diferenças de distribuição da F_0 entre as três cidades, o padrão geral de distribuição das ocorrências desse parâmetro para os sujeitos masculinos segue o dos sujeitos femininos. Como foi observado nos dados dos sujeitos femininos, as três localidades apresentam perfis diferentes da F_0 na vogal /a/ de sílaba tônica, embora haja pontos de contato entre Teófilo Otoni e Vitória da Conquista. Teófilo Otoni e Salvador encontram-se distante no que se refere a esse parâmetro. Para os dados masculinos podemos observar ainda uma oposição entre Vitória da Conquista e Salvador na sílaba tônica.

Como ocorre com os sujeitos femininos, a produção da vogal /a/ em sílaba pretônica por sujeitos masculinos também não constitui um parâmetro satisfatório para identificar os falantes das três cidades avaliadas (cf. gráfico 4). Não há delimitação dessas produções vocálicas, sendo possível encontrar ocorrências de Teófilo Otoni nos quadrantes I, II, III e IV, e produções conquistenses dispersas nos quadrantes I, II e III. Vale ressaltar que, apesar de não verificarmos limites precisos das produções dos sujeitos de Teófilo Otoni e de Vitória da Conquista, as produções dos soteropolitanos estão circunscritas ao quadrante II.

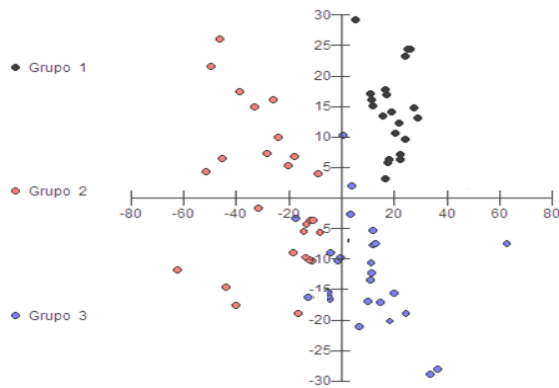
4.2 A frequência fundamental das sílabas tônica e pretônica nucleadas pelas vogais /i/ e /u/ como marca dialetal

Diante das constatações e reflexões até aqui desenvolvidas para a vogal /a/, resta-nos perguntar como será a distribuição da frequência fundamental nas três cidades investigadas no que se refere às vogais altas /i, u/. A F_0 sobre essas vogais pode delimitar dialetos, como o faz a vogal /a/ em sílaba tônica? A resposta a essa pergunta pode ser obtida a partir das análises dos dados dispostos nos gráficos de 5 a 8.

Um olhar geral lançado sobre esses gráficos permite afirmar que estas duas vogais, /i, u/, apresentam a distribuição dos valores de F_0 entre as localidades investigadas muito similar àquela encontrada para a vogal /a/.

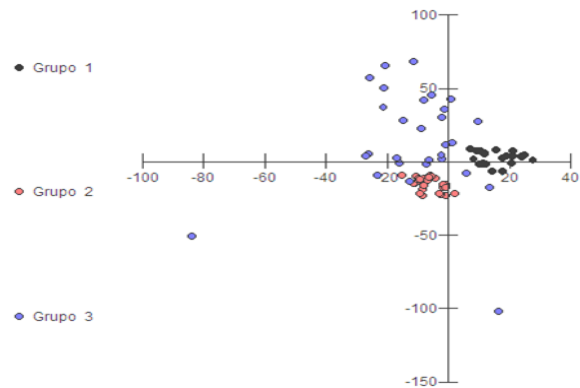
Avaliando a análise multifatorial discriminante das produções da vogal /i/ em sílaba tônica de sujeitos femininos, podemos verificar que os três grupos podem ser delimitados em três zonas relativamente bem delimitadas. As ocorrências de Teófilo Otoni se encontram distribuídas entre os quadrantes I e IV; as de Salvador encontram-se centradas no quadrante II e Vitória da Conquista majoritariamente no quadrante III, com algumas poucas ocorrências no quadrante IV, coocorrendo com produções de Teófilo Otoni e uma única ocorrência no quadrante II, próxima das ocorrências de Salvador (ver gráfico 5).

Gráfico 5 - Análise multifatorial discriminante da F_0 da vogal /i/ **tônica** produzida por sujeitos femininos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Gráfico 6 - Análise multifatorial discriminante da F_0 da vogal /i/ **pretônica** produzida por sujeitos femininos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria

Diante dessa distribuição, podemos afirmar que, em sílaba tônica, Teófilo Otoni e Salvador e Vitória da Conquista e Salvador encontram-se distantes sem qualquer ponto de interseção, e Vitória da Conquista e Teófilo Otoni podem ter traços que as aproximam, mesmo que levemente. A única ocorrência de produção de Vitória da Conquista no quadrante no qual predominam as produções de Salvador não é suficiente para afirmarmos que haja um ponto de contato substancial entre esses dois falares nesse contexto.

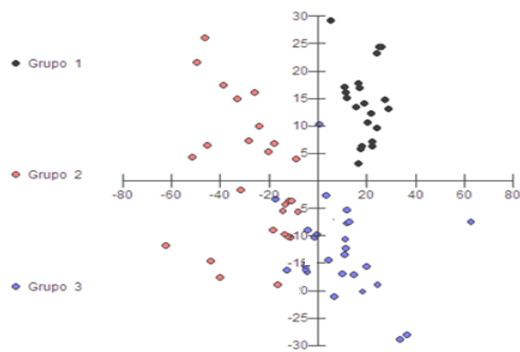
Na sílaba pretônica, de acordo com o gráfico 6, a vogal /i/ produzida por conquistenses encontra-se preferencialmente no quadrante I, havendo, contudo, algumas poucas ocorrências no quadrante II e III, nos quais estão concentradas as ocorrências soteropolitanas, além de encontrarmos algumas poucas ocorrências também no quadrante IV, no qual estão as ocorrências de Teófilo Otoni, as quais ocorrem exclusivamente nesse quadrante, mais próximas do zero do que na sílaba tônica.

Ainda observando as zonas de distribuição das produções do /i/ no gráfico 6, podemos verificar que as produções do /i/ dos três grupos giram em torno do ponto de interseção dos eixos x e y: Salvador e Teófilo Otoni com maior aglomeração nesse ponto e Vitória da Conquista com alguma.

Esse dado lança-nos pistas para a possibilidade de, na sílaba pretônica, a vogal /i/, à semelhança da vogal /a/, não ser exatamente um parâmetro diferenciador de dialetos. O limite separador desses três dialetos, para a vogal /i/, na sílaba pretônica, é pequeno, o que abre margem para pouca diferenciação entre eles.

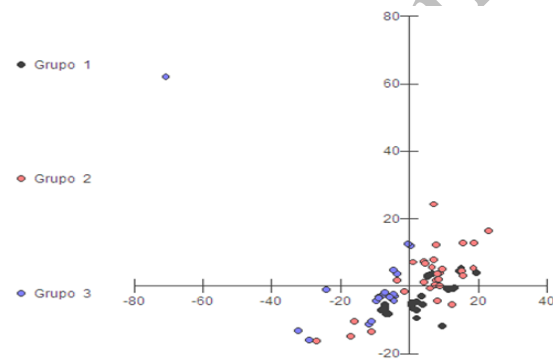
O quadro de distribuição da F_0 da vogal /i/ produzida por sujeitos masculinos em quase nada se difere dos quadros de distribuição da frequência fundamental delineados até aqui: na sílaba tônica, observamos limites melhor delineados entre os três grupos e frágil delimitação deles na sílaba pretônica, conforme dados dos gráficos 7 e 8, respectivamente.

Gráfico 7 - Análise multifatorial discriminante da F_0 da vogal /i/ **tônica** produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Gráfico 8 - Análise multifatorial discriminante da F_0 da vogal /i/ **pretônica** produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Assim, temos, novamente, as produções soteropolitanas concentradas exclusivamente no quadrante II; produções teófilo-otonenses distribuídas entre os quadrantes I e IV, sem qualquer interseção com as produções de Salvador, e as produções conquistenses, na sua maioria, localizadas no quadrante III com alguns pontos de interseção com as produções de Teófilo Otoni no quadrante IV, sem qualquer interseção com as produções de Salvador.

É possível, dessa forma, visualizar, de forma bem clara, as zonas de distribuições das produções da vogal /i/ em sílaba tônica dos falantes de cada uma das cidades, situação bem diferente daquela encontra para a sílaba pretônica (cf. gráfico 8), como tem sido a tendência.

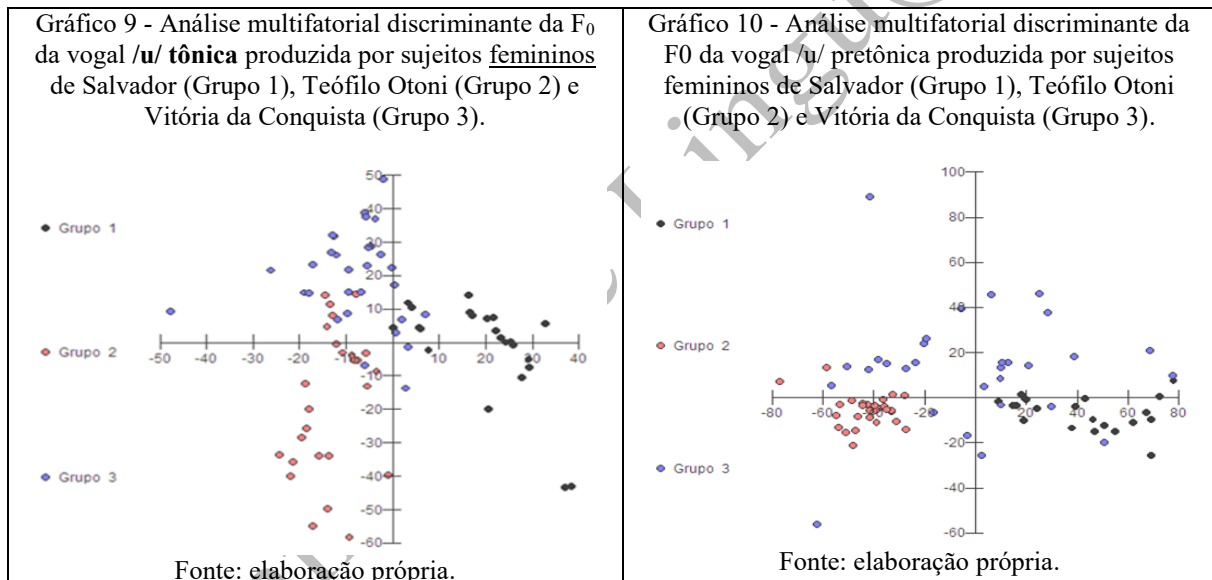
Na distribuição dos valores da F_0 da vogal /i/ em sílaba pretônica, para sujeitos masculinos, o quadrante I comporta somente algumas poucas produções de conquistenses, produções essas que se concentram no quadrante IV que recebe também algumas ocorrências de Teófilo Otoni cujo maior montante de ocorrências encontra-se no quadrante II e algumas outras poucas no quadrante III (cf. gráfico 8).

Ainda de acordo com os dados do gráfico 8, podemos verificar que as produções dos soteropolitanos masculinos para a vogal átona podem ser encontradas em três dos quatro quadrantes: algumas ocorrências no quadrante II e III, juntamente com as ocorrências de

produções de Teófilo Otoni, e no quadrante IV, em coocorrência com produções de conquistenses e de teófilo-otonenses.

É, pois, com a distribuição que ora apresentamos, impossível delimitar o falar de Salvador, Teófilo Otoni e Vitória da Conquista, quando se tem de considerar a vogal /i/ em sílaba pretônica produzida por sujeitos masculinos; o que não é, vale ressaltar, diferente do que foi também encontrado para os sujeitos femininos.

Passando, finalmente, para a avaliação da F_0 de sílabas tônicas e pretônicas com /u/ produzida por sujeitos femininos e masculinos (gráficos 9, 10, 11 e 12), encontramos uma distribuição dessa variável bem próxima ao que até agora encontramos para as demais vogais, exceto para /u/ pretônica produzida por sujeitos femininos, que, como veremos, mostrou possuir da F_0 bem delimitada pelo menos para Salvador e Teófilo Otoni.



Verificamos nos dados encontrados no gráfico 9 grande quantidade de ocorrências das produções conquistenses no quadrante I e algumas poucas nos quadrantes II e IV, coocorrendo, respectivamente, com dados de Salvador, que se distribuem entre os quadrantes II e III, e de Teófilo Otoni, cujas ocorrências encontram-se majoritariamente no quadrante IV, podendo ser observadas algumas outras no quadrante I.

Na sílaba tônica com vogal /u/ produzida por sujeitos femininos, à semelhança do que já observamos anteriormente, as F_0 médias possuem zonas de distribuição para as três cidades avaliadas bem delimitadas, observando pequena sobreposição entre Vitória da Conquista e Teófilo Otoni e Vitória da Conquista e Salvador, bem como ausência de sobreposição entre

ocorrências de Teófilo Otoni e Salvador. As produções dessas duas cidades encontram-se distantes entre si.

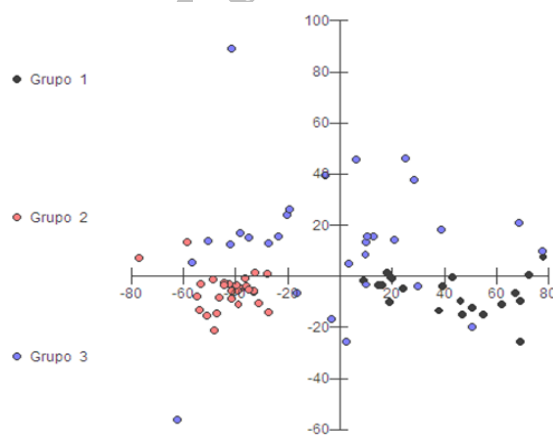
Em sílaba pretônica, a vogal /u/ produzida por sujeitos femininos apresentou distribuição da F_0 diferente da observada para as vogais /a/ e /i/, pois podemos encontrar as produções das três cidades em zonas mais delimitadas do que se espera para as pretônicas, sem confluência entre as produções de Salvador e Teófilo ou proximidades entre as duas, à semelhança do que encontramos para /a/ e /i/ tônicos.

Assim, temos as produções de Vitória da Conquista bem concentradas nos quadrantes I e II com algumas poucas ocorrências no quadrante III, junto com Salvador, e ainda algumas outras poucas no quadrante IV, junto com Teófilo Otoni. Salvador e Teófilo Otoni possuem suas produções localizadas majoritariamente nos quadrantes III e IV, respectivamente, com produções mínimas no quadrante superior (I para Teófilo Otoni e II para Salvador), juntamente com ocorrências de Vitória da Conquista.

Apesar de pequenas sobreposições entre as zonas de ocorrências, a vogal /u/ de sílaba pretônica de sujeitos femininos pode contribuir para a identificação de falantes como pertencentes a essa ou aquela localidade.

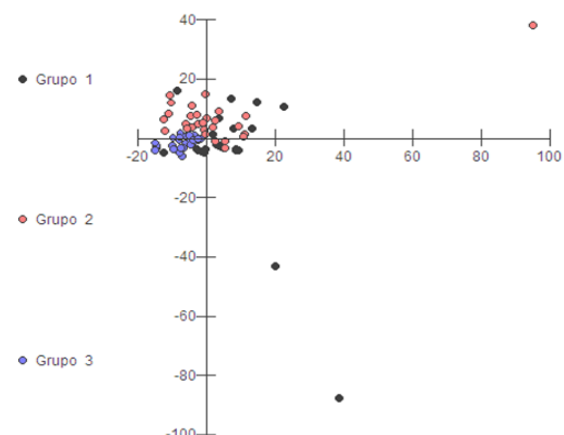
A vogal /u/ produzida por sujeitos masculinos apresenta distribuição da F_0 diferenciada em sílaba tônica e pretônica, como ocorre com as demais vogais /a/ e /i/ nesses contextos de tonicidade, como podemos verificar, respectivamente, nos gráficos 11 e 12.

Gráfico 11 - Análise multifatorial discriminante da F_0 da vogal /u/ tônica produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

Gráfico 12 - Análise multifatorial discriminante da F_0 da vogal /u/ pretônica produzida por sujeitos masculinos de Salvador (Grupo 1), Teófilo Otoni (Grupo 2) e Vitória da Conquista (Grupo 3).



Fonte: elaboração própria.

As produções soteropolitanas encontram-se no quadrante III e as teófilo-otonenses no quadrante IV, sem qualquer ponto de contato entre elas. A produção conquistense distribui-se majoritariamente entre os quadrantes I e II e timidamente entre os quadrantes III e IV (cf. gráfico 11).

Um cenário diferente é o encontrado para a vogal /u/ em sílaba pretônica (cf. gráfico 12), que se caracteriza por uma concentração das ocorrências das três localidades no ponto de interseção dos eixos x e y. Não há qualquer distância entre as zonas de produção dessa vogal, nesse contexto, o que nos sugere que há uma semelhança de produção dessa vogal na produção vocálica dos falantes dessas cidades, nesse tipo de sílaba.

Após avaliarmos as médias de F_0 das vogais /a/, /i/ e /u/ em sílabas tônica e pretônica, somos capazes de determinar em que medida esse parâmetro acústico pode contribuir para a identidade dialetal.

Nossas análises permitem afirmar que há uma diferença do comportamento da F_0 nas vogais tônicas /a,i,u/ dessas três cidades. Nossas análises permitem também afirmar, ainda com relação a esse tipo silábico, que, de fato, não há nenhuma semelhança nesse parâmetro acústico nos falares de sujeitos nativos de Salvador e de Teófilo Otoni. Falantes de Vitória da Conquista, por sua vez, possuem uma produção bem característica, mas com alguns pontos de contatos com Salvador e Teófilo Otoni.

Os nossos resultados concordam com os resultados encontrados por Cunha (2000) para os dialetos de São Paulo e Rio de Janeiro. Vitória da Conquista, à semelhança de São Paulo e Rio de Janeiro, possui um padrão acústico que evidencia certa influência de outros falares. São Paulo e Rio de Janeiro apresentaram padrões prosódicos semelhantes aos das cidades nordestinas e gaúchas ao mesmo tempo. Esse resultado seria consequência da grande concentração de diferentes falares nessas duas cidades? É possível que sim. Da mesma forma, Vitória da Conquista, que recebe cidadãos de várias outras cidades baianas e mineiras pode possuir um padrão de F_0 que esteja sofrendo influência desses diferentes falares.

O dado encontrado para Vitória da Conquista talvez possa ser explicado pelo fato de essa cidade, dada à proximidade geográfica, ter uma afinidade grande com Teófilo Otoni, havendo grande contato entre os moradores das duas cidades. Por outro lado, Vitória da Conquista, apesar da distância geográfica que se encontra da capital do estado, Salvador, mantém estreito contato com a metrópole. Conquistenses com frequência vão estudar e trabalhar em Salvador, da mesma forma que há muitos soteropolitanos vivendo em Vitória da

Conquista,. Assim, é possível que os falantes de Vitória da Conquista possuam uma forma de dizer que permita encontrar traços de influência de Teófilo Otoni e de Salvador.

Podemos afirmar que a relação existente entre os falantes nativos de Vitória da Conquista e Teófilo Otoni e de Vitória da Conquista e Salvador foi expressa na análise discriminante da F_0 . Nas variedades linguísticas estudadas, esse parâmetro acústico pode ajudar a compreender as demarcações dialetais, assim como foi verificado para algumas variedades do inglês americano (JACEWICZ; FOX, 2015) e de certa maneira para os dialetos catarinenses, como investigado por Sosa, Nunes e Seara (2013).

Dados os resultados encontrados que mostram que esses falares possuem padrão de F_0 delineado em dois grandes grupos (Vitória da Conquista/ Teófilo Otoni x Salvador), podemos inferir que esse parâmetro ao ser processado pelo ouvinte, pode despertar o sentimento de pertencimento à comunidade de fala de Vitória da Conquista, por exemplo; isto porque a F_0 pode atuar significativamente na percepção e reconhecimento de dialetos (PERES, 2011), além, obviamente, da ocorrência das vogais médias abertas em sílabas pretônicas muito comuns nesse dialeto e do padrão de curva de F_0 das vogais médias abertas que não apresenta diferença entre as sílabas tônicas e pretônicas, o que é pouco esperado, pois sílabas tônicas tendem a ter maiores valores de F_0 (OLIVEIRA; RIBEIRO; PACHECO, 2007, PACHECO; OLIVEIRA; RIBEIRO, 2013).

E o que dizer dos demais parâmetros acústicos prosódicos, intensidade e duração, que são outros dois parâmetros importantes na análise acústica prosódica? Trabalhos futuros responderão a essa pergunta, para a qual temos a hipótese de que, a duração, a intensidade somada a F_0 contribuem simultaneamente para delimitação desses falares.

5. Considerações finais

Neste artigo, investigamos o quanto os falares de Vitória da Conquista, Salvador e Teófilo Otoni se distanciam, comparando-os a partir da avaliação de F_0 das vogais /a/, /i/ e /u/ de sílabas tônicas e pretônicas.

Por outro lado, Vitória da Conquista, uma cidade baiana com modos de mineiro, possui um falar baiano-mineiro. A produção vocálica conquistense, apesar de possuir uma delimitação própria, pode igualmente estender-se para um padrão mineiro ou baiano.

Nossos resultados mostram que as diferenças dialetais podem estar para além das vogais médias pretônicas, que podem sofrer abaixamento ou alteamento, como tem sido

tradicionalmente descrito na literatura. As outras vogais do sistema fonológico, na sílaba tônica ou na pretônica podem ser coparticipantes de delimitação dialetal.

Diante disso, nosso trabalho descortina novas perguntas para trabalhos futuros, tais como: qual o papel da intensidade e da duração na delimitação dialetal?; cidades contíguas possuem falares contíguos, como uma espécie de empatia dialetal?; qual o alcance dessa “empatia”, caso ela exista?; as mesorregiões, além das afinidades econômicas e sociais podem ter afinidades no seu modo de falar?.

Assim, nosso trabalho não se esgota por aqui. Temos ciência de que trouxemos uma pequena contribuição para os estudos linguísticos, mas muito ainda deve ser feito para uma compreensão maior do falar dos brasileiros.

Referências bibliográficas

ATKINSON, K. Language identification from non-segmental cues. **Working Papers in Phonetics** (UCLA), v. 10, p. 85-89, 1968.

AYRES, M.; AYRES JÚNIOR, M.; AYRES, D. L.; SANTOS, A. A. 2007. **BIOESTAT – Aplicações estatísticas nas áreas das ciências bio-médicas**. Ong Mamiraua. Belém: Ong Mamiraua; 2007.

BISOL, L. **Harmonização vocálica: uma regra variável**. Rio de Janeiro, UFRJ, 1981. Tese de Doutorado.

BOERSMA, P.; WEENINK, D. **Praat: doing phonetics by computer**. (Version 5.1.43) [Programa de computador]. 2006. Disponível em <http://www.praat.org/>.

BONTE, R. **Can you identify a language by its prosody?** Dissertação (Mestrado) s.f. 1975. University of California, 1975.

BORTONI, S. et al. Um estudo preliminar do [e] pretônico. **Cadernos de estudos lingüísticos**, v. 20, 75-90, 1991.

CARDOSO, S. A. L. M.; MOTA, J. A., O atlas linguístico do Brasil (uma descrição do português brasileiro do Oiapoque ao Chuí. In: FONSECA-SILVA; PACHECO, V. LESSA-DE-OLIVEIRA, A. S. C. (Orgs.) **Em torno da língua(gem): questões de análises**. Vitória da Conquista: Edições Uesb, p. 139-159, 2007.

CELIA, G. F. **Variação das vogais médias pretônicas no português de Nova Venécia (ES)**. (Dissertação de Mestrado) Campinas, UNICAMP, 2004.

CLOPPER, C. G.; PISONI, D. B., Perception of dialect variation. In: PISONI, D. B.; REMEZ, R. E. (EDS.) **The handbook of speech perception**. Oxford: Blackwell Publishing, 2005. p. 313-337. **crossref** <http://dx.doi.org/10.1002/9780470757024.ch13>

COLOMARCO, M. Aspectos prosódicos do corpus do projeto ALiB: o padrão interrogativo na fala baiana e carioca. In: _____. SANTOS, D. V. (org.) **inicia** – Revista da Graduação em Letras da UFRJ. Rio de Janeiro. Ed. Faculdade de Letras/URFJ, 2005, pp. 35-43

CRANDAL, J. B. The sounds of speech. **Bell Sysy. Tech J.**, v. 4, p. 586-626, 1925.

CRUZ, R. et al. As vogais médias pretônicas no português falado nas ilhas de Belém (PA). In: _____. ARAGÃO, M. do S. (org.). **Estudos em fonética e fonologia no Brasil**. João Pessoa: GT Fonética e Fonologia / ANPOLL, 2008.

CUNHA, C. **Entoação Regional no Português do Brasil**. 2000. s.f Tese (Doutorado) Faculdade de Letras, UFRJ, Rio de Janeiro.

GUIMARÃES, E. D. **A variação entoacional da ilha de Mosqueiro**: contribuições para a formação do Atlas Prosódico Multimídia do Português do Norte do Brasil (AMPER – NORTE). Dissertação (Mestrado). 2013. 143f. Instituto de Letras e Comunicação, Universidade Federal do Pará.

IBGE. **População brasileira**: Dados das pesquisas do Censo/2014. Disponível em: http://www.ibge.gov.br/censo2010/dados_divulgados/index.php.

JACEWICZ, E.; FOX, R. A. I Intrinsic fundamental frequency of vowels is moderated by regional dialect. **The Journal of the Acoustical Society of America**, v. 138, n. 4, p. 405-410, 2015.

KAILER, D. A. Alçamento da vogal pretônica /o/ em duas regiões paranaenses. **SIGNUM**: Estud. Ling., Londrina, n. 15/1, p. 201-221, jun. 2012. Disponível em: <http://www.uel.br/revistas/uel/index.php/signum/article/view/11668/11175>.

KENT, R. D.; READ, C. **The acoustic analysis of speech**. 2. ed. Cambridge: Singular, 1992.

LEE, S. H.; OLIVEIRA, M. A. de. Variação Inter- e Intra-Dialetal no Português Brasileiro: Um Problema para a Teoria Fonológica. In: _____. OLIVEIRA, D. da H.; COLLISCHONN, G. (Org.). **Teoria Linguística: fonologia e outros temas**. João Pessoa, 2003. p. 67-91.

LEHISTE, I.; PETERSON, G. Some basic considerations in the analysis of intonation. **The Journal of the Acoustical Society of America**, v. 33, p. 419-425, 1961. **crossref** <http://dx.doi.org/10.1121/1.1908681>

LIRA, Z. **A entoação modal em cinco falares do nordeste brasileiro**. 2009. s.f. Tese (Doutorado) – Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa.

MAIDMENT, J. Voice fundamental frequency characteristics as language differentiators. **Speech and Hearing Working in Progress**, v. 16, 1976. p. 74-93.

NASCENTES, A. **O linguajar carioca**. Rio de Janeiro: Simões, 1953.

NINA, T. **Aspectos da variação fonético-fonológica na fala de Belém**. 1991. 216f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

OHALA, J. J.; GILBERT, J. B. Listener's ability to identify languages from their prosody. **Report of the Phonology Laboratory**, v. 2, 1978. p. 126-132.

OHALA, J.J.; EUKEL, B. W. Explaining the intrinsic pitch of vowels. In: Channon, R.; Shockey, L. (eds.) **In honor of Ilse Lehiste**. Dordrecht: Foris, 1987. p. 207-215. **crossref** <http://dx.doi.org/10.1515/9783110886078.207>

OLIVEIRA, J. S. N. **Análise acústico-perceptual das frases exclamativas e interrogativas realizadas por falantes de Vitória da Conquista/Ba**. 2014. 79 p. Dissertação (Mestrado em Linguística). Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, 2014.

OLIVEIRA, J. S. N.; RIBEIRO, P. J.; PACHECO, V. Realização das vogais médias abertas no dialeto de Vitória da Conquista/BA. In: _____. FONSECA-SILVA, M.C.; PACHECO, V. SILVA, E.G. (Org.). **Pesquisas em Estudos da Linguagem III**, Vitória da Conquista, v. 3, n.1, 2007. p. 67-74.

PACHECO, V.; OLIVEIRA, M.; RIBEIRO, P. de J. **Em busca da melodia nordestina: as vogais médias pretônicas de um dialeto baiano**. *Linguística* / Vol. 29 (1), junho 2013: 165-187. Disponível em: <http://www.scielo.edu.uy/pdf/ling/v29n1/v29n1a08.pdf>. Acesso em outubro de 2014.

PERES, D. O. **O papel da prosódia na identificação das variedades regionais do português brasileiro**. Dissertação (Mestrado). 151 f. 2011. Faculdade de Ciências Humanas. Universidade de São Paulo. 2011.

PETERSON, G. E. BARNEY, H.L. Control methods used in a study of the vowels. **The Journal of the Acoustical Society of America**, v. 24, 1952. p. 175-184.

SEARA, I. C. REBOLLO-COUTO, L. Entoação de frases declarativas e interrogativas no falar fluminense e catarinense. Congresso Internacional de la ALFAL, 16, 2011, Alcalá de Henares (Espanha). **Actas....** Alcalá de Henares (Espanha): Universidade de Alcalá, 2011. p.4203-4208.

SILVA, J. C. B.; CUNHA, C.S. Caracterização prosódica dos falares brasileiros: a questão total em Recife, Rio de Janeiro e Florianópolis. **Antares**, v 3, n. 6, 2011. p. 282-294.

SOSA, J. M.; NUNES, G. V. SEARA, I. C. Variação prosódica das sentenças interrogativas totais no falar catarinense: um estudo experimental. **Leitura**, n. 52, 2013. p. 139-163.

van BEZOOIJEN, R.; GOOSKENS, C. Identification of language varieties: the contribution of different linguistic levels. **Journal of Language and Social Psychology**, v. 18, , 1999. p. 31-48.

VIEGAS, M. C. **Alçamento de vogais médias pretônicas**: uma abordagem sociolinguística. Dissertação (Mestrado) 1987. s.f. Belo Horizonte, UFMG, 1987.

VOGELY, A. N.; HORA, D. Harmonia Vocálica no dialeto recifense. **Organon**, v. 28, n.54, 2013. p. 63-81.

WHALEN, D. H.; LEVITT, A. G. The universality of intrinsic F₀ of vowels. **Journal of Phonetics**, v. 23, 1995. p. 349-366. **crossref** [http://dx.doi.org/10.1016/S0095-4470\(95\)80165-0](http://dx.doi.org/10.1016/S0095-4470(95)80165-0)

ZENDRON DA CUNHA, K. **Sentenças exclamativas do português brasileiro**: padrão entoacional e sintaxe. Dissertação (Mestrado em Letras). Curitiba: Universidade Federal do Paraná, 2012

Artigo recebido em: 01.11.2015

Artigo aprovado em: 30.04.2016

A aquisição do sistema vocálico do português por falantes nativos da variedade rio-platense de espanhol: uma discussão sobre a bidirecionalidade da transferência vocálica

The acquisition of the Portuguese vowel system by speakers of the Riverplate Spanish variety: a discussion on the bi-directionality of vowel transfer

Letícia Pereyron*
Ubiratã Kickhöfel Alves**

RESUMO: Este estudo visa a investigar a produção vocálica de hispano-falantes (variedade Rioplatense) em português (L2), bem como verificar alterações no espaço vocálico materno destes falantes em termos de influências exercidas pela segunda língua (L2). Os dados foram coletados de 10 participantes argentinos - 5 monolíngues e 5 bilíngues. Verificaram-se os valores de F1, F2 e de duração, em espanhol (L1) e em português (L2), de falantes bilíngues. Verificou-se, também, a produção em espanhol por parte do grupo de monolíngues, residentes na Argentina, tomados como controle. Os resultados mostraram, em termos de altura e anterioridade/posterioridade, que os participantes não realizam as distinções entre as vogais médias-baixas e médias-altas na L2. Entretanto, verificou-se que suas vogais médias-altas ocupam um espaço mais alto no espaço acústico, como se o sistema estivesse se adaptando para a formação das médias-baixas. Tal tendência foi verificada tanto nas produções da L2 como na L1, mas não nas produções dos monolíngues argentinos. Quanto à duração, os bilíngues apresentaram vogais mais longas na L2 e produziram, na L1, vogais mais longas do que aquelas dos monolíngues. Tais evidências, analisadas sob uma concepção de língua enquanto Sistema Adaptativo Complexo (DE BOT *et al.*, 2013), indicam uma multidirecionalidade na transferência linguística.

ABSTRACT: This study aims to investigate the vowels produced by L1 Riverplate Spanish speakers of Portuguese (L2) and verify modifications in the vowel space of their L1 due to L2 influences. Data were collected from 10 participants - 5 bilinguals and 5 monolinguals. F1, F2, and duration values were verified in the Spanish (L1) and in the Portuguese (L2) of bilingual speakers, while such values were also verified in the Spanish produced by the monolingual participants (control group). Results show that, in terms of the front-back dimension, bilingual participants do not differ in mid-low and mid-high vowel production in Portuguese. However, it was found that their production of mid-high vowels is located at a higher position in the acoustic space, as if the system were being prepared for mid-low vowel formation. This pattern was verified in both their L2 and L1. This tendency was not found in the productions of monolinguals. As for vowel duration, bilinguals show longer vowels in the L2 than in their L1 and their vowels in Spanish are longer than those produced by monolinguals. These results, in light of a conception of language as a Complex Adaptive System (DE BOT *et al.*, 2013), suggests a multidirectionality in terms of language transfer.

* Doutoranda do PPG em Letras da UFRGS. Bolsista CAPES.

** Doutor em Letras. Professor do PPG em Letras da UFRGS. Bolsista de Produtividade do CNPq.

PALAVRAS-CHAVE: Aquisição do português brasileiro. Sistema vocálico. Bílingues espanhol-português.

KEYWORDS: Brazilian Portuguese phonetic acquisition. Vowel system. Portuguese-Spanish bilinguals

1. Introdução

Pesquisas recentes mostram que a influência interlinguística no âmbito do bilinguismo não segue uma direção unilateral, mas sim bidirecional (GROSJEAN, 2013), já que depende de múltiplos fatores externos, como a tipologia linguística, a proficiência em cada língua, a consciência metalinguística, a idade, dentre outros (CENOZ, 2001; HERDINA; JESSNER, 2002; JESSNER, 2008; ORTEGA, 2009; BRITO, 2011). No âmbito bilíngue, um sistema é capaz de influenciar o outro e, ainda, essa influência pode dar-se da língua menos entrincheirada para a dominante, mesmo que amenamente, conforme os autores supracitados. Além disso, o desenvolvimento de um novo sistema pode exercer efeitos, de diversas maneiras, sobre sistemas prévios (LARSEN-FREEMAN, 1997, 2011; BECKNER *et al.*, 2009; ELLIS, 2011; DE BOT, LARSEN-FREEMAN, 2011).

As teorias que conseguem acolher essas modificações nos sistemas linguísticos são a Teoria dos Sistemas Dinâmicos (doravante TSD) e a Teoria da Complexidade (LARSEN-FREEMAN, 1997, 2011; LARSEN-FREEMAN; ELLIS, 2006; ELLIS, 2007, 2011; LARSEN-FREEMAN; CAMERON, 2008; DE BOT; LOWIE; VERSPOOR, 2007; DE BOT *et al.*, 2013), visto que tendem a explicar a ‘mudança’, no sentido de que postulam que uma simples modificação em uma parte do sistema pode causar alterações em todo o seu restante. Assim como Flege (1995), tais teorias consideram que os aspectos de L1 e L2 interagem em um espaço fonológico comum.

Sob a perspectiva dinâmica que considera a língua como um sistema adaptativo complexo (doravante, CAS¹), uma mudança pode gerar outra e um ciclo ilimitado é formado, já que alteração gera alteração. Assim, pode-se pensar que o processo de aquisição (ou desenvolvimento²) de uma língua - materna, segunda ou até mesmo terceira – nunca alcança um estágio final, uma vez que está sempre sujeito à mudança e, ainda, é repleto de crescimento

¹ Do inglês, Complex Adaptive System.

² Os termos ‘aquisição’ e ‘desenvolvimento’ serão empregados, neste trabalho, com um caráter intercambiável. O termo ‘desenvolvimento’, e não ‘aquisição’, é empregado por dinamicistas como Larsen-Freeman (1997, 2011), De Bot, Lowie e Verspoor (2007) e Ellis (2011), uma vez que o referido termo não remete ao caráter de alvo, término, pré-definido e estanque, como o termo ‘aquisição’ remete. Esses autores focam no processo e não no produto final.

ao longo do percurso. Além disso, em um CAS, não há linearidade nem unilateralidade, portanto, em um falante bilíngue, é possível prever que a presença da L2 possa, por sua vez, alterar a L1.

Com base nessas premissas, pretende-se, com esta pesquisa, verificar o dinamismo dos sistemas linguísticos no que diz respeito à questão da direcionalidade das influências entre sistemas de primeira e segunda língua, respectivamente, o espanhol e o português. Desse modo, almeja-se investigar a influência de uma língua adquirida posteriormente, o português porto-alegrense, sobre o sistema materno, o espanhol rio-platense, de modo a verificar a possibilidade da direcionalidade L2-L1 em termos de transferência entre os sistemas, verificação essa ainda escassa nos estudos de bilinguismo. Para o cumprimento desse objetivo, foram comparados os sistemas vocálicos (valores de F1, F2 e duração) de falantes monolíngues de espanhol rio-platense, residentes na Província de Buenos Aires (Argentina), com a produção, em espanhol (L1) e português (L2), por falantes deste mesmo dialeto de língua-materna, residentes na cidade de Porto Alegre - RS. Com base em Feiden, Alves e Finger (2014) e Machry da Silva (2014), aprendizes hispano-falantes de português apresentam dificuldades na distinção entre as vogais médias-baixas [ɛ] e [ɔ] e médias-altas [e] e [o]; mais precisamente, esses aprendizes tendem a apresentar dificuldades, tanto em termos de percepção quanto em termos de produção, em distinguir tais sons em diferentes categorias fonético-fonológicas, visto que o espanhol dispõe de apenas cinco vogais tanto em sílabas átonas quanto tônicas.

Com o intuito de atender aos propósitos desta pesquisa e verificar a bidirecionalidade (L1-L2; L2-L1) da transferência vocálica na produção destes falantes hispânicos residentes em Porto Alegre, este trabalho encontra-se organizado da seguinte forma: a próxima seção corresponde ao referencial teórico, com base na Teoria dos Sistemas Dinâmicos Complexo-Adaptativos; a seção em seguida expõe os objetivos, as perguntas de pesquisa e os procedimentos metodológicos (as coletas, os participantes e a descrição das etapas de análise de dados); a seção subsequente traz os resultados, a próxima seção apresenta as conclusões seguidas de discussão e, por fim, são apresentadas as referências bibliográficas.

2. Pressupostos teóricos

Nesta seção, serão apresentados e discutidos os pressupostos teóricos utilizados para fundamentar a noção de língua como sistema dinâmico, complexo e adaptativo. De acordo com tais pressupostos, autores como Larsen-Freeman (1997, 2011), Ellis e Larsen-Freeman (2006),

De Bot, Lowie e Verspoor (2007), Larsen-Freeman e Cameron (2008), Beckner *et al.* (2009), Ellis (2007, 2011) e De Bot *et al.* (2013) aplicam a Teoria dos Sistemas Dinâmicos (TSD) aos estudos de desenvolvimento de línguas estrangeiras, mostrando que os processos envolvidos englobam múltiplos agentes em constante interação, de modo a conceberem língua como um CAS. A partir desta discussão, será discutida, também, a noção de não-linearidade dos sistemas envolvidos quando o assunto é complexidade.

2.1 Teoria dos Sistemas Dinâmicos

Sistemas dinâmicos recebem este nome porque resultam em mudança através de um fator essencial: o tempo (LARSEN-FREEMAN, 1997, 2011; DE BOT; LOWIE; VERSPOOR, 2007; BECKNER *et al.*, 2009; DE BOT *et al.*, 2013). Esses sistemas são compostos de múltiplos agentes em constante interação (daí vem a noção de ‘mudança’), e tal multiplicidade de agentes é o que deriva a complexidade desses sistemas.

Em sistemas complexos, o resultado do desenvolvimento ao longo do tempo não pode ser calculado com exatidão, não por falta de ferramentas certas, mas porque as variáveis que fazem parte deste sistema interagem e se modificam constantemente. O desenlace dessas variáveis, a menos que formassem parte de um sistema simples, não pode ser resolvido analiticamente. A trajetória dinâmica, nesse sentido, deve incluir as interações frequentes e recorrentes do próprio sistema e, dessa forma, não há equação que ofereça um valor exato, estanque, do sistema, até porque ele se encontra em mudança contínua (DE BOT; LOWIE; VERSPOOR, 2007).

Ainda em referência à trajetória dinâmica, De Bot, Lowie e Verspoor (2007) afirmam que, em sistemas dinâmicos, cada sistema é parte de outro sistema - da menor partícula molecular ao universo - e todos seguem os mesmos princípios dinâmicos operando em todos os níveis. Com o desenvolvimento ao longo do tempo, os subsistemas tendem a se estabelecer em estados específicos, conhecidos por estados atratores. Esses estados demonstram uma preferência pelo sistema, mas não são necessariamente estados completamente previsíveis.

Estados não preferíveis pelo sistema são chamados de estados repulsores. Os autores ilustram tanto o estado atrator quanto o repulsor com a analogia de uma bola rolando sobre uma superfície. Um buraco na superfície funciona como um estado atrator, já uma lombada, como repulsor. A trajetória da bola é o desenvolvimento. Os buracos podem ser rasos ou profundos

e, quanto mais profundidade apresentar o buraco, mais energia a bola precisará para rolar até o próximo buraco, ou seja, o atrator seguinte (DE BOT; LOWIE; VERSPOOR, 2007).

Pode-se adicionar a esta trajetória dinâmica, além dos estados atratores e repulsores, a dependência dos estados iniciais, isto é, diferenças mínimas no início do caminho podem trazer consequências dramáticas ao longo do percurso. O comportamento imprevisível de sistemas complexos dá-se devido a sua dependência nas condições iniciais. Uma pequena mudança em algum estágio inicial pode causar amplas implicações em um comportamento futuro. O exemplo mais utilizado na literatura para explicar a relevância dos estados iniciais é o efeito borboleta de Lorenz, que demonstra a interdependência entre todos os componentes do sistema: uma pequena mudança em um único elemento, em um determinado momento, pode causar mudanças drásticas em todo o sistema (LARSEN-FREEMAN, 1997).

Dessa forma, as características encontradas no processo de aquisição de línguas adicionais são semelhantes às características apresentadas pela Teoria dos Sistemas Dinâmicos Complexo-Adaptativos. Por essa razão, as línguas e seu desenvolvimento começam a ser analisados a partir desta concepção. A seguir, serão apresentadas características dos sistemas dinâmicos.

2.1.1 Características dos Sistemas dinâmicos, complexos e adaptativos

2.1.1.1 A língua é dinâmica

Larsen-Freeman (1997) sugere que os modelos empregados para descrever o processo de desenvolvimento de línguas adicionais não levam em consideração o dinamismo e a variabilidade da língua em uso como realmente é. Ao adotar a perspectiva de língua como CAS, a língua é considerada dinâmica, uma vez que dinamismo significa crescimento (*growth*) e mudança (*change*), isto é, envolve o desenvolvimento, a variação de algo ao longo do tempo. O sentido da palavra “dinâmico” refere-se a um “processo sincrônico” além de “crescimento/mudança diacrônica”, de modo que tal termo possa também ser interpretado como um processo isomórfico, uma vez que a utilização da língua sempre implica mudança. Tal mudança pode ocorrer no nível individual, como também no nível global. Sob essa perspectiva, o desenvolvimento da língua e sua organização são realizados de maneira orgânica, isto é, *bottom up*, em que o sistema não é um produto de regras, mas um produto de uso.

Ellis (2011) apresenta uma abordagem referente à língua como sistema dinâmico semelhante à abordagem descrita acima e afirma que as estruturas linguísticas emergem

ontogeneticamente, isto é, transformam-se do início à “completitude” do desenvolvimento, a partir do uso em contextos particulares. O autor afirma que o desenvolvimento é vagaroso e gradual e parte de uma grande dependência de aspectos concretos a esquemas linguísticos mais abstratos. Esse processo é extremamente dependente da frequência de tipo e de *token* no input. Ainda sob a abordagem dinâmica, Ellis (2011) pontua que o aprendizado individual de língua é uma propriedade emergente de um sistema dinâmico e inclui um grande número de influências sociais, dialetais, individuais e contextuais, envolvendo o aprendiz numa tensão consciente entre as produções da interlíngua e a evidência do *feedback* linguístico, pragmático ou metalinguístico, o que possibilita um desenvolvimento estruturado.

A língua é vista como dinâmica uma vez que, de acordo com Ellis (2011), seu uso engloba agentes e processos de níveis variados e sua emergência se origina a partir não apenas das interações entre eles como dentre eles através de uma variável fundamental, o tempo. O sistema é composto por interações entre diversos falantes da comunidade de fala (os agentes) e também por interações entre os agentes linguísticos do próprio sistema: pragmática, sintaxe, semântica, etc. Fatores cognitivos, sociais e ambientais interagem constantemente e geram relações não-lineares, multiplamente variadas e interativas com o tempo. Em outras palavras, o sistema, que é composto por vários agentes interagindo entre e dentre eles mesmos, desenvolve padrões linguísticos. Conforme mencionado anteriormente, os agentes podem ser os falantes de uma comunidade de fala – interagindo entre si – ou padrões da organização linguística como a fonologia, o léxico, a sintaxe, a semântica, a pragmática, dentre outros, que estão em constante interação e, enquanto interagem, causam mudança um no outro.

Larsen-Freeman (1997) discute não apenas o dinamismo da língua materna, mas também o dinamismo do desenvolvimento de línguas adicionais. A autora defende que o desenvolvimento dessas línguas é um processo dinâmico, uma vez que abrange a evolução do sistema linguístico do aprendiz. As gramáticas que empregam regras estáticas não levam em consideração o constante caráter de mudança da gramática interna da Língua Adicional (LA) do aprendiz. Já o novo sistema em desenvolvimento, assim como as línguas naturais, é instável, uma vez que é desenvolvida com empréstimos, inovações e preenchida com novas formas linguísticas. Tal instabilidade, no entanto, não deve ser considerada um risco à sistematicidade. A sistematicidade e instabilidade, sob uma perspectiva dinâmica, não representam ameaça uma a outra, mas se alimentam uma da outra.

2.1.1.2 A língua é adaptativa e complexa

Segundo Larsen-Freeman (1997), Beckner *et al.* (2009) e Ellis (2011), dentre outros, a língua é considerada complexa porque consiste em diversos subsistemas como, por exemplo, a fonologia, a morfologia, o léxico, a sintaxe, a semântica, a pragmática, e tais subsistemas apresentam uma relação de interdependência, em que uma mudança em algum dos subsistemas pode gerar mudança nos outros. O comportamento de um todo emerge a partir das interações dos subsistemas e não de um subsistema somente, conforme já mencionado anteriormente.

Larsen-Freeman (1997) não restringe a complexidade apenas para o estudo da L1. Ao analisar o desenvolvimento de línguas adicionais, a autora também assegura que o processo empregado é complexo, uma vez que diversos fatores envolvidos estão em interação entre si e são fundamentais para o percurso do desenvolvimento da interlíngua: a língua materna, a língua adicional (LA), a quantidade e tipo de input, a quantidade e tipo de interação, a quantidade e tipo de *feedback* recebido, etc. Além disso, para que o processo de aquisição de língua adicional ocorra satisfatoriamente, há fatores que precisam interagir entre si, tais como idade; aptidão; fatores sócio psicológicos como motivação e atitude; fatores relacionados à personalidade; estilo cognitivo, dentre outros. A autora acrescenta que um fator apenas pode não exercer um impacto no processo de aquisição/aprendizagem da língua adicional, mas a interação entre eles pode vir a ter impacto em tal desenvolvimento (LARSEN-FREEMAN, 1997).

Larsen-Freeman (2011) assegura que a complexidade dos sistemas é emergente: não emerge a partir de um elemento ou agente somente, mas a partir de suas interações. Além disso, a autora complementa que o termo ‘complexo’ não significa “complicado”, pois o termo denota um comportamento que prospera da interação de seus constituintes. Ainda quanto à complexidade, a autora relaciona o construto à criatividade: a construção de língua não é meramente uma repetição, mas implica inovação. Larsen-Freeman e Cameron (2008), em referência ao desenvolvimento de línguas adicionais, afirmam que o aprendizado dessas não abrange apenas a aquisição de estruturas linguísticas pelos aprendizes da mesma, mas uma adaptação contínua e realização de padrões de uso da língua a fim de construir significado, para que uma resposta emergja em uma situação comunicativa e dinâmica. Esta perspectiva assegura que o processo de aprendizagem da língua não significa aprender símbolos abstratos, mas endossa experiências de vida real que requerem uma coadaptação entre dois ou mais interlocutores em interação.

A interação entre os fatores descrita acima não se restringe apenas aos sistemas complexos e dinâmicos, conforme já mencionado. Tal interação também apresenta relevância ao se definir a língua como sistema adaptativo. O comportamento dos falantes parece ser construído com base em suas interações anteriores, enquanto que interações presentes e passadas, juntas, formam o comportamento futuro. O comportamento do falante é a consequência de fatores concorrentes entre restrições perceptuais a motivações sociais. As estruturas de uma língua emergem a partir dos padrões inter-relacionados de experiência, interação social e mecanismos cognitivos (BECKNER *et al.*, 2009). Adicionalmente, os autores apontam que a abordagem de Sistemas Complexos Adaptativos é encontrada em diversas áreas de pesquisa linguística como, por exemplo, aquisição de L1 e LA, história da linguística, evolução linguística, dentre outras.

Assim, a língua, à luz desta perspectiva, é inserida na ciência da complexidade devido aos múltiplos agentes interagindo entre e dentre si, fazendo com que um tenha efeito sobre o outro, bem como que interações também exerçam efeito sobre os outros agentes. A língua é vista como adaptativa, uma vez que tais interações formam o comportamento e as estruturas linguísticas do falante que emergem não de um componente, mas a partir dessas interações.

2.1.1.3 A língua é não-linear

Em L1 ou LA, isto é, L2, L3, L4, e assim por diante, os componentes são contingentes uns com os outros, além de estarem em constante interação, conforme já mencionado. Assim, conforme já afirmado, mudanças em uma parte do sistema poderão causar mudanças em outras partes do sistema, não sendo possível prever completamente o efeito causado. Nesse sentido, sistemas complexos são descritos como não-lineares. Larsen-Freeman (1997) afirma que a identificação de todos os fatores separadamente não é possível de modo exato e, ainda que fosse plausível tal identificação, a previsão do resultado da combinação deles não é possível de ser alcançada.

O aprendizado de língua adicional, de acordo com a autora, deve ser discutido à Luz da Teoria da Complexidade, a qual rejeita a perspectiva simplista de causa e efeito de aprendizagem linguística e sugere uma natureza não-linear e dinâmica, que pode variar entre indivíduos e vários fatores contextuais. Larsen-Freeman (2011), ao enumerar diversas características do CAS, sugere que os sistemas se adaptam tanto através de interação com o ambiente (fatores externos) como através de auto-organização (fatores internos), e a força das

interações se modifica com o percurso do tempo. Assim, múltiplas rotas são possíveis entre os componentes, mediadas por diferentes maneiras. Dessa forma, não há previsibilidade na trajetória dinâmica. Já que tais sistemas são abertos, isto é, qualquer elemento ou agente externo pode alterar o percurso, uma relação não-linear emerge, e uma ocorrência imprevista pode ocorrer a qualquer instante. Finalmente, a não-linearidade do CAS explica-se, também, através da ocasional variação caótica, que ocorre nestes sistemas. Sistemas complexos revisitam o mesmo território múltiplas vezes, o que significa que o nível presente do desenvolvimento depende criticamente daquele que o antecedeu.

Percebe-se que a noção de interação está sempre presente em sistemas complexos, tanto entre os fatores internos, quanto entre fatores externos. Os resultados, apesar de imprevisíveis, originam-se não de um fator, mas da interação entre eles. O caos é outro componente desses sistemas, uma vez que tais sistemas não operam em condições de equilíbrio. De acordo com Larsen-Freeman (1997), sistemas não-lineares complexos comportam-se regularmente, atingem um ponto crítico, tornam-se caóticos e, após isso, retornam ao seu estado regular novamente.

Para concluir, é importante ressaltar que as abordagens mais recentes (LARSEN-FREEMAN, 1997, 2011; ELLIS; LARSEN-FREEMAN, 2006; DE BOT; LOWIE; VERSPOOR, 2007; ELLIS, 2007, 2011; BECKNER *et al.*, 2009; DE BOT *et al.*, 2013) sugerem que tanto o desenvolvimento de L1 quanto o de uma LA sejam considerados sistemas complexos, dinâmicos e adaptativos, uma vez que o processo de aquisição destes dois tipos de sistemas envolve características específicas como a interação entre e dentre agentes, o que causa mudança (*status dinâmico*) entre os múltiplos agentes envolvidos no processo (*status complexo*). A partir de tal interação, esses sistemas se reorganizam, tornando-se **adaptativos**. Os autores sugerem, também, que o desenvolvimento do processo de aprendizagem da LA seja visto como **aberto**, uma vez que este recebe energia do exterior, e **não-linear**, já que não seguem uma rota linear de causa e efeito.

Pode-se esperar, a partir das premissas acima expressas, que não somente o sistema de língua materna exerça efeitos sobre o sistema de L2 em desenvolvimento, mas que o próprio sistema que está se desenvolvendo venha a agir, também, sobre a língua materna, uma vez que ambos os sistemas se encontram em constante evolução. A partir dessa premissa, propõe-se a seguinte investigação, cujos aspectos metodológicos são descritos no que segue.

3. O Estudo

3.1 Objetivo e perguntas de pesquisa

O presente estudo teve como objetivo geral investigar a produção vocálica de falantes nativos do espanhol rio-platense aprendizes de português porto-alegrense, a fim de se comparar a realização das vogais desses falantes com a realização das vogais dos falantes de espanhol rio-platense residentes na Argentina e monolíngues. Através da comparação desses espaços vocálicos, foi possível averiguar não somente a ação da L1 sobre o sistema da L2, mas, também, as influências que a língua não dominante, o português, pôde exercer na língua materna, o espanhol. Além disso, almejou-se responder, à luz da TSD e da Teoria da Complexidade, tomando por pressuposto que ambos os repertórios fonético-fonológicos estão juntos em um mesmo espaço acústico, as seguintes perguntas de pesquisa:

1. Os aprendizes residentes no Brasil (L1: espanhol rio-platense) já conseguem estabelecer a distinção entre as vogais médias baixas e médias altas do PB? Há diferenças referentes às produções das vogais /e/ e /o/ na L1 e na L2 desses aprendizes?
2. Há diferenças, comparando-se as produções, em espanhol, desses aprendizes com a de monolíngues? Os resultados podem sugerir que o desenvolvimento da L2 pode implicar mudanças no próprio sistema de L1 dos aprendizes?

Ao se contemplar a primeira questão, espera-se prover insumos para a discussão dos efeitos do sistema de L1 sobre o sistema de L2 em desenvolvimento dos aprendizes. Por sua vez, o cumprimento dos objetivos referentes à segunda questão possibilitará a discussão acerca da ação do sistema de L2 sobre a própria L1 dos bilíngues. As respostas para os dois questionamentos, dessa forma, possibilitarão a discussão sobre a bidirecionalidade da transferência vocálica.

3.2 Os participantes

Os dados para a realização deste estudo foram coletados de dois grupos distintos: o primeiro grupo contou com 5 monolíngues nativos do espanhol rio-platense, gravados pelo segundo autor deste estudo, na região da cidade de Mar del Plata, no ano de 2014; o segundo grupo contou com 5 bilíngues nativos da mesma variedade de espanhol supracitada, falantes de português e residentes em Porto Alegre, RS, por mais de 3 anos, gravados pela primeira autora, no ano de 2015. Do grupo monolíngue, 3 eram do sexo masculino e 2 do sexo feminino; por

sua vez, no grupo bilíngue, 2 participantes eram do sexo masculino e 3, do feminino. Todos os sujeitos participaram de forma voluntária da pesquisa e assinaram um Termo de Consentimento Livre e Esclarecido, cuja cópia foi fornecida aos participantes³.

3.3 Instrumento e procedimentos de coleta de dados

Com a finalidade de realizar a presente pesquisa, foram formulados dois instrumentos de leitura, um em espanhol – lido pelos participantes de ambos os grupos-, e outro em português – lido apenas pelos participantes do grupo bilíngue. Ambos foram formados por palavras inseridas em frases-veículos (Diga “palavra-alvo”), apresentadas em slides, como “diga tosse” e “diga pipoca”, em português, e “diga coche” e “diga machaca”, em espanhol. Os contextos antecedentes e seguintes da vogal tônica eram surdos, de modo a não causar um alongamento da referida vogal devido ao vozeamento da consoante precedente ou seguinte (cf. ALVES, 2015). Para a elaboração do instrumento, foram escolhidas apenas palavras não-cognatas, a fim de não ativar a língua materna do participante e, quanto ao número de sílabas, foram escolhidas 3 palavras dissilábicas e 3 palavras trissilábicas, de padrão acentual paroxítono, para cada vogal. A utilização de palavras tanto dissilábicas quanto trissilábicas se deve à limitação, no léxico, para se conseguir um número de 6 palavras, para cada vogal, que apresentassem índices de frequência próximos. Dessa forma, contou-se com um número igual de palavras dissilábicas e trissilábicas, dado que o número distinto de sílabas implica cálculos de duração relativa separados, em função do *status* bi ou trissilábico da palavra-alvo. A lista de palavras do espanhol contou com 105 palavras, das quais 90 eram palavras-alvo e 15, distratoras; por sua vez, a lista de palavras do português contou com 145 palavras, das quais 126 eram palavras-alvo e 19, distratoras. Assim, cada instrumento foi composto por 6 *types* por vogal, com 3 repetições, totalizando 18 *tokens* por vogal para cada participante.

O *software* utilizado para a gravação de áudio dos participantes foi o *Audacity*⁴, e a gravação foi realizada em um computador *Sony Vaio*, com um microfone adaptado *Sony*, modelo ECM x1. Após a gravação em espanhol, houve uma pausa de aproximadamente 5 minutos para que os informantes pudessem descansar e tomar água. Além disso, os participantes bilíngues da pesquisa também foram solicitados a preencher um questionário sobre seu histórico de aquisição da segunda língua; por sua vez, os membros do grupo de monolíngues

³ Esta pesquisa foi aprovada pelo Comitê de Ética da Instituição dos autores (Parecer no. 1.144.074).

⁴ *Software* livre. Download gratuito em www.audacity.sourceforge.net.

também foram solicitados a preencherem uma Ficha de Informações. As informações fornecidas em tal ficha confirmariam que os participantes considerados como monolíngues não tiveram contato com qualquer língua estrangeira, em termos instrucionais, nem tampouco se mostram capazes de falar ou entender qualquer outro idioma que não a sua L1. Todos os participantes assinaram um termo de consentimento livre e esclarecido, conforme já mencionado na seção 2.2.

3.4 Análises acústica e estatística dos dados

Quanto à análise acústica, foi utilizado o *software* Praat -versão 5.4.08 (BOERSMA & WEENINK, 2015) para a medição dos formantes F1 e F2 e da duração absoluta e a obtenção dos valores de duração relativa das vogais. Para a obtenção dos valores de duração relativa, procedeu-se à medição da duração de toda a frase-veículo e, após isso, os dados de duração da vogal e da frase foram inseridos no cálculo obtido através da duração absoluta do evento acústico multiplicado por 100, dividido pelo valor da duração absoluta da frase, para que se obtivesse, como resultado final, o percentual ocupado pela vogal no interior da frase-veículo. Segundo Silva (2014, p. 15), “esta normalização é indicada para evitar que fatores prosódicos, como velocidade de fala, possam enviesar os resultados”.

Para a normalização dos valores de F1 e F2, utilizou-se o método de normalização Lobanov, através do site de normalização “Norm” (<http://lvc.uoregon.edu/norm>), que permitiu, também, as plotagens desses valores nos gráficos a serem apresentados neste trabalho. A partir dos dados normalizados de F1 e F2, bem como dos valores de duração absoluta e relativa das vogais, a análise estatística foi realizada através do pacote de estatística SPSS (versão 16).

4. Resultados

A presente seção será organizada em dois diferentes momentos. Primeiramente, serão descritos e discutidos os dados referentes aos valores de F1 e F2 das produções dos participantes bilíngues e monolíngues. Em um segundo momento, serão apresentados e discutidos os valores de duração absoluta e relativa encontradas nas produções de ambos os grupos. Em ambos os momentos, procurar-se-á contemplar as duas questões de investigação, apresentadas na seção anterior.

4.1 Formantes

O gráfico a seguir mostra a organização do espaçamento vocálico do português (L2) por falantes nativos do espanhol, e a Tabela 1 apresenta as Médias, os Desvios-Padrão (DP) e os Coeficientes de Variação (CV) de F1 e F2 (em Hertz) de cada vogal do português do grupo bilíngue após a normalização.

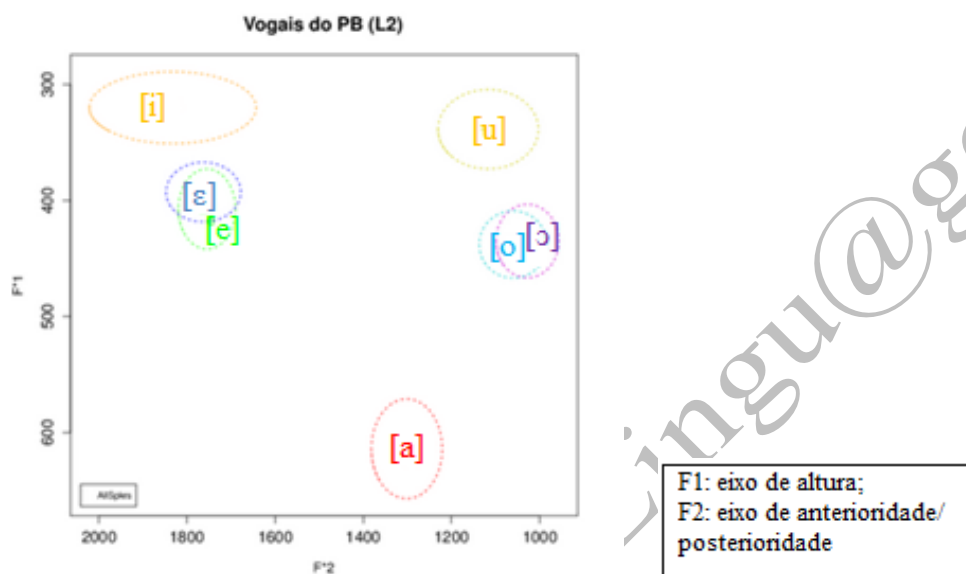


Figura 1: Espaço vocálico do português (L2) produzido pelos bilíngues.

Tabela 1: Médias de F1 e F2 (em Hz), Desvios-Padrão (DP) e Coeficientes de Variação (CV) do português (L2) após normalização.

	[a]	[ε]	[e]	[i]	[ɔ]	[o]	[u]
F1	614,1 (DP=27,27) (CV=4,44%)	392,7 (DP=13,77) (CV=3,50%)	407,35 (DP=24,71) (CV=6,06%)	319,9 (DP=19,41) (CV=6,06%)	434,94 (DP=22,68) (CV=5,21%)	437,9 (DP=22,72) (CV=5,18%)	338,5 (DP=26,99) (CV=7,97%)
F2	1301,3 (DP=23,32) (CV=1,79%)	1762,6 (DP=39,37) (CV=2,23%)	1754,6 (DP=17,72) (CV=1,00%)	1832,1 (DP=63,93) (CV=3,48%)	1028,01 (DP=31,61) (CV=3,07%)	1060,4 (DP=40,65) (CV=3,83%)	1116,8 (DP=60,90) (CV=5,45%)

Através de uma simples visualização da Fig. 1, é possível constatar que a formação das novas categorias vocálicas que correspondam às vogais médias-baixas [ε] e [ɔ] não ocorre, uma vez que as produções das vogais médias-baixas pelos bilíngues se sobrepõem às categorias das vogais médias-altas, que já ocorrem na L1 destes falantes. O que também é possível de se observar na Fig. 1 é um espaço acústico “vazio”, justamente onde estariam as vogais médias-baixas; assim, esses bilíngues podem estar abrindo uma “brecha”, no seu espaço acústico, para a posterior formação destas novas categorias.

A fim de responder à primeira pergunta de pesquisa deste trabalho, que indagava se os estrangeiros residentes no Brasil se mostravam capazes de realizar a distinção entre as vogais médias-altas e médias-baixas, foram realizados uma série de testes- t^5 para a verificação das diferenças produzidas pelos bilíngues em termos de altura (F1) e anterioridade/posterioridade (F2) destas vogais. Os resultados dos testes- t negritados indicam uma diferença estatística significativa.

Tabela 2: Resultados dos testes- t pareados: comparações intra-grupo (bilíngues) entre as vogais médias do PB.

	[ɛ] vs [e]	[ɔ] vs [o]
F1	T (4) = 1,894; p=0,131	T (4) = -1,079; p= 0,341
F2	T (4) = -0,478; p= 0,658	T (4) = -2,263; p= 0,086

Os testes- t para amostras emparelhadas (*Paired-Sample T-Test*) não revelaram diferenças significativas entre os membros dos pares de vogais, tendo apontado apenas uma diferença marginalmente significativa em F2 para as vogais [o] e [ɔ]. No que diz respeito ao eixo da altura, conforme se vê na Fig. 1, é preciso ressaltar que a vogal [ɛ] é mais alta, em termos descritivos, do que a vogal [e], ainda que não tenha havido uma diferença significativa entre os valores de F1 de [e] e [ɛ]. Isso sugere que os aprendizes ainda não adquiriram as diferenças entre médias baixas e médias altas, e, além disso, podem estar usando as vogais intercambiavelmente, de modo a produzir, por exemplo, uma palavra como “peso” (verbo) com uma altura mais alta do que em “peso” (substantivo).

O gráfico a seguir apresenta a organização do espaçamento das vogais do espanhol (L1) produzidas pelos bilíngues.

⁵ Para a realização de testes paramétricos, foram aplicados os testes Kolmogorov-Smirnov e Shapiro-Wilk, de verificação de normalidade das amostras. Esses testes apontaram, na maior parte dos casos, uma distribuição normal dos dados. Seguindo-se Fife-Schaw (2006) e Matos *et al.* (2009), nos casos de violação dos pressupostos de utilização de testes paramétricos, foram realizados tanto testes paramétricos quanto não-paramétricos. Seguindo-se, também, os autores supracitado, uma vez que não houve diferenças referentes às conclusões de significância por ambos os tipos de testes, optou-se por reportar os resultados dos testes paramétricos, em função de sua maior robustez.

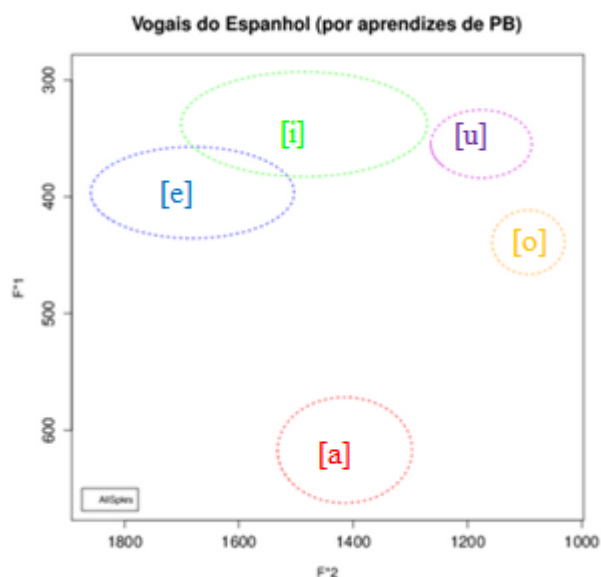


Figura 2: Espaço vocálico do espanhol (L1) produzido pelos bilíngues.

Através da visualização na Fig. 2, é possível constatar a mesma elevação das vogais médias-altas [e] e [o] que esses bilíngues também realizaram no espaço acústico das vogais do português, de modo a haverem realizado um espaço maior entre a vogal baixa [a] e o restante das vogais. Esta elevação das vogais médias parece surpreendente, especialmente se for considerado o fato de que, uma vez que o espanhol apresenta apenas vogais médias-altas, o espaço acústico a ser ocupado por esta categoria fonológica poderia ser mais amplo do que aquele ocupado pelas vogais médias-altas do português, língua essa que apresenta um espaço acústico médio que deve ser dividido entre médias-altas e médias-baixas. Maiores detalhes sobre esta elevação das médias-altas na L1 dos bilíngues serão discutidos a seguir, quando forem comparadas tais produções às vogais do espanhol realizadas por monolíngues. De qualquer forma, a observação da Fig. 2 já permite sugerir que os falantes bilíngues elevaram suas categorias nativas, assim como fizeram na L2.

A Tabela 3 apresenta as Médias de F1 e F2 (em Hertz), o Desvio-Padrão (DP) e o Coeficiente de Variação (CV) após normalização das vogais do espanhol (L1) e do português (L2), pelos falantes bilíngues.

Tabela 3: Médias de F1 e F2 (em Hz), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das vogais dos bilíngues, em espanhol e em português.

Espanhol	[a]	[e]	[i]	[o]	[u]
F1	617,1 (DP=9,73) (CV=1,57%)	396,4 (DP=24,65) (CV=6,21%)	337,9 (DP=19,65) (CV=5,81%)	438,9 (DP=10,84) (CV=2,46%)	354,8 (DP=9,74) (CV=2,74%)
F2	1414,1 (DP=82,98) (CV=5,86%)	1681,2 (DP=75,10) (CV=4,46%)	1485,9 (DP=164,6) (CV=11,07%)	1093,6 (DP=33,00) (CV=3,01%)	1175,9 (DP=39,92) (CV=3,39%)
Português					
F1	614,1 (DP=27,27) (CV=4,44%)	407,35 (DP=24,71) (CV=6,06%)	319,9 (DP=19,41) (CV=6,06%)	437,9 (DP=22,72) (CV=5,18%)	338,5 (DP=26,99) (CV=7,97%)
F2	1301,3 (DP=23,32) (CV=1,79%)	1754,6 (DP=17,72) (CV=1,00%)	1832,1 (DP=63,93) (CV=3,48%)	1060,4 (DP=40,65) (CV=3,83%)	1116,8 (DP=60,90) (CV=5,45%)

Também com vistas a atender à primeira questão norteadora, testes-*t* para amostras pareadas (intra-grupo) foram conduzidos para verificar se as vogais [e] e [o] são distinguidas pelos participantes bilíngues em português e em espanhol em termos de altura (F1) e anterioridade/posterioridade (F2).

Tabela 4: Resultados dos testes-*t* pareados: comparação intra-grupo (bilíngues) entre as vogais médias-altas /e/ e /o/ do português e do espanhol.

	[e]	[o]
F1	T (4) = 0,664; p=0,543	T (4) = -0,092; p=0,931
F2	T (4) = -2,513; p=0,066	T (4) = -2,473; p=0,069

A tabela acima mostra que não há diferença em relação à altura na produção de [e] destes bilíngues, em português e em espanhol, tendo havido apenas uma diferença marginalmente significativa no que diz respeito ao eixo anteroposterior. Em termos de estatística descritiva, é necessário apontar que, no que diz respeito à altura, a vogal [e] do português se mostrou mais baixa do que a do espanhol. Este fato é, também, surpreendente, visto que, pelo fato de o português apresentar duas vogais médias que se diferenciam, esperava-se que a média-alta da L2 fosse mais alta do que a única média em espanhol, que teria mais liberdade de ocupar um

espaço acústico maior em termos de altura. Tem-se, assim, mais um fato que nos leva a sugerir que a vogal [e], em espanhol (L1), parece ter sido elevada, no sistema destes bilíngues. Já no que diz respeito à produção da vogal [o], de modo semelhante ao verificado com [e], apenas a diferença no eixo anteroposterior resultou marginalmente significativa. De fato, conforme a descrição realizada na Tabela 3, as médias de F1 de [o], em português e em espanhol, são praticamente iguais (437,9 e 438,9, respectivamente). Conforme já dito, esperava-se que a média em espanhol fosse mais baixa, uma vez que há apenas um fonema vocálico posterior médio em espanhol.

A seguir, passa-se a contemplar a segunda questão de investigação proposta para este estudo. No que segue, serão apresentadas as comparações entre a produção vocálica em espanhol dos bilíngues com a produção vocálica dos falantes monolíngues. A Fig. 3 mostra a organização do espaçamento vocálico dos falantes monolíngues residentes na Argentina.



Figura 3: Espaço vocálico do espanhol (L1) produzido pelos monolíngues.

Por meio de uma simples visualização da Fig. 3, já se faz possível verificar uma melhor distribuição, ao longo do espaço acústico, das cinco vogais do espanhol. Percebe-se que as vogais médias altas estão ocupando um espaço mais centralizado do trato, em termos de altura, apesar de a vogal anterior ocupar uma posição mais elevada do que a vogal posterior.

Para fins de comparação, na Fig. 4, são apresentados ambos os espaçamentos das vogais do espanhol, realizadas pelos bilíngues e monolíngues.

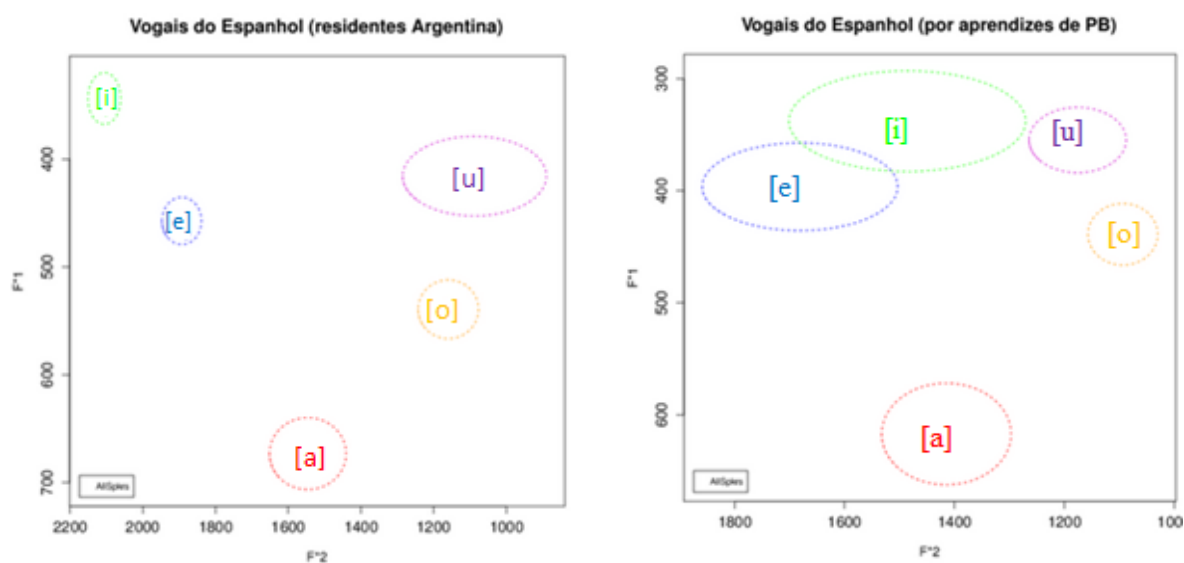


Figura 4: Espaços vocálicos do espanhol (L1) produzidos pelos monolíngues e bilíngues.

É possível perceber, na Fig. 4, que as vogais dos monolíngues se mostram mais bem distribuídas ao longo do espaço acústico. As dos bilíngues encontram-se, sobretudo, em posições mais altas. Ainda que não estejam sendo feitas, em português (L2), as distinções entre médias baixas e altas, não se pode descartar a possibilidade de que estes valores mais altos, tanto na L1 quanto na L2 dos bilíngues, correspondam a uma espécie de “acomodação”, a partir da qual, conforme já mencionado, os aprendizes estejam “abrindo espaço” para a possibilidade de vogais ocupando um espaço mais baixo (justamente, as médias-baixas). Essa “acomodação” na L2 pode estar, também, exercendo efeitos no sistema de L1. Chama a atenção, também nos dados em L1 dos bilíngues, a centralização do [i]. Tal centralização também parece estar relacionada à elevação da média anterior: sendo elevada a vogal [e], o que poderia ocasionar, inclusive, uma pressuposição entre a vogal média e a alta, [i] acaba assumindo uma posição mais centralizada, para evitar a sobreposição entre categorias. Essa modificação em [e] que acaba exercendo efeitos, também, em [i], mostra-se em consonância com a teoria de base que rege o presente estudo, visto que subentende que uma modificação em parte do sistema pode vir a exercer efeitos em todo o restante do sistema.

A Tabela 5, em consonância com os valores apresentados na figura acima, apresenta as Médias de F1 e F2 (em Hertz), Desvio-Padrão (DP) e Coeficientes de Variação (CV) das vogais de monolíngues e bilíngues, após a normalização.

Tabela 5: Médias de F1 e F2 (em Hz), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das vogais do espanhol (L1) para cada grupo após normalização.

Grupo	Formantes	[a]	[e]	[i]	[o]	[u]
Monolíngues	F1	673,1 (DP=11,35) (CV=1,68%)	457,2 (DP=18,02) (CV=3,94%)	343,6 (DP=14,81) (CV=4,31%)	539,2 (DP=19,44) (CV=3,60%)	415,7 (DP=19,62) (CV=4,71%)
	F2	1546,0 (DP=30,11) (CV=1,97%)	1892,8 (DP=14,70) (CV=0,77%)	2105,8 (DP=24,05) (CV=1,14%)	1159,5 (DP=31,75) (CV=2,73%)	1087,4 (DP=13,63) (CV=1,25%)
Bilíngues	F1	617,1 (DP=9,73) (CV=1,57%)	396,4 (DP=24,65) (CV=6,21%)	337,9 (DP=19,65) (CV=5,81%)	438,9 (DP=10,84) (CV=2,46%)	354,8 (DP=9,74) (CV=2,74%)
	F2	1414,1 (DP=82,98) (CV=5,86%)	1681,2 (DP=75,10) (CV=4,46%)	1485,9 (DP=164,6) (CV=11,07%)	1093,6 (DP=33,00) (CV=3,01%)	1175,9 (DP=39,92) (CV=3,39%)

A partir destes dados, pode-se ver que, em média, as produções das vogais [e] e [o] se apresentam bastante distintas entre os grupos em termos de altura e em termos de anterioridade/posterioridade. Os bilíngues apresentam uma produção mais elevada destas vogais, com valores de F1 mais baixos. A comparação dos dados da Tabela 5, bem como das plotagens, permite ver que as vogais médias dos bilíngues tendem a ser mais altas do que as dos monolíngues.

Com o objetivo de responder à segunda questão norteadora, foram realizados testes-*t* para amostras independentes (entre grupos), visando a verificar possíveis diferenças estatísticas entre as vogais produzidas por ambos os grupos.

Tabela 6: Resultados dos testes-*t* para Amostras Independentes (entre grupos): comparação entre as vogais do espanhol (L1) de monolíngues e bilíngues.

Formantes	/a/	/e/	/i/	/o/	/u/
F1	t(8)= -8,384; p=0,000	t(8)= -4,452; p=0,002	t(8)= -0,514; p=0,621	t(8)=-10,079; p=0,000	t(8)= -6,216; p=0,000
F2	t(8)=-3,342; p=0,010	t(8)= -6,183; p=0,000	t(8)= -8,330; p=0,000	t(8)= -3,218; p=0,012	t(8)= 4,690; p=0,002

Todas as vogais, com exceção dos valores de F1 de [i], apresentaram uma diferença significativa, nas produções de monolíngues e bilíngues, tanto em função do eixo de altura (F1), quanto no eixo anteroposterior. Tal fato sugere uma alteração no sistema vocálico de L1 dos bilíngues, possivelmente devido à influência de outro sistema linguístico, o português (L2), visto que suas realizações vocálicas se diferem das realizações dos monolíngues. Conforme já esclarecido anteriormente, o fato de haverem sido encontradas diferenças significativas mostra-se, também, em consonância com a noção de Língua como Sistema Adaptativo Complexo. Além das alterações nas vogais médias, discutidas ao longo de toda a seção, é importante, ainda neste aspecto, ressaltar os valores distintos de [i] (mais centralizado) e de [a] (com um maior desvio padrão, de modo a ocupar um espaço maior, que havia sido deixado “vazio”, no espaço acústico, após o alçamento das médias), que confirmam, também, a complexidade do sistema.

4.2 Duração

Ao se iniciar a descrição dos valores absolutos e relativos das durações dos segmentos vocálicos, serão contempladas as duas questões de pesquisa já abordadas na descrição das frequências formânticas, uma vez que os valores de duração poderão fornecer, também, insumos para a discussão acerca da bidirecionalidade da transferência linguística. A seção será organizada em dois momentos: primeiramente, serão discutidos os valores absolutos de duração; em um segundo momento, serão apresentados os dados referentes às durações relativas. Em ambas as subseções, as duas questões de pesquisa previamente apresentadas serão abordadas.

4.2.1 Valores absolutos de duração⁶

Inicia-se, pois, pela discussão referente à primeira questão norteadora, que indagava se haveria diferenças entre as produções em espanhol (L1) e em português (L2), por parte dos informantes bilíngues. A seguir, a Tabela 7 apresenta as Médias, Desvio-Padrão (DP) e Coeficiente de Variação (CV) da duração das vogais pelos bilíngues (em milissegundos).

⁶ Ainda que se reconheça o fato de que a literatura aponta os valores relativos de duração como mais expressivos, tendo em vista que minimizam efeitos de velocidade de fala, optou-se, neste trabalho, por reportar tanto os valores absolutos quanto relativos, pelo fato de ambas as medidas terem claramente evidenciado a noção de transferência bidirecional que rege o presente trabalho, conforme será visto a partir da presente seção.

Tabela 7: Médias (em milissegundos), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das durações absolutas das vogais produzidas pelos bilíngues, em ambas as línguas.

	[a]	[ɛ]	[e]	[i]	[ɔ]	[o]	[u]
Espanhol (L1)	115,8 (DP=25,05) (CV=21,6%)	-----	114,74 (DP=25,73) (CV=22,4%)	88,7 (DP=14,15) (CV=15,94%)	-----	121,7 (DP=22,76) (CV=18,7%)	107,5 (DP=20,36) (CV=18,9%)
Português (L2)	141,5 (DP=40,3) (CV=28,45%)	129,03 (DP=37,7) (CV=29,2%)	139,3 (DP=37,45) (CV=26,75%)	105,04 (DP=20,6) (CV=19,6%)	136,9 (DP=32,61) (CV=23,8%)	131,8 (DP=36,2) (CV=27,44%)	112,3 (DP=30,6) (CV=27,24%)

Para todas as vogais, a duração dos segmentos em português é mais longa do que em espanhol. Isso vai ao encontro do que afirma Santos (2014), que atesta que as vogais do PB são mais longas do que as do espanhol rio-platense. Dessa forma, com base no afirmado acima, pode-se verificar que os aprendizes residentes no Brasil já apresentam valores de duração que tendem a se aproximar aos da L2, em termos de duração vocálica. Há de se ressaltar, também, que os valores de Desvio-Padrão (DP) e de Coeficiente de Variação (CV) das cinco vogais se mostram mais altos em português do que em espanhol, de modo a demonstrar mais variabilidade na língua-alvo.

Testes-*t* para amostras emparelhadas (*Paired-sample T-tests*) foram realizados a fim de verificar se havia diferenças significativas entre os valores em espanhol e português. Tais resultados são mostrados na Tabela 8, a seguir.

Tabela 8: Resultados dos testes-*t* para Amostras Emparelhadas (intra-grupo), referentes às diferenças de duração absoluta das vogais, produzidas pelo grupo bilíngue, em português e em espanhol.

[a]	[ɛ]	[i]	[o]	[u]
t(4)=-2,628; p=0,058	t(4)=-2,564; p=0,062	t(4)=-2,784; p=0,050	t(4)=-1,292; p=0,266	t(4)=-0,730; p=0,506

Os resultados da Tabela acima indicam que as diferenças de duração são marginalmente significativas (exceto para as vogais posteriores [o] e [u], que não mostraram significância estatística) quando ambos os sistemas são comparados para cada uma das vogais.

Passa-se, agora, à discussão referente à segunda questão norteadora, que indagava se haveria diferenças, comparando-se as produções, em espanhol (L1), dos falantes bilíngues com a de monolíngues, verificação essa que possibilitaria a discussão acerca da possibilidade de o desenvolvimento linguístico em L2 implicar mudanças no próprio sistema de L1 dos aprendizes. A seguir, são apresentadas, nas Tabelas 9 e 10, as comparações (em termos

descritivos e inferenciais, respectivamente) entre as produções em espanhol de monolíngues e bilíngues.

Tabela 9: Médias (em milissegundos), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das durações absolutas das vogais produzidas pelos monolíngues e bilíngues, em espanhol (L1).

Grupo	[a]	[e]	[i]	[o]	[u]
Bilíngues	115,8 (DP=25,05) (CV=21,6%)	114,74 (DP=25,73) (CV=22,4%)	88,7 (DP=14,15) (CV=15,94%)	121,7 (DP=22,76) (CV=18,7%)	107,5 (DP=20,36) (CV=18,9%)
Monolíngues	86,74 (DP=6,05) (CV=6,97%)	78,03 (DP=6,06) (CV=7,76%)	70,49 (DP=6,17) (CV=8,75%)	89,4 (DP=5,70) (CV=6,37%)	85,82 (DP=5,33) (CV=6,21%)

As durações, bem como os Desvios-Padrão e os Coeficientes de Variação, das vogais dos bilíngues são maiores do que as dos monolíngues. Diante de tal resultado, foram realizados testes-*t* para Amostras Independentes, a fim de investigar a relevância estatística destas diferenças.

Tabela 10: Resultados dos testes-*t* para Amostras Independentes (entre grupos), referentes às diferenças da duração absoluta das vogais produzidas por monolíngues e bilíngues, em espanhol (L1).

[a]	[e]	[i]	[o]	[u]
t(8)=2,523; p=0,036	t(8)=3,104; p=0,031	t(8)=2,643; p=0,030	t(8)=-3,079; p=0,015	t(8)=-2,298; p=0,050

Os resultados da Tabela 10 indicam que as diferenças de duração de todas as vogais se mostraram estatisticamente significativas nas comparações das produções entre monolíngues e bilíngues, o que pode indicar uma alteração na duração das vogais do espanhol dos bilíngues devido à influência do sistema de língua portuguesa. Tal achado vai ao encontro da premissa dinâmica, que prevê a emergência de alterações, no próprio sistema de L1, devido a qualquer interação com agentes adicionais, que neste caso, é o próprio português. A perspectiva de língua como CAS também prevê que a direção da influência linguística não é unilateral, mas bidirecional, como pode ser observado com o alongamento das vogais no espanhol pelos falantes bilíngues do português, idioma esse que, conforme Santos (2014), dispõe de vogais mais longas do que o espanhol.

4.2.2 Valores relativos de duração

A fim de garantir que as diferenças significativas encontradas na seção anterior não se deviam, unicamente, a possíveis diferenças da taxa de elocução da leitura das frases-veículo em cada idioma, procedeu-se à verificação, também, dos valores de duração relativa das vogais-alvo, com relação à duração da frase veículo em que se encontravam inseridas. Uma vez que o instrumento contava tanto com palavras dissilábicas quanto com itens lexicais trissilábicos, o que poderia, de certa forma, diminuir a duração relativa das vogais em frases cujas palavras-alvo continham três sílabas, realizou-se a organização e descrição dos dados separadamente, considerando-se o número de sílabas da palavra-alvo. Procurou-se, também, organizar a apresentação dos dados de acordo com o ordenamento das duas questões que motivam o estudo. Dessa forma, primeiramente, serão comparadas as durações relativas das vogais em português e espanhol dos bilíngues; em seguida, serão verificadas as diferenças nas durações relativas em língua espanhola (L1), por parte de monolíngues e bilíngues.

A seguir, serão apresentadas as Médias, os Desvios-Padrão (DP) e os Coeficientes de Variação (CV) da duração relativa das vogais dos bilíngues, em português e espanhol.

Tabela 11: Médias (em porcentagens), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das durações relativas das vogais produzidas pelos bilíngues, em ambas as línguas.

	[a]	[ɛ]	[e]	[i]	[ɔ]	[o]	[u]
Português Palavras dissilabas	16,23 (DP= 2,96) (CV=18,23%)	14,04 (DP=1,96) (CV=13,96%)	16,93 (DP=2,23) (CV=13,17%)	12,26 (DP=1,48) (CV=12,1%)	15,91 (DP=1,71) (CV=10,74%)	14,85 (DP=2,15) (CV=14,47%)	13,09 (DP=1,78) (CV=13,6%)
Espanhol Palavras dissilabas	13,37 (DP=1,49) (CV=11,14%)	—	12,16 (DP=1,21) (CV=9,95%)	10,09 (DP=1,24) (CV=12,3%)	—	14,03 (DP=1,22) (CV=8,69%)	13,19 (DP=1,35) (CV=10,23%)
Português Palavras trissilabas	14,00 (DP=2,31) (CV=16,5%)	13,42 DP=2,28 (CV=16,98%)	13,17 DP=2,23 (CV=16,93%)	9,87 DP=1,20 (CV=12,15%)	12,67 DP=1,61 (CV=12,7%)	12,20 DP=1,67 (CV=13,7%)	10,74 DP=1,86 (CV=17,31%)
Espanhol Palavras trissilabas	11,49 (DP=2,12) (CV=18,45%)	—	10,65 (DP=1,32) (CV=12,5%)	8,86 (DP=1,09) (CV=12,3%)	—	12,24 (DP=0,87) (CV=7,10%)	9,49 (DP=1,13) (CV=11,9%)

Os resultados da Tabela 11 acima mostram que, quando comparadas as durações relativas das vogais em palavras dissilabas nas duas línguas do participante bilíngue, em português, a duração é sempre maior do que as vogais em espanhol, como exceção da vogal [u], que apesar de apresentar durações bastante semelhantes em ambas as línguas, é produzida mais longamente em espanhol. Em referência às palavras de três sílabas, as durações relativas

das vogais do português são mais longas do que as do espanhol, quando produzidas pelos falantes bilíngues, com exceção da vogal [o], que é levemente mais longa em espanhol. Para a maioria das vogais, os valores de DP e CV foram mais altos, também, nas produções em língua portuguesa.

Uma série de testes-*t* para Amostras Emparelhadas (intra-grupo) foi conduzida para a verificação do contraste estatístico referente à duração relativa dessas vogais produzidas pelo falante bilíngue. Os resultados são apresentados a seguir. Os resultados em negrito indicam relevância estatística.

Tabela 12: Resultados dos testes-*t* para Amostras Emparelhadas (intra-grupo), referentes às diferenças da duração relativa das vogais, produzidas pelo grupo bilíngue, em português (L2) e em espanhol (L1).

Tipo de palavra	[a]	[e]	[i]	[o]	[u]
Dissílabas	t (4) = -2,783; p = 0,050	t (4) = -6,800; p = 0,002	t (4) = -3,812; p = 0,019	t (4) = -0,916; p = 0,412	t (4) = 0,145; p = 0,892
Trissílabas	t (4) = -3,829; p = 0,019	t (4) = -3,154; p = 0,034	t (4) = -1,788; p = 0,148	t (4) = -0,068; p = 0,949	t (4) = -2,080; p = 0,106

A Tabela 12 mostra que as durações relativas das vogais realizadas pelos bilíngues se distinguem de uma língua para a outra com as vogais [a, e], em ambos os tipos de palavras, dissílabas e trissílabas. Para a vogal [i], foi encontrada diferença significativa apenas nas dissilábicas. Já as vogais [o] e [u], conforme pode ser visualizado nos dados descritivos da Tabela 11, são produzidas com durações relativas estatisticamente semelhantes em ambos os sistemas, espanhol e português. Tal resultado indica que os falantes bilíngues estão adquirindo os padrões de duração das vogais do português e deixando de aplicar os padrões nativos na L2. Ao se compararem os resultados da Tabela 12 com os observados com durações absolutas (Tabela 08), verifica-se que, com exceção do observado na vogal [i] (que, com durações absolutas, apresentava diferenças significativas), resultados semelhantes aos já apontados na seção anterior são, aqui, apresentados. Verifica-se, assim, que as diferenças de duração tendem a prevalecer nas vogais mais anteriores, talvez pelo fato de o arredondamento implicar um fator que afeta uma maior duração das vogais da L2 dos bilíngues.

Ao se discutir a segunda questão norteadora, cabe verificar se as durações relativas das vogais do espanhol destes falantes são semelhantes ou distintas das durações relativas das vogais dos falantes monolíngues. A tabela 13, a seguir, apresenta os valores correspondentes.

Tabela 13: Médias (em porcentagens), Desvios-Padrão (DP) e Coeficientes de Variação (CV) das durações relativas das vogais produzidas pelos monolíngues e pelos bilíngues, em espanhol (L1).

Grupo	Tipo de palavra	[a]	[e]	[i]	[o]	[u]
Monolíngues	Dissílabas	11,06 (DP= 0,74) (CV=6,37%)	9,64 (DP=1,17) (CV=12,13%)	8,69 (DP=1,16) (CV=13,34%)	10,88 (DP=0,59) (CV=5,42%)	11,16 (DP=1,18) (CV=10,57%)
Bilíngue	Dissílabas	13,37 (DP=1,49) (CV=11,14%)	12,16 (DP=1,21) (CV=9,95%)	10,09 (DP=1,24) (CV=12,3%)	14,03 (DP=1,22) (CV=8,69%)	13,19 (DP=1,35) (CV=10,23%)
Monolíngue	Trissílabas	9,12 (DP=0,63) (CV=6,9%)	8,70 (DP=0,53) (CV=6,09%)	7,80 (DP=0,38) (CV=4,87%)	9,51 (DP=0,37) (CV=3,89%)	8,57 (DP=0,36) (CV=4,20%)
Bilíngue	Trissílabas	11,49 (DP=2,12) (CV=18,45%)	10,65 (DP=1,32) (CV=12,5%)	8,86 (DP=1,09) (CV=12,3%)	12,24 (DP=0,87) (CV=7,10%)	9,49 (DP=1,13) (CV=11,9%)

Conforme a Tabela 13 mostra, todas as vogais em espanhol produzidas pelos falantes bilíngues, tanto em palavras dissílabas quanto em palavras trissílabas, apresentam durações relativas descritivamente maiores do que as das vogais produzidas pelos falantes monolíngues, além de tenderem a valores de Desvio-Padrão e Coeficientes de Variação também mais longos. Tais resultados sugerem que estas durações maiores nas produções das vogais pelos falantes bilíngues, em sua própria L1, possa ser influência do português, sistema que dispõe de vogais mais longas do que a variedade do espanhol, segundo Santos (2014). Este resultado, semelhante ao já observado com os valores absolutos de duração, sugere que a influência vocálica não segue, unicamente, uma direção unilateral de L1 → L2, mas também pode se manifestar a partir de uma direção bidirecional: L2 → L1.

Com a finalidade de investigar se estas diferenças de durações apresentam relevância estatística, testes-*t* de Amostras Independentes (entre grupos), que compraram as durações relativas do espanhol dos bilíngues com as durações relativas do espanhol dos monolíngues, foram conduzidos. Os resultados são apresentados a seguir.

Tabela 14: Resultados dos testes-*t* para Amostras Independentes (entre grupos), referentes às diferenças de duração relativa das vogais produzidas pelos bilíngues e pelos monolíngues, em espanhol (L1).

Tipo de Palavra	[a]	[e]	[i]	[o]	[u]
Dissílabas	t (8) = 3,094; p = 0,015	t (8) = 3,337; p = 0,010	t (8) = 1,836; p = 0,104	t (8) = 5,181; p = 0,002	t (8) = 2,524; p = 0,036
Trissílabas	t (8) = 2,336; p = 0,048	t (8) = 2,917; p = 0,020	t (8) = 2,046; p = 0,075	t (8) = 6,455; p = 0,000	t (8) = 1,730; p = 0,122

Como pode ser observado através da Tabela 14, de modo semelhante ao que já havia sido apontado nas durações absolutas, todas as durações relativas (com exceção dos casos de [u], no caso das trissilábicas, e [i], nas dissilábicas) das vogais apresentaram contraste significativo quando comparadas às produções de ambos os grupos falantes de espanhol, bilíngues e monolíngues. Esse resultado indica que as durações relativas referentes às vogais do grupo bilíngue se distinguem estatisticamente e são maiores do que as referentes às vogais do grupo monolíngue, de modo a sugerir que o padrão referente à duração do português (L2) pode estar agindo sobre o próprio sistema de espanhol (L1) desses participantes.

5. Conclusão

Inicialmente, investigou-se se os valores formânticos referentes à altura e à anterioridade/posterioridade das vogais médias-baixas de falantes bilíngues aprendizes de português, a fim de verificar a formação de novas categorias vocálicas da L2 (cf. FLEGE, 1995) por estes participantes. Percebeu-se que estes falantes, ainda, não chegam a formar novas categorias no seu espaço articulatório. O que ocorre é uma acomodação no espaço vocálico, possibilitando uma abertura neste espaçamento para o surgimento das vogais médias-baixas, mas elas não chegam a ser produzidas de fato. Não há distinção das médias-baixas com as vogais já produzidas no dialeto materno (médias-altas), não havendo, portanto, formação de novas categorias⁷.

⁷ Em estudos futuros, procurar-se-á comparar as produções de português (L2) destes participantes argentinos com a de monolíngues brasileiros, falantes do dialeto porto-alegrense do português. Através de comparações, será possível, também, estabelecer as distâncias euclidianas entre os pares de média-baixa e média-alta em L1 e em L2, visando à verificação de diferenças significativas entre os dois sistemas. Ainda que tal tarefa constitua um exercício interessante para o mapeamento do processo desenvolvimental em L2, de modo a já estar sendo levada a cabo por nosso grupo de investigações, devemos reconhecer que, dado o objetivo deste trabalho de verificar a direcionalidade L2-L1 da transferência linguística, tal tarefa não se faz necessária para a satisfação do tema de investigação do presente artigo.

Quando os valores de F1 e de F2 foram comparados no espanhol (L1), por falantes monolíngues e bilíngues, verificou-se que, em termos descritivos, esses últimos elevaram todas as suas vogais, inclusive, na sua língua materna, seguindo a tendência previamente verificada em suas produções em L2. Tais aprendizes apresentaram um espaçamento “em aberto” na região vocálica média baixa. Apesar de os aprendizes de português não terem formado novas categorias para a realização das médias-baixas, eles modificaram o espaço articulatório tanto para o sistema vocálico do português quanto para o do espanhol, de modo a sugerir que uma modificação em um sistema pode modificar o outro devido à completa interconexão entre L2 e L1. Além disso, pode-se concluir, com tal resultado, que todos os componentes vocálicos estão presentes em um mesmo espaço acústico (cf. FLEGE, 1995), e que não há um espaço articulatório específico para cada sistema linguístico, mas sim um em comum para o falante bilíngue.

Verificaram-se, ainda, diferenças significativas na produção vocálica dos participantes bilíngues e monolíngues em termos de duração, tanto absoluta quanto relativa, que tenderam a mostrar o mesmo padrão. As produções em português (L2) tenderam a mostrar-se significativamente mais longas do que em espanhol (L1). Por sua vez, uma comparação entre essas produções em L1, por parte de bilíngues e de monolíngues, sugerem, também, uma duração significativamente mais longa para as vogais dos bilíngues. Esses resultados corroboram a premissa, levantada para este trabalho a partir da noção de língua como CAS, de que o próprio sistema de L1 pode sofrer efeitos da L2, corroborando, assim, a noção de bidirecionalidade da transferência vocálica.

Os resultados apresentados neste trabalho vão ao encontro da premissa dinâmica de que os sistemas de L1 e L2 se influenciam mutuamente. De acordo com a noção de língua como CAS, todas as partes de um sistema estão conectadas a todas as suas outras partes. Analisar a língua como um sistema dinâmico complexo adaptativo implica considerar que todos seus subsistemas estão inter-relacionados e, além disso, que uma mudança em um subsistema poderá resultar em mudanças em todos os outros subsistemas. Seguindo a mesma premissa, pode-se pensar que, quanto à aquisição de línguas adicionais, uma modificação em um sistema, como por exemplo, a L2, poderá causar mudança no outro sistema, isto é, o materno. Em outras palavras, conforme já citado, não há apenas um direcionamento envolvido, $L1 \rightarrow L2$, mas pode haver outras direções quanto à transferência entre os sistemas, isto é, a L2 pode vir a influenciar a L1. Há interação contínua entre todos os sistemas, o que deriva um processo dinâmico.

Os resultados apresentados neste estudo sugerem que a língua deve ser vista como um sistema dinâmico complexo e adaptativo, visto que L1 e L2 se influenciam mutuamente através de múltiplas interações. As descobertas deste estudo, dessa forma, apontam para a bidirecionalidade da influência linguística L1-L2, de modo que se rejeite a assunção de um direcionamento unilateral, da língua mais entrincheirada para a menos dominante, e se passe a conceber, também, um direcionamento tido como “não convencional”, isto é, da língua adicional em direção à língua nativa.

Referências bibliográficas

ALVES, M. A.. **Estudo dos parâmetros acústicos relacionados à produção das plosivas do Português Brasileiro na fala adulta: análise acústico-quantitativa.** Tese (Doutorado em Letras). Universidade Federal de Santa Catarina, Florianópolis, 2015.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M.; CROFT, W.; ELLIS, N.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a Complex Adaptive System: Position Paper. **Language Learning**, v. 59, n. 1, p. 1-26, 2009.

BOERMA, P.; WEENINK, D. **Praat: doing phonetics by computer.** Versão 5.4.22, 8 de Outubro, 2015. Disponível em <http://www.praat.org/>. Acesso em 11 outubro 2015.

BRITO, K. S. **Influências interlinguísticas na mente multilíngue: perspectivas psicolinguísticas e (psico)tipológicas.** Tese (Doutorado em Letras). Universidade Federal do Paraná, Curitiba, 2011.

CENOZ, J. The Effect of Linguistic Distance, L2 Status and Age on Cross-linguistic Influence in Third Language Acquisition. In CENOZ, J.; HUFSEISEN, B.; JESSNER, U. **Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives.** Clevedon: Multilingual Matters, 2001.

DE BOT, K.; LOWIE, W.; VERSPOOR, M. A Dynamic System Theory approach to second language acquisition. **Bilingualism: Language and Cognition**, v. 10, n.1, p. 7-21, 2007.

crossref <http://dx.doi.org/10.1017/S1366728906002732>

DE BOT, K.; LARSEN-FREEMAN, D. Researching second language development from a dynamic systems perspective. In: VERSPOOR, M. H.; DE BOT, K.; LOWIE, W. (eds). **A Dynamic Approach to Second Language Development: Methods and Techniques.** Amsterdam: John Benjamins Publishing Company, p. 5-24, 2011. **crossref** <http://dx.doi.org/10.1075/llt.29.01deb>

DE BOT, K.; LOWIE, W.; THORNE, S. L.; VERSPOOR, M. Dynamic System Theory as a comprehensive theory of second language development. In GARCIA MAYO, P. ; MANGADO J. G. ; ADRIAN, M. M. (Eds.). **Contemporary perspectives on second language acquisition** (pp. 167-189). John Benjamins Publishers, 2013. **crossref** <http://dx.doi.org/10.1075/aals.9.13ch10>

ELLIS, N. C. **Dynamic Systems and SLA: The wood and the trees.** *Bilingualism: Language and Cognition*, v. 10, n. 1, p. 23-25, 2007. **crossref**
<http://dx.doi.org/10.1017/S1366728906002744>

_____. The Emergence of Language as a Complex Adaptive System. In: SIMPSON, J. (ed.). **Routledge Handbook of Applied Linguistics.** Routledge /Taylor Francis, 2011.

_____; LARSEN-FREEMAN, D. Language Emergence: Implications for Applied Linguistics—Introduction to the Special Issue. **Applied Linguistics**, v. 27, n. 4, p. 558–589. Oxford University Press, 2006.

FEIDEN, J. A.; ALVES, U. K.; FINGER, I. O efeito da anterioridade e da altura na identificação das vogais médias altas e médias baixas do Português Brasileiro por falantes de espanhol. **Letras de Hoje**, Porto Alegre, v. 49, n. 1, p. 85-94, 2014. **crossref**
<http://dx.doi.org/10.15448/1984-7726.2014.1.14715>

FIFE-SCHAW, C. Levels of Measurement. In: BREAKWELL, G. M.; HAMMOND, S.; FIFE-SCHAW, C.; SMITH, J. A. (eds). **Research Methods in Psychology** (3a. ed.). London: Sabe, 2006.

FLEGE, J. E. Second Language Speech Learning: Theory, Findings, and Problems. In: STRANGE, W. (ed.) **Speech perception and linguistic experience: issues in cross language research.** Timonium, MD: York press, 1995.

GROJEAN, F. In: GROSJEAN, F.; LI, P. **The psycholinguistics of Bilingualism.** Wiley-Blackwell, 2013.

HERDINA, P.; JESSNER, U. **A Dynamical Model of Multilingualism: perspectives of Change in Psycholinguistics.** Cromwell Press Ltd., 2002.

JESSNER, U. Teaching third languages: Findings, trends and challenges. **Language Teaching**, v. 41, n.1, p. 15-58, 2008. **crossref** <http://dx.doi.org/10.1017/S0261444807004739>

LARSEN-FREEMAN, D. Chaos/Complexity Science and Second Language Acquisition. **Applied Linguistics**, v. 18, n. 2, 141-165, Oxford University Press, 1997.

_____. Complex, dynamics systems: A new transdisciplinary theme for applied linguistics? **Language Teaching**, v. 1, n. 13, Cambridge University Press, 2011.

_____; CAMERON, L. **Complex Systems and Applied Linguistics.** Oxford University Press, 2008.

MACHRY DA SILVA, S. **Aprendizagem fonológica e alofônica em L2: percepção e produção das vogais médias do português por falantes nativos do espanhol.** Tese (Doutorado em Letras). Porto Alegre:PUCRS, 2014.

MATOS, M.; SANTOS, A.; GONÇALVES, M. M.; MARTINS, C. Innovative moments and change in narrative therapy. **Psychotherapy Research**, v. 19, p. 68-80, 2009. **crossref** <http://dx.doi.org/10.1080/10503300802430657>

ORTEGA, L. **Understanding Second Language Acquisition**. Hachette UH Company, 2009.

SANTOS, G. **Percepção e produção das vogais médias do espanhol por falantes do português brasileiro**. Tese (Doutorado em Letras). Universidade Católica de Pelotas, Pelotas, 2014.

SILVA, A. H. P. Organização temporal de encontros vocálicos no Português Brasileiro e a relação entre Fonética e Fonologia. **Letras de Hoje**, v. 49, n. 1, p. 11-18, 2014. **crossref** <http://dx.doi.org/10.15448/1984-7726.2014.1.14891>

THOMAS, E. R.; KENDALL, T. **NORM**: The vowel normalization and plotting suite. 2007. Disponível em <http://ncslaap.lib.ncsu.edu/tools/norm/>. Acesso em 30 julho 2015.

Artigo recebido em: 30.10.2015

Artigo aprovado em: 29.04.2016

A Aquisição Fonológica Variável da Nasal Velar por Aprendizes de Inglês-L2: Análise pela Teoria da Otimidade Estocástica

The Phonological Acquisition of the Velar Nasal by L2 Learners of English: Analysis by The Stochastic Optimality Theory

Athany Gutierrez*

RESUMO: este artigo apresenta uma análise formal da Interlíngua de aprendizes brasileiros (gaúchos) de Inglês quanto à aquisição da nasal velar em coda silábica. A análise está associada ao algoritmo de aprendizagem da Teoria da Otimidade Estocástica, o *GLA* (*Gradual Learning Algorithm*) (BOERSMA e HAYES, 2001), que considera a aquisição como o ordenamento hierárquico de restrições linguísticas, em que a variação é representada pela sobreposição de pontos de seleção das restrições numa escala contínua. O *status* da variação na L1 (primeira língua, Português) e na L2 (segunda língua, Inglês) é distinto, tem condicionamento fonético na primeira e fonológico na segunda. O ponto de partida da aquisição da L2 é o *ranking* da L1, de configuração M(arcação)>>F(idelidade). Na formalização proposta, uma restrição AGREE (M), que exige partilha de ponto de articulação na sequência vN_#, possui valor de ranqueamento próximo ao de uma restrição IDENT (F), que requer identidade quanto ao ponto da nasal no *input* e no *output*. O modelamento feito reitera o potencial do algoritmo no que diz respeito à formalização de sistemas variáveis de L2. Adquirir a nasal velar é promover Fidelidade em direção ao *input*, através do ajuste gradual das restrições.

ABSTRACT: this article presents a formal analysis of the Interlanguage of Brazilian learners (*gaúchos*) of English regarding the acquisition of the velar nasal in syllabic coda. The analysis is associated with the learning algorithm of the Stochastic Optimality Theory, the *GLA* (*Gradual Learning Algorithm*) (BOERSMA and HAYES, 2001), which views acquisition as a hierarchical ranking of linguistic constraints, in which variation is represented by the overlapping of the selection points of the constraints in a continuous scale. The status of the variation in the L1 (first language, Portuguese) and in the L2 (second language, English) is distinct, it is due to phonetic conditioning in the first language and phonological conditioning in the second language. The starting point of the L2 acquisition is the L1 ranking, whose configuration is M(arkedness)>>F(aithfulness). In the modelling proposed, a constraint AGREE (M), which demands sharing of manner of articulation in the sequence vN_#, has a similar ranking value to a constraint IDENT (F), which requires identity concerning nasal manner of articulation in the input and in the output. The modelling confirms the potential of the algorithm concerning the formalization of variable L2 systems. Acquiring the velar nasal is to promote Faithfulness towards the input, through the gradual adjustment of constraints.

PALAVRAS-CHAVE: Aquisição. Variação. TO Estocástica. *GLA*. Nasal velar.

KEYWORDS: Acquisition. Variation. Stochastic OT. *GLA*. Velar nasal.

* Doutora em Estudos da Linguagem - Instituto de Letras (UFRGS).

1. Introdução

Neste artigo, é apresentada uma análise formal da aquisição da nasal velar em Inglês como segunda língua (L2), por aprendizes brasileiros do município de Caxias do Sul - RS, na Serra Gaúcha. Os dados utilizados para a formalização da Interlíngua advêm de uma Análise de Regra Variável (ARV) (WEINREICH; LABOV; HERZOG, [1968] 2006; LABOV, 1994; 2001; SANKOFF; TAGLIAMONTE; SMITH, 2014) realizada, constituindo, portanto, amostras autênticas da produção linguística dos aprendizes. A análise foi executada através do algoritmo de aprendizagem gradual (*GLA*) (BOERMA; HAYES, 2001).

Em Inglês, as consoantes nasais em coda silábica são fonemas distintivos, como se percebe no par *si*[ŋ] ‘cantar’ e *si*[n] ‘pecado’, por exemplo. Em Português Brasileiro (PB), as nasais em coda são subespecificadas, não possuem ponto de articulação definido, sendo sua realização condicionada foneticamente pela assimilação de ponto da vogal que as precede, no contexto de coda final¹: palatal, se a vogal precedente for anterior, como em ‘jasmin’ [ʒas’miŋ], ou velar, se a vogal precedente for posterior, como em ‘atum’ [a’tuŋ], por exemplo. O dados desta análise referem-se a ocorrências com o sufixo {ing} em Inglês, em palavras como ‘*living*’ (vivendo) e ‘*playing*’ (jogando)².

Dado o *status* diferenciado das nasais em coda silábica nos dois sistemas linguísticos, a questão que orienta este trabalho é: como acontece a aquisição da nasal velar na Interlíngua Português-Inglês, sob a perspectiva da Teoria da Otimidade (TO) Estocástica? Para responder a essa questão, este artigo organiza-se em outras quatro seções, além desta (1) Introdução: (2) revisão da literatura, (3) metodologia, (4) resultados da formalização pelo *GLA* e (5) breve discussão sobre as contribuições do estudo.

¹ Contextos de nasais em coda silábica medial foram eliminados da análise por não constituírem ambiente variável, já que a nasal que se realiza é sempre homorgânica à consoante seguinte.

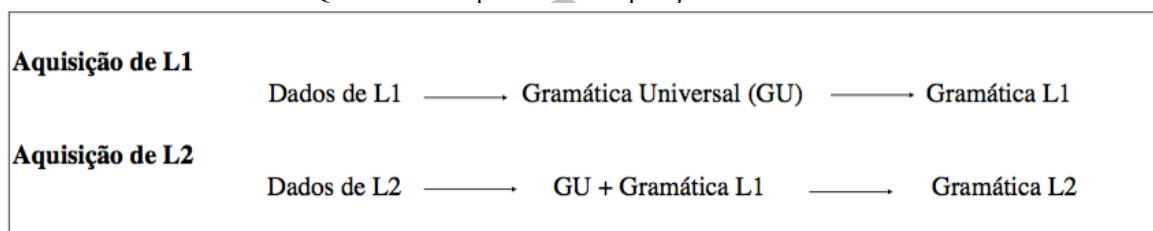
² No *corpus*, foram verificadas ocorrências com {ing} que não constituem palavras sufixadas, como *morning* ‘manhã’, além de palavras formadas pela sequência -ng, como *song* ‘canção’. Dados como esses foram eliminados da ARV por nocaute de aplicação da regra: a realização da nasal velar manifestou-se em 100% das ocorrências desses tipos.

2. Pressupostos teóricos

2.1 Fonologia nasal da Interlíngua

‘Interlíngua’ é o nome que se atribui ao sistema linguístico instável construído pelo aprendiz, que comporta características da L1 e da L2. Para Selinker (1969, *apud* SELINKER, 1994, p. 231; 214), a Interlíngua é uma “tentativa do falante em produzir uma norma estrangeira”³; “um processo que reflete hipóteses (universais) sobre o *input* de L2” e institui-se como um “comportamento altamente estruturado”. A partir da consolidação do termo, os estudos sobre a aquisição de L2 passaram a considerar a produção dos aprendizes não apenas como uma coleção aleatória de erros, mas como um *sistema linguístico per se*. Tem-se a pressuposição de que, quanto mais inicial o estágio de aprendizagem da L2, mais similar à L1 é a Interlíngua; e quanto mais avançado, mais semelhante à L2. O Quadro 1 sintetiza a estrutura do processo de aquisição linguística.

Quadro 1. Componentes da aquisição de L1 e de L2.



Fonte: adaptado de Mohanan e Mohanan (2003, p. 8)

Como se observa no Quadro 1, durante a aquisição de primeira língua, a criança é exposta a dados linguísticos do ambiente familiar que, com o componente da GU (Gramática Universal)⁴, estruturam gradativamente o sistema de L1. Na aquisição de segunda língua, a criança ou o adulto são expostos a dados de L2 (provenientes de falantes estrangeiros, no caso de a língua alvo ser adquirida no país de origem do aprendiz) que, associados à GU e à gramática da L1, estruturam gradativamente o sistema da L2 em desenvolvimento, a Interlíngua.

Em Português Brasileiro (PB), não há palavras terminadas por oclusivas velares como há em Inglês (*think* ‘pensar’, *sing* ‘cantar’). Assume-se que, diante de uma palavra com tal

³ Todas as traduções do Inglês são traduções livres feitas pela autora.

⁴ É importante mencionar que existem propostas teóricas que desconsideram o acesso da GU na aquisição, seja de L1 ou de L2. Autores como Bley-Vroman (1989, p. 50), por exemplo, advogam que o acesso à GU dá-se apenas durante a aquisição da língua na infância, já que, na fase adulta, o sistema inato já estaria “preenchido” por conhecimentos da primeira língua e por um sistema responsável pela resolução de problemas abstratos gerais, não apenas linguísticos.

especificação em coda, um falante de PB que desconhece a Língua Inglesa, ao tentar pronunciá-la, apoia-se na grafia da palavra e insere uma vogal epentética para satisfazer a condição silábica de sua língua, produzindo uma forma de saída como [ˈsɪŋ.gɪ], diferente do que é comumente produzido por um falante nativo de Inglês, [sɪŋ], já que a oclusiva velar após a nasal não é produzida no dialeto padrão.

A nasal velar produzida em [ˈsɪŋ.gɪ] é manifestação fonética, homorgânica ao segmento anterior (CÂMARA JR., [1970] 2005; [1971] 2002; CAGLIARI, 1977; BATTISTI, 2014). Sem ser instruído⁵, o falante de PB provavelmente desconheceria a não realização da oclusiva velar final na fala dessa palavra. O mesmo acontece com palavras do Inglês utilizadas cotidianamente no Português, como o nome do hipermercado ‘Big’, comumente pronunciado como [ˈbɪgɪ].

Por outro lado, ao se considerar um falante de PB que é aprendiz de Inglês e que, por essa razão, é orientado a não pronunciar a oclusiva final de palavras terminadas por {ing} ou por /ŋg/ (muito embora isso não o impeça de criar uma epêntese após a oclusiva), uma palavra como ‘sing’ passa a ter duas possibilidades de realização da nasal em coda: com a nasal palatal [sɪŋ], assimilando traços da anterioridade da vogal precedente (realização fonética do PB), ou com a nasal velar [sɪŋ], resultante de um esforço articulatório do falante, que precisa combinar o traço [+anterior] da vogal alta com o traço [+dorsal] da velar seguinte. A realização com a nasal velar aproxima-se da língua alvo, o Inglês, e pode ser produto de uma representação subjacente advinda da instrução explícita, com a nasal velar. Os dois exemplos constituem hipóteses de realizações para a nasal em coda nos dados da Interlíngua, dispensando, por ora, possíveis ocorrências com a oclusiva após a nasal, provavelmente com epêntese, que são resultantes da transferência de padrões da L1 para a L2 (relação grafema x fonema⁶) e não exibem contexto de variação.

Diante do exposto, assume-se uma (inter)fonologia que inclui características da nasalidade da L1 e da L2 dos falantes. Tem-se por base teórica (i) informações da fonologia do PB em relação à realização do arquifonema nasal /N/ em coda, como foram nessa seção

⁵ Instruído, neste caso, significa receber instruções de um professor quanto à pronúncia dos sons do Inglês. Embora os informantes deste estudo sejam alunos de Inglês de um Programa de Línguas, a instrução explícita não foi uma variável medida nesta investigação. Todos os comentários a esse respeito constituem observações da pesquisadora a partir de sua experiência na época como docente na Instituição onde os dados foram coletados.

⁶ Para saber mais sobre transferência grafo-fono-fonológica, ver ZIMMER e ALVES (2006), ALVES e CABANERO (2008).

expostas, e (ii) premissas da Teoria da Otimidade, que entende a aquisição de um sistema linguístico como o ordenamento de restrições universais. No caso de uma Interlíngua, restrições de ambos os sistemas (L1 e L2) estão operantes variavelmente em direção à aquisição da língua alvo. A seção seguinte explicita como se dá a aquisição variável de uma L2 pela TO Estocástica.

2.2. Teoria da Otimidade Estocástica

A Teoria da Otimidade (TO) *standard* (PRINCE; SMOLENSKY, [1993] 2004 doravante P&S, [1993] 2004) é um modelo representacional da gramática gerativa, surgido na década de 90, cujo foco principal é a descrição formal das línguas através de restrições linguísticas, e não mais de regras, como se fazia na teoria gerativa clássica. Apesar de não se restringir à fonologia, é nessa área que se concentra a maior parte dos estudos desenvolvidos recentemente.

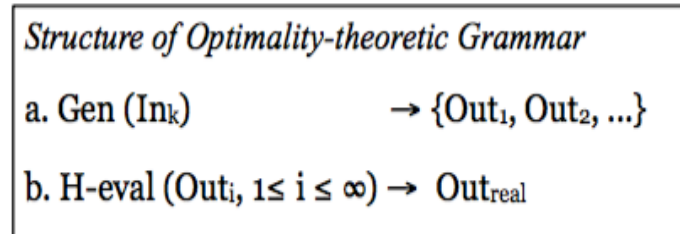
A TO é um modelo gramatical que interpreta e analisa a estrutura das línguas com base na interação entre restrições no sistema linguístico do falante. Com a TO, o processamento da gramática deixa de ser traduzido via regras e de forma serial, e passa a ser compreendido a partir de restrições e de forma paralela. A premissa básica do modelo é a de que as restrições são universais e o ordenamento dessas restrições é particular, o que torna possível a diferenciação das línguas; ou seja, todas as restrições estão presentes em todas as gramáticas, apesar de algumas restrições causarem pouco ou nenhum efeito em determinadas estruturas linguísticas em razão de seu baixo ordenamento.

Pela TO, a fala é resultado de um mapeamento entre formas de entrada (*input*), as representações mentais, e formas de saída (*output*), as representações que se efetivam na produção linguística. O processamento da gramática do falante é operado por componentes da GU: CON (*CON*straints), o conjunto universal de restrições linguísticas; GEN (*GEN*erator), o constituinte que gera candidatos a *output* e EVAL (*EVAL*uation), o constituinte que avalia esses candidatos, com base em CON; além de componentes particulares da língua do falante, como o LEXICON, que contém as representações fonológicas e informações morfológicas dos itens lexicais, além da hierarquia de CON.

A gramática opera da seguinte maneira: o dispositivo GEN cria diferentes candidatos a *output* (Out_1, Out_2, \dots), que são avaliados por EVAL com base num conjunto CON de restrições ($Out_i, 1 \leq i \leq \infty$) hierarquicamente dispostas da esquerda para a direita. Dessa avaliação, emerge a forma “ótima” (Out_{real}), que se manifesta na fala do indivíduo como derivação fonética,

representação física desse processo subjacente. Tal processamento pode ser observado na Figura 1.

Figura 1. Processamento da gramática na TO.



Fonte: P&S, [1993] 2004, p.4.

Inicialmente, a TO foi pensada para a formalização de gramáticas categóricas, já que, segundo a teoria, apenas um candidato pode ser selecionado como ótimo. Derivada da visão gerativista de linguagem, que trabalha com a hipótese de falantes e gramáticas ideais, e preocupada essencialmente com a competência linguística, a proposta é de certa forma limitada ao considerar-se a natureza variável da linguagem, evidenciada no desempenho linguístico dos falantes. Essa dualidade categoricidade \times variabilidade impõe alguns desafios ao modelo, cujas origens são de natureza categórica.

Tendo em vista tal limitação, modelos da TO advindos da proposta *standard* foram desenvolvidos a fim de dar conta de dados variáveis, como a TO Estocástica de Boersma e Hayes (2001). O *GLA*, algoritmo operante nessa proposta, é de natureza estocástica: lida com processos probabilísticos, não-aleatórios. Atua sob os princípios da teoria padrão, mas diferencia-se desta por atribuir valores numéricos (pesos) às restrições linguísticas e justificar mais de uma forma de saída ótima, dando conta da variação encontrada na fala.

Por se tratar de um mecanismo que simula a aquisição da língua, o *GLA* tem sido utilizado numa série de investigações sobre a aquisição de L2 no Brasil (ALVES, 2008, 2009, 2010, 2013; AZEVEDO, 2011; GARCIA, 2012; GUIMARÃES, 2012; SCHMITT; ALVES, 2014; ALVES; LUCENA, 2014). O algoritmo pressupõe que, se o aprendiz tem acesso a um inventário fonológico de restrições universais, então existe a possibilidade de que qualquer gramática seja aprendida, considerando-se a existência de um mecanismo disponível que ranqueie as restrições com base em dados de *input* (BOERSMA; HAYES, 2001). O *GLA* é também sensível ao erro do aprendiz (*error-driven*): compara as formas produzidas (*outputs*) com a evidência positiva, a forma “esperada” ou “alvo” (*input*).

A aquisição de determinada estrutura linguística é processada de modo contínuo e gradual, incluindo estágios de variação em seu desenvolvimento. É formalizada pela demissão e promoção de restrições, que recebem um peso numérico inicialmente arbitrário e são dispostas numa escala contínua. O deslocamento das restrições para a direita ou para a esquerda na escala e a sua eventual sobreposição de faixa de valores indica a variação da língua do falante em direção à aquisição da forma-alvo.

Em termos de funcionamento, o *GLA* é executado através do *software Praat*, versão 5.4.08 (BOERSMA; WEENINK, 2015). Nesse programa, os dados são carregados em dois arquivos de extensão .txt (*scripts*), simultaneamente. Um deles (o arquivo de distribuição) contém as informações de base da gramática: o *input* e os *outputs* com suas respectivas frequências de realização. O outro arquivo (arquivo de especificações) carrega as demais informações necessárias ao processamento: as restrições, com valores para marcação e fidelidade, e um *quasi-tableau*⁷, indicando as violações de cada candidato para as restrições especificadas. Os valores atribuídos para a frequência de realização dos *outputs* são geralmente aleatórios, refletindo padrões observados na fala pelo analista, mas devem totalizar 100 para que o algoritmo opere corretamente.

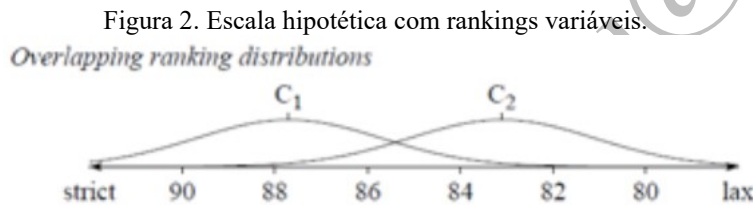
Ao ser alimentado de dados e restrições, o *GLA* executa seu processamento mediado por um valor de plasticidade (*default* definido pelo programa: 1.0) e um valor de ruído (*default* definido pelo programa: 2.0), adicionado a cada novo evento de fala. As restrições recebem valores numéricos de dois tipos:

- (i) de ranqueamento (*ranking values*), pesos mais ou menos fixos que determinam a faixa de abrangência dos pontos de seleção indicativos da variação (a partir do valor de ranqueamento, 5 pontos para a esquerda e 5 pontos para a direita), e
- (ii) de ponto de seleção (*selection points*), pesos variáveis que localizam o momento de fala na faixa de valores de ranqueamento (em uma área de abrangência de 10 pontos).

⁷ Um *quasi-tableau* é uma tabela que permite a visualização das marcas de violação incorridas pelos candidatos a *output*. Constitui-se como um “*quasi-tableau*” por não existir concorrência entre candidatos e, assim, nem a escolha de um candidato ótimo (de LACY, 2002).

A faixa de abrangência das restrições é uma distribuição de probabilidades, em que é possível calcular a probabilidade de ocorrência de uma ou outra(s) gramática(s), através da suposição de que um ponto de seleção vai ocorrer mais à esquerda ou mais à direita do valor de ranqueamento. Para a medida probabilística, assume-se uma distribuição normal (Gaussiana)⁸ com desvio padrão de 2.0 (ruído) para mais ou para menos, e ainda somando-se ao valor de plasticidade.

Na Figura 2, a faixa de abrangência das restrições hipotéticas C1 e C2 é representada através de ondas Gaussianas. Quanto mais à esquerda o ponto de seleção localizar-se, menor é sua chance de movimentação no *ranking*.



Fonte: Boerma e Hayes (2001, p. 5).

A respeito da gramática representada na Figura 2, pode-se afirmar que:

- (i) o ponto de seleção da restrição C1 é aproximadamente 88; se a faixa de abrangência varia entre 82 e 92, o valor de ranqueamento é provavelmente 87;
- (ii) o ponto de seleção da restrição C2 é aproximadamente 83; se a faixa de abrangência varia entre 78 e 88, o valor de ranqueamento é provavelmente 83;
- (iii) dados tais valores e, desse modo, a possibilidade de sobreposição de valores de seleção (observada na área comum das duas restrições, entre os pontos 83 e 87, aproximadamente), a variação é verificada pelos ranqueamentos $C1 \gg C2 \sim C2 \gg C1$.

Como pode ser observado, é a proximidade ou afastamento dos valores de seleção que determinam a variação: se duas ou mais restrições apresentarem valores de ranqueamento cuja diferença é inferior a 10 (no exemplo dado: $C1 \cong 88$ e $C2 \cong 83$, com diferença de 5 pontos na

⁸ Uma distribuição Gaussiana prevê um único pico no centro da curva, cujos valores próximos a esse ponto são os mais prováveis de ocorrerem.

escala), há probabilidade de que as restrições sobreponham-se, gerando *rankings* variáveis. Caso contrário, o ordenamento é estrito e não há variação.

Trabalhos de base empírica indicam o estágio inicial da aquisição como uma etapa governada por restrições de Marcação (GNANADESIKAN 1995; LEVELT, 1995; LEVELT; SCHILLER; LEVELT, 2000; PATER; WERLE, 2001; CURTIN; ZYRAW, 2002; ADAM, 2003; MOHANAN; MOHANAN, 2003; BOERSMA; LEVELT, 2004). Alguns desses estudos investigam a aquisição da linguagem na infância, em que o estágio inicial da aquisição é a própria GU e o alvo é a gramática do adulto; outros tratam da aquisição de L2 por adultos, que tem como etapa inicial a L1 e o alvo é aquisição de uma gramática bastante próxima a de um falante nativo da língua alvo.

De qualquer forma, seja na infância, seja no início da aquisição de L2 na idade adulta, a gramática é bastante restrita e as produções linguísticas dão preferência a formas não marcadas na língua. Durante o seu desenvolvimento, a gramática do aprendiz vai modificando-se com base na evidência positiva, até atingir estágios mais avançados da Interlíngua, ou um estágio final, mais estável. Em se tratando de aquisição pela TO, Boersma e Levelt (2004, p. 02) pontuam:

Na TO, a ideia básica é que restrições são inatas e universais e apresentam-se com um ranqueamento inicial em que todas as restrições de marcação (segmentais, silábicas e de boa formação prosódica) estão ranqueadas acima de todas as restrições de fidelidade (restrições que exigem similaridade entre as representações subjacente e de superfície). O aprendiz precisa adquirir um ranking linguo-específico dessas restrições. Através de reordenamentos subsequentes, a gramática inicial gradualmente desenvolve-se até a gramática alvo final.

Nesse sentido, o desenvolvimento da Interlíngua, que é caracterizado por variação, é entendido como um reordenamento de restrições. Conforme Boersma e Levelt (2004, p. 08), “o reordenamento encerra quando os *outputs* da gramática em desenvolvimento e aqueles da gramática alvo são idênticos”; o que seria, nesse caso, o estágio final da aquisição.

Os dados empíricos deste estudo são oriundos de uma análise variacionista realizada, cujas frequências variáveis observadas quanto à aquisição da nasal velar constituem valores estatisticamente validados para o modelamento da gramática. As informações essenciais a respeito da metodologia de coleta de dados são fornecidas na próxima seção.

3. Metodologia

O método empregado para coleta dos dados foi a ARV⁹ (WEINREICH; LABOV; HERZOG, [1968] 2006; LABOV, 1994; 2001; SANKOFF; TAGLIAMONTE; SMITH, 2014). O *corpus* é proveniente da fala gravada de aprendizes de Inglês de dois níveis iniciais de proficiência, organizados em dois grupos: 5 informantes de nível básico e 5 de nível pré-intermediário. Foram gravadas 10 horas de fala de cada grupo de aprendizes em encontros de conversação organizados pela pesquisadora, no Programa de Línguas Estrangeiras da Universidade de Caxias do Sul (PLE/UCS).

Os informantes preencheram questionários, termos de consentimento para a gravação de sua fala e realizaram o *Oxford Online Placement Test*, para confirmação de seu nível de proficiência em Língua Inglesa, que foram equivalentes à categorização recebida pelo PLE/UCS. As conversas gravadas seguiram propostas de tópicos pré-organizados pela pesquisadora, de modo a elicitarem as formas esperadas, com a nasal velar. Os estudantes ficaram dispostos numa mesa redonda, com o gravador no centro da mesa, apoiado com um suporte móvel de modo que o microfone ficasse posicionado para cima. O dispositivo utilizado para a gravação dos dados foi um gravador modelo Sony, ICD-PX312. Como os encontros foram realizados em Janeiro/2014, não havia circulação de estudantes e funcionários pelo prédio, garantindo o silêncio do ambiente.

O tratamento do *corpus* seguiu as normas da ARV, na tentativa de replicar um estudo variacionista da linha de Labov para dados de L2. Para a seleção e submissão dos dados à análise perceptual, a analista ouviu, transcreveu e codificou todas as ocorrências com nasais em coda três vezes, de modo que pudessem ser classificadas como (i) palatal, (ii) velar ou (iii) velar seguida de oclusiva, com ou sem epêntese. Cada ocorrência foi ouvida três vezes, e dados duvidosos foram submetidos a um juiz. Ocorrências como aquelas em (iii) foram eliminadas da análise pelas razões já mencionadas.

⁹ A Análise de Regra Variável é um modelo de análise empírica cunhado por Labov na década de 60. Concebe a variação fonológica como um processo correlacionado a fatores sociais e estruturais, passível de quantificação (a variação é sistemática e apresenta taxas maiores ou menores de ocorrência em determinados contextos). O objetivo da ARV é separar, quantificar e testar o efeito de diferentes fatores sobre um fenômeno variável da língua. A obra de Weinreich, Labov e Herzog constitui uma das primeiras tentativas de estabelecimento de um método para a investigação da variação e mudança linguística, e o modelo de Sankoff, Tagliamonte e Smith refere-se ao programa de regressão logística desenvolvido para medir estatisticamente a variação e os seus condicionadores sociais e estruturais.

Dados como (i) e (ii) constituíram 385 ocorrências variáveis de nasal em coda, palatal ou velar (ver nota de rodapé 13). Após carregados e rodados no *Goldvarb*¹⁰ (SANKOFF; TAGLIAMONTE; SMITH, 2014), obtiveram-se as seguintes frequências de realização das nasais em coda: 37% nasal velar e 63% nasal palatal. A seção 4 explicará como tais dados são computados via algoritmo de aprendizagem gradual, representativo da aquisição da L2.

4. Resultados

4.1 Estágio inicial: o ranking da L1

Pensar na aquisição de uma estrutura fonológica na L2 implica considerar a pré-existência de um *ranking* de L1, que é o estágio inicial da aquisição nessas condições (ALVES, 2013). Assim, o processo de aquisição da língua alvo terá como alicerce restrições que organizam a L1 dos falantes, em competição com restrições da L2, variavelmente (MONAHAN, 2001; MONAHAN; MONAHAN, 2003; BROSELOW, 2004; entre os já citados). Dada a L1 dos falantes (PB) e a L2 sendo aprendida (Inglês), a hierarquia inicial da Interlíngua terá uma configuração semelhante ao *ranking* da L1, em que restrições de Marcação estão no topo (são dominantes) e restrições de Fidelidade estão na base (são dominadas): M>>F. Tal configuração adapta-se às exigências dos *outputs*, representando sua infidelidade em relação aos *inputs*, ou a tendência do falante em produzir formas menos marcadas nessa etapa da aprendizagem.

Em se tratando de coda silábica de fronteira de palavra, que é, particularmente, o contexto fonológico em análise neste estudo, a nasal final manifesta-se naturalmente pela assimilação de traços da vogal precedente (ou, eventualmente, pela produção ‘ilegal’ da oclusiva final de {ing}¹¹, desencadeando a realização da nasal velar). Levando-se em conta os

¹⁰ O *Goldvarb* é um *software* de análise estatística utilizado para a quantificação de dados linguísticos e para a verificação das proporções de aplicação e não aplicação de uma regra variável, além dos pesos relativos dos diferentes fatores considerados, o que resulta na escolha do indivíduo sobre duas ou mais formas alternantes, sem distinção de significado. Nesse tipo de análise, os graus de realização das variantes linguísticas e a relação entre fatores sociais e estruturais são estabelecidos por um modelo de regressão logística.

¹¹ Dada a terminação dos dados que compõem o *corpus* (palavras com {ing}) e as condições de exposição e uso da língua alvo pelo aprendiz, que incluem instrução explícita, há a possibilidade de (i) realização da consoante oclusiva seguinte à nasal (*studying* [‘stʌdɪŋg] ‘estudando’), por razões de transferência da relação grafema-fonema da L1 para a L2; e, conseqüentemente, (ii) inserção de uma vogal epentética após a oclusiva, já que esse segmento consonantal é proibido em coda no PB. Em um ou outro caso, a nasal que se realiza é velar, por assimilação de ponto de articulação. Ocorrências como estas manifestaram-se na produção dos aprendizes (119/504), mas foram excluídas da ARV por não apresentarem contexto de variação, já que a nasal, seguida de oclusiva velar, sempre será velar.

outputs produzidos *sem* a oclusiva velar /g/, quais restrições estruturam a Interlíngua e estão competindo variavelmente para a produção das nasais velar e palatal em coda silábica? Na perspectiva da TO Estocástica, todas as ocorrências encontradas no *corpus* (incluindo-se formas produzidas com a oclusiva velar seguinte à nasal, com ou sem epêntese) são consideradas como tentativas de acerto por parte do aprendiz em direção à evidência positiva e, por essa razão, candidatos a *output* em sua gramática.

O falante de PB, aprendiz de Inglês-L2, aprende a comunicar-se na língua alvo com apoio do registro escrito, apresentando uma tendência a pronunciar as palavras conforme a sua grafia. Pensando-se num aprendiz inicial, *real beginner*, que nunca teve contato com o Inglês formal através de instrução, e que se depara pela primeira vez com uma forma linguística como ‘*looking*’ (olhando), por exemplo, entende-se que sua tentativa inicial de produção refletirá uma relação idêntica entre grafemas e fonemas na palavra, atribuindo um fonema a cada grafema do item lexical. Ou seja: o falante aprendiz vai pronunciar a palavra tal qual ela é escrita.

Levando-se em consideração tal perfil, uma palavra como ‘*looking*’, em PB, teria uma forma de *input* fiel à grafia, /lukɪŋ/, já que esse falante, inicialmente, não tem ainda contato com a língua e nem está exposto a qualquer tipo de evidência positiva que lhe dê indícios sobre como essa palavra é pronunciada na língua alvo. Afinal, que razões haveria para que uma forma lexical com a nasal velar, sem a oclusiva, (/‘lukɪŋ/) (forma de *input* a ser assumida pela Interlíngua), constituísse a representação subjacente da gramática da L1?

Frente ao exposto, pergunta-se: que possibilidades de realização linguística estão previstas pela gramática do PB, uma língua caracterizada por Marcação, principalmente no que se refere ao licenciamento da nasal em coda silábica?

Em primeiro lugar, é pouco provável que o falante produza um *output* fiel ao *input*, porque no PB consoantes oclusivas não são permitidas em coda. Um candidato a *output* como (i) [‘lukɪŋg], fiel ao *input*, seria eliminado por uma restrição de Marcação que proíbe segmentos oclusivos nessa posição na sílaba, *CODA(stop)¹², formalizada em (1):

(1) *CODA(stop)

Atribua uma marca de violação (*) para cada consoante oclusiva presente em coda silábica.

¹² Todas as restrições aqui apresentadas foram adaptadas da lista de restrições em McCarthy (2004, p. 595-597).

Não sendo permitida a produção de segmentos oclusivos em coda final, o falante pode empregar duas estratégias de reparo silábico, de modo a adaptar a produção às condições fonotáticas de sua língua: (a) a inserção de uma vogal epentética, gerando um candidato a *output* como (ii) [ˈlukɪŋɨ], ou (b) o cancelamento da oclusiva /g/, gerando um candidato como [ˈlukɪN] que, dado o condicionamento do arquifonema nasal, manifesta-se com a nasal palatal (iii) [ˈlukɪŋ] em PB.

Se o falante optar por (a), tem-se um candidato epentetizado (ii) violando uma restrição de Fidelidade DEP, que proíbe inserção de segmentos no *output*. Esta restrição é formalizada em (2):

(2) DEP

Atribua uma marca de violação (*) para cada segmento do *output* que não estiver presente no *input* (não insira segmentos).

Ao produzir uma forma como (ii) [ˈlukɪŋɨ], que respeita as condições fonotáticas do Português, a gramática da L1 assumiria uma configuração tal qual expressa no *Tableau 1*:

Tableau 1 - Ranking parcial da L1

/lukɪŋg/	*CODA(stop)	DEP
lukɪŋg	*	
→ lukɪŋɨ		*

Fonte: a autora (2015)

Se o falante optar por (b), tem-se um candidato (iii) [ˈlukɪŋ] que viola MAX, restrição de Fidelidade que proíbe o apagamento de segmentos no *output*. Essa restrição é formalizada em (3):

(3) MAX

Atribua uma marca de violação (*) para cada segmento presente no *input* que não estiver presente no *output* (não apague segmentos).

Ao cancelar a consoante oclusiva do *input*, o candidato (iii), com a nasal palatal, se realiza com a nasal homorgânica ao ponto de articulação da vogal precedente (anterior). A gramática ficaria assim modelada:

Tableau 2 - Ranking da L1 (primeira possibilidade)

/ʌkɪŋ/	*CODA(stop)	DEP	MAX
ʌkɪŋg	*		
ʌkɪŋɣ		*	
→ ʌkɪŋ			*

Fonte: a autora (2015).

Se um candidato com a nasal velar (iv) [ˈʌkɪŋ] fosse considerado, as restrições da gramática no Tableau 2 não seria suficientes para resolver o conflito, como se vê no Tableau 3:

Tableau 3 - Ranking problemático da L1

/ʌkɪŋ/	*CODA(stop)	DEP	MAX
ʌkɪŋg	*		
ʌkɪŋɣ		*	
→ ʌkɪŋ			*
→ ʌkɪŋ			*

Fonte: a autora (2015).

O Tableau 3 não resolveria a disputa da gramática pelo candidato ótimo, já que tanto o *output* com a nasal velar quanto aquele com a nasal palatal incorrem o mesmo número de violações para a restrição MAX. Isso explica porque um candidato com a nasal velar em coda final não emerge naturalmente em PB. Considerando-se que o *output* (iii), com a nasal palatal, é a realização da nasal em coda quando precedida pela vogal alta anterior, seria necessária uma restrição que eliminasse o candidato (iv), proposta em (4):

(4) AGREE(place)VN#¹³

Atribua uma marca de violação (*) para a sequência Vogal + Nasal em coda silábica final que não partilhar do mesmo ponto de articulação.

Finalmente, o *Tableau* 4 mostra a resolução do conflito entre as nasais velar e palatal no PB:

Tableau 4 - Ranking da L1 (segunda possibilidade)

/lɔkɪŋg/	*CODA(stop)	AGREE(place)VN#	DEP	MAX
lɔkɪŋg	*			
lɔkɪŋɣ			*	
→ lɔkɪɲ				*
lɔkɪɲ		*		*

Fonte: a autora (2015).

No *Tableau* 4, as duas restrições de Marcação são as restrições dominantes neste *ranking* (*CODA e AGREE) e, apesar de altamente ranqueadas, não se apresentam em relação de dominância estrita, o que é indicado pela linha pontilhada no *Tableau* (*CODA(stop)>>AGREE(place)VN# ou AGREE(place)VN#>>*CODA(stop)). A dominância da Marcação em L1 elimina de imediato os candidatos (i) e (iv). O candidato (ii) é eliminado ao inserir a vogal epentética, ferindo DEP, a terceira restrição mais alta no ordenamento. O candidato vencedor é (iii), com a palatal, incorrendo apenas uma violação na restrição mais baixa da hierarquia (MAX).

Perante as gramáticas apresentadas nos *Tableaux* 1-4, é possível verificar que o *Tableau* 1 pode se referir, especificamente, a um falante de PB sem conhecimento de Inglês, ou a um falante *real beginner*, sem nenhuma experiência prévia com a língua. Os *Tableaux* 2-4 representam um *ranking* de L1 de um falante de PB que é aprendiz de Inglês-L2, que tem conhecimento sobre a não produção da oclusiva velar /g/ em palavras sufixadas por {ing}.

¹³ Bisol (2008) utiliza a mesma restrição, assim nomeada: “AGREE(VC) = o segmento nasal em coda deve concordar em ponto de articulação com a vogal precedente”.

Dessa maneira, o *output* (iii), [ˈlukɪŋ], é o candidato mais harmônico¹⁴ escolhido pela gramática da L1, expressa pelo *Tableau* 4. Assim chega-se à configuração da gramática da L1 que é o ponto de partida da aquisição da L2. Tomando-se como início da aquisição da L2 o *Tableau* 4, a gramática da L1 reitera o estágio inicial da aprendizagem como uma etapa caracterizada por *outputs* infieis ao *input*. Sendo a nasal velar em L1 uma realização fonética possível de /N/ e, em L2, um fonema contrastivo, /ŋ/ apresenta-se como uma fonte de dificuldade ao falante de PB, aprendiz de Inglês. Investigações sobre a aquisição de L2 (KOERICH, 2002; BAPTISTA; SILVA FILHO, 2006; ALVES, 2008) apontam o traço [+dorsal], característico da nasal velar, como sendo o mais marcado em relação ao traço [+coronal], característico da nasal palatal. Tal afirmação sustenta-se, nos estudos supracitados, com base nas evidências empíricas relativas à maior ocorrência de epêntese após segmentos oclusivos, servindo também como evidência para o modelamento proposto no *Tableau* 1.

4.2 Estágio atual: o ranking da Interlíngua

Após realizada a simulação da aquisição da nasal velar com o *GLA*, obtiveram-se duas gramáticas variáveis, responsáveis pela produção das nasais palatal e velar, nomeadas, respectivamente, de gramática *GLA1* e *GLA2*, cuja formalização é apresentada nesta seção.

Os valores das restrições para a gramática *GLA1* podem ser observados na Tabela 1:

Tabela 1. Valores de ranqueamento e desarmonia na gramática *GLA1*.

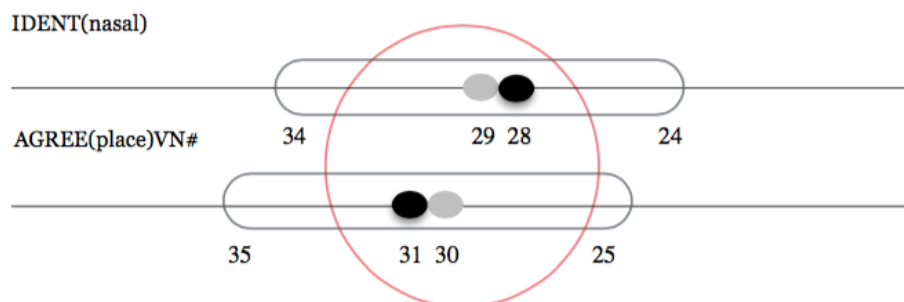
	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
*CODAstop	100.000	100.319	1.000000
DEP	40.808	41.214	1.000000
AGREEplaceVN	30.061	30.775	1.000000
IDENTnasal	29.132	28.042	1.000000

Fonte: a autora (2015).

¹⁴ É importante fazer a ressalva de que um *output* como o (ii), [lu.kɪŋ.gɪ], constituiria-se como um bom candidato na gramática da L1 (dadas as realizações da oclusiva foi verificadas na ARV), pois respeita condições de boa estruturação do PB. Entretanto, a eliminação de tal candidato para a formalização da aquisição variável da nasal velar na Interlíngua é uma decisão da analista, cujo foco da investigação reside somente na análise variável das *nasais* produzidas em coda silábica, eliminando contextos em que a sua produção é categórica, muito embora a realização da oclusiva epentetizada sinalize um desenvolvimento em direção à aquisição da nasal em coda final no Inglês.

Conforme apresentado em 2.2, a variação na TO Estocástica é verificada pela proximidade dos valores das restrições (<10). Como se pode observar na Tabela 1, AGREE e IDENT apresentam uma diferença bem inferior a 10 pontos, o que vai ocasionar uma sobreposição de seus pontos de seleção (determinados pelo valor de desarmonia) na área de abrangência das restrições (delimitada em 5 pontos para a esquerda e 5 pontos para a direita a partir do valor de ranqueamento), gerando variação. Na Figura 3, IDENT(nasal) e AGREE(place)VN# são representadas simultaneamente: a área de abrangência das restrições são definidas a partir do valor de ranqueamento (pequeno círculo cinza) e o ponto de seleção/desarmonia é marcado pelo pequeno círculo preto. O círculo vermelho indica, aproximadamente, a faixa comum entre as restrições, área onde os valores de desarmonia¹⁵ podem situar-se em distintos momentos de produção linguística.

Figura 3 - Valores de ranqueamento e desarmonia para IDENT(nasal) e AGREE(place)VN# na GLA1



Fonte: autora (2015).

Os valores de desarmonia/seleção para as restrições IDENT e AGREE são, respectivamente, 28 e 31, sendo a diferença entre esses valores igual a 3 (<10). Em diferentes momentos de fala, esses pontos de seleção movimentarão-se pouco mais para a esquerda ou para a direita, representando a variação na fala dos aprendizes.

¹⁵ Os valores de desarmonia e ranqueamento obtidos no algoritmo foram arredondados desta forma: se a primeira casa decimal for até 5, o valor mantém-se (exemplo: 29.132 = 29); se esse valor ultrapassar 5, adiciona-se 1 (exemplo: 40.808 = 41).

Se *CODA(stop) proíbe a oclusiva em coda, o candidato [ˈlukɪŋ] será eliminado da disputa. Se a segunda restrição dominante na gramática é DEP, o candidato [ˈlukɪŋɡɪ] é também eliminado. Em um momento de fala tal qual apresentado pelos valores da primeira rodada do *GLA*, a restrição AGREE(place)VN# tem ponto de seleção superior a IDENT(nasal) (a primeira está ranqueada acima da segunda), impondo à gramática a preferência por um *output* que partilhe de ponto de articulação no ambiente vN_# em detrimento a um que partilhe da mesma especificação de traço nasal com o *input*. Desse modo, o vencedor, nesse momento, é o candidato com a nasal palatal ([ˈlukɪŋ]), como se vê no *Tableau 5*:

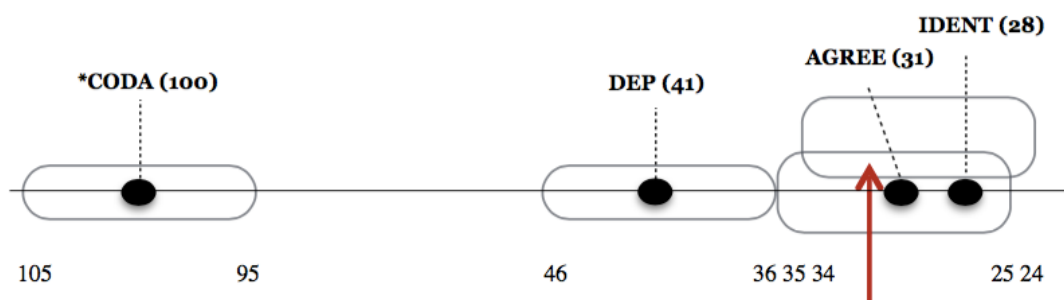
Tableau 5 - Gramática GLA1

lukɪŋ	*CODAstop	DEP	AGREEplaceVN	IDENTnasal
lukɪŋ			*!	
lukɪŋ				*
lukɪŋɡ	*!	*		
lukɪŋɡɪ		*!*		

Fonte: a autora (2015).

Quando o aprendiz realiza a nasal palatal, a gramática variável da Interlíngua configura-se desta maneira, tal qual expressa pela Figura 4:

Figura 4 - Representação da gramática variável GLA1



Fonte: a autora (2015).

De acordo com a Figura 4, a restrição *CODA permanece estável e afastada das demais, de modo a não ser deslocada no *ranking* e não gerar variação; sua posição é fixa nessa hierarquia

indica que, na Interlíngua, tal como na L1, não serão produzidos *outputs* com a oclusiva velar em coda silábica em fronteira de palavra. DEP e AGREE apresentam um diferença exata de pouco mais de 10 pontos (10.439), então é provável que essas restrições também não interajam na produção variável da nasal velar. A variação concentra-se na sobreposição de pontos de seleção/desarmonia entre IDENT e AGREE, indicada pela flecha vermelha na Figura 4.

Em um segundo momento de avaliação (adição do ruído, equivalente a um novo evento de fala), estes são os valores obtidos para as restrições, gerando um novo ordenamento (gramática *GLA2*) exibidos na Tabela 2.

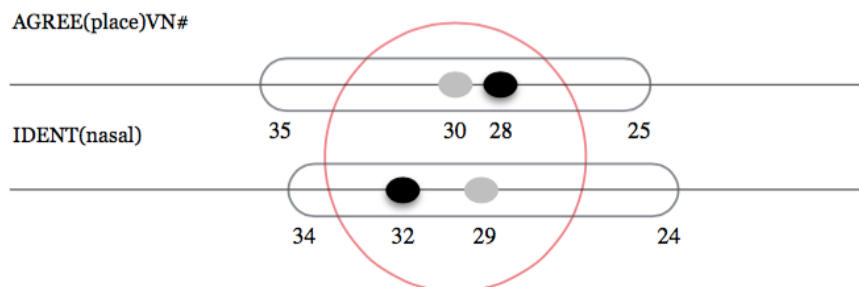
Tabela 2 - Valores de ranqueamento e desarmonia na gramática *GLA2*

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
*CODAstop	100.000	99.154	1.000000
DEP	40.808	38.490	1.000000
IDENTnasal	29.132	32.374	1.000000
AGREEplaceVN	30.061	28.373	1.000000

Fonte: a autora (2015).

Neste momento, a gramática mantém os valores de ranqueamento para a restrições, mas estabelece novos pontos de seleção/desarmonia. O ponto de seleção de IDENT(nasal) (32) é superior ao de AGREE(place)VN# (28), mas a gama de valores dessas restrições segue se sobrepondo, o que faz emergir um segundo candidato ótimo, com a nasal velar (como será explicado na sequência). A Figura 5, a seguir, ilustra a área em que os pontos de seleção das duas restrições podem situar-se, variavelmente, além dos valores de ranqueamento e seleção.

Figura 5 - Valores de ranqueamento e desarmonia para AGREE(place)VN# e IDENT(nasal) na *GLA2*



Fonte: a autora (2015).

O aumento do valor de ponto de seleção de IDENT (a promoção de Identidade) faz com que a gramática do aprendiz mova-se em direção à gramática-alvo (Inglês-L2), já que, nesse momento de avaliação, respeitar o traço [+velar] da nasal faz com que a forma de *output* seja fiel à forma subjacente. A produção da nasal velar, em casos como esse, diz respeito ao esforço articulatorio do aprendiz na tentativa de adquirir um fonema da língua alvo que não se manifesta por assimilação do traço [+dorsal], como acontece na L1. A dificuldade reside, justamente, em associar um ponto de articulação anterior da vogal alta a um ponto de articulação posterior da nasal velar.

Os pontos de seleção para *CODA(stop) e DEP alteraram-se minimamente (100 para 99 e 41 para 38), de modo que seu afastamento das demais restrições as conservam numa posição estrita no *ranking*. Isso significa que, em seu processo de aquisição da nasal em coda silábica em Inglês, o aprendiz rejeita a produção da oclusiva, seja por especificações de formação silábica do PB, seja por fidelidade ao *input*. Assim como na gramática GLA1, a segunda avaliação do algoritmo mostra que a distância entre a desarmonia de DEP e AGREE conserva-se superior a 10 pontos na escala, deixando a restrição de fidelidade DEP fora do escopo da variação.

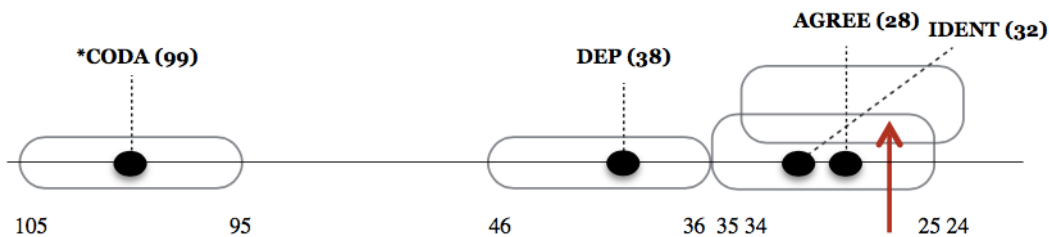
Nesta segunda avaliação, a gramática da Interlíngua continua coibindo a emergência de *outputs* com a oclusiva e também com a oclusiva seguida de epêntese, devido aos altos valores de seleção para *CODA e DEP (99 e 38). O *output* escolhido como ótimo é aquele com a nasal velar ([ˈlukɪŋ]), já que IDENT apresenta ponto de seleção superior a AGREE (32>28). A escolha pelo candidato fiel, neste momento de fala, é expressa pelo *Tableau 6*:

Tableau 6 - Gramática GLA2

lukɪŋ	*CODAstop	DEP	IDENTnasal	AGREEplaceVN
lukɪŋ				*
lukɪŋ			*!	
lukiŋg	*!	*		
lukiŋgi		*!*		

Fonte: a autora (2015).

Quando o aprendiz realiza a nasal velar, variavelmente, a gramática da Interlíngua configura-se assim:

Figura 6 - Representação da gramática variável *GLA2*

Fonte: a autora (2015).

A Figura 6 reitera a dominância plena de *CODA sobre as demais restrições do ordenamento, assim como DEP que, apesar de próxima de AGREE, não apresenta a diferença mínima para um alcance comum das áreas de abrangência representativas da variação. A alternância entre as nasais palatal e velar em coda silábica final na fala dos aprendizes de Inglês é confirmada pela flutuação de valores entre as restrições AGREE e IDENT, sendo que, num segundo momento de fala, o falante respeita a identidade de traço [+velar] da nasal, produzindo a nasal velar.

Desse modo, comprova-se que a variação é provocada pela sobreposição da faixa de valores de duas restrições, uma de Fidelidade e outra de Marcação: IDENT_(nasal) e AGREE_{(place)VN#}. Enquanto que na primeira avaliação escolheu-se como ótimo um candidato (palatal) devido ao maior valor de ponto de seleção de AGREE, seu oponente variável (velar) é promovido pelo aumento do valor de seleção de IDENT, restrição que favorece o candidato fiel ao *input*.

A variação encontrada entre as nasais velar e palatal torna evidente a característica *error-driven* do algoritmo, que promove e demove restrições de modo a aproximar-se paulatinamente do *input*. A emergência de uma ou outra nasal pode ser generalizada na colocação de Boersma e Levelt (2004, p.04):

(...) quando o aprendiz detecta uma disparidade entre o output de sua gramática e o output da gramática alvo, ele “abaixará” o ranking de uma ou mais restrições a favor de seu próprio output, assim como pode também elevar os rankings de algumas ou de todas as restrições que favorecem o alvo. Em estágios iniciais de $M \gg F$, isso significa que se a forma alvo contém uma estrutura marcada que é proibida por uma restrição de marcação altamente ranqueada na gramática do aprendiz, esta restrição de marcação tende a ser demovida, e a restrição de fidelidade correspondente, promovida; isso vai, mais cedo ou mais tarde, levar a rankings de configuração $F \gg M$.

No modelamento da gramática da L1 (seção 4.1), estágio inicial da gramática da Interlíngua, viu-se que a restrição de fidelidade MAX é relevante ao ordenamento proposto, proibindo o apagamento de segmentos pela manutenção de fidelidade máxima ao *input*. Na formalização da Interlíngua, tal qual proposta na presente seção, MAX é irrelevante ao ordenamento, pois não foram previstos candidatos que cancelam segmentos no processo de aquisição da nasal velar, dada a forma de *input* com a nasal velar. O aprendiz tem duas escolhas: (i) realiza uma nasal, sabendo que a oclusiva final de {ing} não é pronunciada, colocando em jogo as restrições IDENT e AGREE; (ii) produz a oclusiva final de {ing} seguida ou não de epêntese¹⁶, colocando em competição as restrições DEP e *CODA. O possível apagamento da nasal em coda e nasalização da vogal precedente, tal qual atestado na fonologia do PB, não acontece na Interlíngua.

Assumindo-se as restrições como descrições estruturais que dão conta da boa formação das realizações linguísticas e sendo elas universais (todas estão presentes em todos os sistemas linguísticos), pode-se pensar que, se fossem considerados um maior número de candidatos na gramática em aquisição, uma restrição como MAX seria relevante ao ranqueamento. Tendo em vista as formalizações propostas da L1 e da Interlíngua, o processo de aquisição da nasal velar em L2 pode ser representado pelos seguintes estágios:

Quadro 2 - Estágios de aquisição da nasal velar em Inglês-L2

L1	→ Estágio 1: *CODA _(stop) , AGREE _(place) VN#>>DEP>>MAX>>... [IDENT _(nasal)]
Interlíngua	→ Estágio 2: *CODA _(stop) >>DEP>>AGREE _(place) VN#>>IDENT _(nasal) >>... [MAX]
	→ Estágio 3: *CODA _(stop) >>DEP>>IDENT _(nasal) >>AGREE _(place) VN#>>... [MAX]
L2	→ Estágio 4: IDENT _(nasal) >>DEP>>AGREE _(place) VN#>>*CODA _(stop) >>... [MAX]

Fonte: a autora (2015).

¹⁶ Se as frequências de realização da nasal velar seguida da oclusiva e da oclusiva epentetizada tivessem sido controladas, a gramática seria capaz de modelar a seleção de tais candidatos pelo *ranking*, com a demissão dos valores de DEP e *CODA no ordenamento. Uma análise acústica minuciosa demonstraria, provavelmente, a prevalência de dados com epêntese, o que, em termos estocásticos, comprovaria a dominância estrita de *CODA sobre as demais restrições da gramática da Interlíngua.

Como se observa no Quadro 2, tanto a L1 quanto a Interlíngua são gramáticas dominadas por Marcação (*CODA(stop)). As restrições marcadas em negrito nos estágios 2 e 3 (AGREE(place)VN# e IDENT(nasal)), que representam gramáticas em desenvolvimento, são as responsáveis pela variação na aprendizagem. O estágio 4, que simularia uma aquisição “plena” da nasal velar, corresponde à gramática mais próxima a de um falante nativo de Inglês, dominada por Fidelidade (IDENT(nasal)). A restrição de Fidelidade MAX, tanto na Interlíngua quanto na L2, está grafada em cor cinza para demonstrar a possibilidade de seu provável efeito nas gramáticas se mais candidatos a *output* fossem considerados, assim como a restrição IDENT(nasal) no *ranking* da L1.

5. Considerações finais

Este estudo utilizou um programa algorítmico de aprendizagem gradual para representar a aquisição fonológica variável da nasal velar em Inglês-L2 por falantes gaúchos de Português. O *GLA* executou com eficiência a tarefa de calcular o desenvolvimento e a variação da fonologia de L2 com base em dados advindos da produção linguística dos aprendizes. Por ora, as afirmações podem ser feitas com base no estudo realizado, a partir dos preceitos teóricos assumidos, são as seguintes:

- i. línguas são um *continuum* de restrições universais que, sob diferentes ordenamentos, definem sistemas linguísticos. A aquisição é um processo gradual e não linear, em que o aprendiz utiliza restrições de Marcação da L1, que serão dominantes em etapas iniciais da aprendizagem, e restrições de Fidelidade da L2, que serão inicialmente dominadas para estruturar a língua em aquisição (Interlíngua);
- ii. a Interlíngua é, de fato, um sistema linguístico estruturado, em que operam restrições fonológicas da L1 e da língua alvo variavelmente;
- iii. a variação é uma característica inerente às línguas naturais e foi comprovada tanto pela ARV realizada (embora tal análise não tenha sido o foco deste artigo) quanto pela análise pelo *GLA*, atribuindo o *status* de “língua natural” ao sistema interlinguístico;
- iv. a primeira língua (L1), em uma perspectiva linguística formal, deixa de ser vista como uma “intervenção negativa” na aquisição de L2, já que é o ponto de partida para a construção do sistema alvo;
- v. a inversão do *ranking* M>>F para F>>M é indicativa de que a aquisição aproxima-se paulatinamente de um estágio mais estável na L2, próximo ao de uma aquisição “plena” da língua alvo. Adquirir a nasal velar em Inglês-L2 é promover Fidelidade em direção ao *input*;
- vi. por fim, a TO Estocástica, através do *GLA*, constitui-se como uma teoria gramatical eficaz para representar o processamento linguístico variável de aprendizes de uma segunda língua.

Referências bibliográficas

ADAM, G. **From variable to optimal grammar**: Evidence from language acquisition and language change. PhD dissertation. Tel Aviv: Tel-Aviv University. 2003.

ALVES, U. K. **A Aquisição das Sequências Finais de Obstruintes do Inglês (L2) por Falantes do Sul do Brasil**: análise via Teoria da Otimidade. 2008. 337 f. Tese (Doutorado em Linguística). Faculdade de Letras - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2008. Disponível em: <http://repositorio.pucrs.br:8080/dspace/bitstream/10923/4091/1/000399693-Texto%2BCompleto-0.pdf>, acesso: 09-12-14.

_____. Os Dados de Aquisição de L2 e sua Relação com a Teoria Fonológica: reflexão a partir de uma análise via TO. In: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM-DE-PAULA, M. R. (Org). **Estudos em Aquisição Fonológica**, Santa Maria, v. 2, p. 165-180, 2009.

_____. Teoria da Otimidade, Gramática Harmônica e Restrições Conjuntas. **Alfa**, São Paulo, v. 54, n.1, p. 237-263, 2010. Disponível em: <http://seer.fclar.unesp.br/alfa/article/view/2879>, acesso: 10-12-14.

_____. Aquisição Fonológica de L2: formalização de fenômenos variáveis na língua-fonte, na língua-alvo e em seus sistemas intermediários. In: COLLISCHONN, G.; BISOL, L. **Fonologia**. Teorias e perspectivas. POA: EDIPUCRS, 2013. p. 133-148.

_____; CABAÑERO, M. B. A transferência grafo-fono-fonológica na produção de sequências ortográficas ‘ng’ do Inglês (L2): uma abordagem conexionista. **ReVEL**. Pelotas, v.6, n. 11, ago./2008. Disponível em: http://www.revel.inf.br/files/artigos/revel_11_a_transferencia_grafo_fonico_fonologica_na_producao_de_sequencias_ortograficas.pdf, acesso: 08-03-15.

ALVES, F.C.; LUCENA, R.M. Aquisição da lateral silábica do inglês: uma análise via teoria da otimidade estocástica. **Letrônica**, Porto Alegre, v. 7, n. 2, p. 795-820, jul./dez., 2014.

AZEVEDO, R. Q. **A epêntese no português brasileiro (L2), em segmentos plosivos em codas mediais, por falantes nativos do espanhol colombiano (L1)**: uma análise via Teoria da Otimidade Estocástica e Gramática Harmônica. 2011. 180 f. Dissertação (Mestrado em Linguística Aplicada) – Universidade Católica de Pelotas, Pelotas, 2011.

BAPTISTA, B. O.; SILVA-FILHO, J. L. A. The influence of voicing and sonority relationships on the production of English final consonants. In: BAPTISTA, B. O. WATKINS, M. A. **English with a Latin beat**: Studies in Portuguese/Spanish-English Interphonology. John Benjamins, 2006, p. 73-90.

BATTISTI, E. A representação da nasal em coda silábica e os ditongos nasais do português. In: **Anais do XVII Congresso Internacional Asociación de Lingüística y Filología de América Latina (ALFAL 2014)**. João Pessoa - Paraíba, Brasil. 2014. p. 1433-1453. Disponível em: <http://www.mundoalfal.org/CDAnaisXVII/trabalhos/R0819-1.pdf>, acesso: 02-07-15.

BLEY-VROMAN, R. What is the logical problem of foreign language acquisition? In: GASS, S.M.; SCHACHTER, J. **Linguistic Perspectives on Second Language Acquisition**. Cambridge: CUP, 1989.

BOERSMA, P.; HAYES, B. **Empirical Tests of the Gradual Learning Algorithm**. 2001. Disponível em: <http://www.linguistics.ucla.edu/people/hayes/GLA/gla.pdf>, acesso: 02-07-15.

_____; WEENINCK, D. **Praat – Doing Phonetics by Computer**. 2015. Disponível em: http://www.fon.hum.uva.nl/praat/download_win.html, acesso: 02-07-15.

_____; LEVELT, C. Optimality Theory and phonological acquisition. **Annual Review of Language Acquisition** 3 (2004), 1–50.

BROSELOW, E. Unmarked structures and emergent rankings in second language phonology. In: **International Journal of Bilingualism**, v.8, n.1, 2004. p. 51-65.

CAGLIARI, L. C. **An experimental study of nasality with particular reference to Brazilian Portuguese**. 1977. 321 f. Tese (Doutorado em Filosofia) - Universidade de Edinburgo, Edinburgo, 1977.

CÂMARA JR., J. M. **Estrutura da língua portuguesa**. 8.ed. Petrópolis: Vozes, [1970] 2005.

_____. **Problemas de lingüística descritiva**. 11.ed. Petrópolis: Vozes, [1971] 2002.

CURTIN, S.; ZURAW, K. Explaining constraint demotion in a developing system. In SKARABELA, B.; FISH, S.; and A.H.-J. do (Eds.). **Proceedings of the 26th annual Boston University Conference on Language Development** (p. 118-129). Somerville, MA: Cascadilla Press. 2002.

GARCIA, G. D. **Aquisição de acento primário em inglês por falantes de português: uma análise de derivações com sufixos não neutros via Algoritmo de Aprendizagem Gradual - GLA**. 2012. 148 f. Dissertação (Mestrado em Teoria e Análise Linguística) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.

GNANADESIKAN, A. Markedness and faithfulness constraints in child phonology. In KAGER, R.; PATER, J.; & ZONNEVELD, W. (Eds.). **Fixing priorities: constraints in phonological acquisition**. Cambridge: Cambridge University Press. [ROA 67, 1995]

GUIMARÃES, M. A. A. **Aspectos da fonologia do português como segunda língua por aprendizes anglófonos – uma análise via Teoria da Otimidade**. 2012. 115 f. Dissertação (Mestrado em Letras - Filologia e Língua Portuguesa) - Universidade de São Paulo, São Paulo, 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/8/8142/tde-22022013-153330/en.php>, acesso: 02-07-15.

KOERICH, R. D. **Perception and Production of Vowel Epenthesis in Word-Final Single Consonant Codas**. 261 f. Tese: Doutorado em Letras. Universidade Federal de Santa Catarina. Florianópolis, 2002.

LABOV, W. **Principles of Linguistic Change**. Vol. 1: Internal Factors (Language in Society). Philadelphia: Blackwell Publishers, 1994.

_____. **Principles of Linguistic Change**. Vol. 2: Social Factors (Language in Society). Philadelphia: Blackwell Publishers, 2001.

LACY, P. de. **The formal expression of markedness**. Dissertation on Linguistics. University of Massachusetts Amherst, Massachusetts, 2002. Disponível em: <http://www.pauldelacy.net/webpage/docs/delacy-2002-formal%20expression%20of%20markedness.pdf>, acesso: 28-09-2015.

LEVELT, C. **Unfaithful kids**: place of articulation patterns in early child language. Paper presented at the University of Pennsylvania, October 13, 1995.

_____; SCHILLER, N.; LEVELT, W. The acquisition of syllable types. **Language Acquisition**, 8.3, 237-264. 2000.

LUCENA, R. M.; ALVES, F. C. Análise Variacionista da Aquisição do /p/ em Coda Silábica por Aprendizizes de Inglês como LE. **Revista Intertexto**, v. 5, n. 2, 2012. Disponível em: <http://www.uftm.edu.br/revistaeletronica/index.php/intertexto/article/view/310>, acesso: 09-12-14.

MCCARTHY, J. J. (ed) **Optimality Theory in Phonology**. Blackwell Publishing, 2004.

MOHANAN, P. J. **Evidence of transference and Emergence in Interlanguage**. 2001. Disponível em: <http://roa.rutgers.edu/files/444-0701/444-0701-MONAHAN-0-0.PDF.gz>, acesso: 02-07-15.

MOHANAN, T.; MOHANAN, K. P. **Towards a Theory of Constraints in OT**: Emergence of the not-so-unmarked in Malayalee English. 2003. Disponível em: <http://roa.rutgers.edu/files/601-0503/601-0503-MOHANAN-0-0.PDF>, acesso: 28-09-2015.

PATER, J.; WERLE, A. Typology and variation in child consonant harmony. In: FÉRY, C.; DUBACH GREEN, A.; VAN DE VIJVER, R. (Eds.). **Proceedings of the 5th HIL Phonology Conference** (pp. 119-139). Potsdam: University of Potsdam. 2001.

PRINCE, A.; SMOLENSKY, P. Optimality Theory: constraint interaction in generative grammar. 1993. In: MCCARTHY, J. **Optimality Theory in Phonology**. Oxford: Blackwell Publishing, 2004.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. **Goldvarb X**: a variable rule application for Macintosh and Windows. Department of Linguistics. University of Toronto, 2014.

SCHMITT, B.; ALVES, U. K. The acquisition of /p/ and /k/ word-mid codas of English (L2) by learners from Southern Brazil (L1): a gestural analysis in stochastic optimality theory. **Letrônica**, Porto Alegre, v. 7, n. 2, p. 765-794, jul./dez., 2014.

SELKIRK, E. The Syllable. In: HULST; SMITH. (eds.). **The Structure of Phonological Representations** (Part II). Dordrecht Foris. p. 337-383. 1982.

SELINKER, L. Language Transfer. **General Linguistics**, n.9 (2), p.67-92, 1969.

_____. **Rediscovering Interlanguage**. NY: Longman, 1994.

WEINREICH, U; LABOV, W.; HERZOG, M. I. **Fundamentos Empíricos para uma Teoria da Mudança Linguística**. Trad. de Marcos Bagno. São Paulo: Parábola Editorial, [1968] 2006.

ZIMMER, M. C.; ALVES, U. K. **A produção de aspectos fonéticos/fonológicos da L2: instrução explícita e conexãoismo**. Revista Linguagem e Ensino, Pelotas, v. 9, n. 2, p. 101-143, 2006.

Artigo recebido em: 10.10.2015

Artigo aprovado em: 09.04.2016

Nuclear stress placement by Brazilian users of English as an international language

A alocação do acento nuclear por falantes brasileiros de inglês como língua internacional

Leonice Passarella dos Reis*
Rosane Silveira**

RESUMO: There have been some studies on the pronunciation features that are important to guarantee speech intelligibility by users of English as an international language (IL). In light of the importance of nuclear stress placement for successful communication between speakers in the international community (JENKINS, 2000), the present article reports the results of a pilot study that investigated the way nuclear stress is placed by four Brazilian Portuguese (BP) intermediate users of English when interacting with other BP users of English (as listeners). Participants met in pairs and engaged in a controlled pairwork oral task, which yielded the production of 160 audio-recorded utterances. With the use of *Praat*, perceptual and acoustic analysis of the dataset were performed in order to examine if nuclear stress was placed as expected according to the discursive contexts set. The analysis revealed that speakers had difficulties in placing the expected nuclear stress at sentence initial, medial, and, surprisingly, final position. Additionally, it was found that for the participants in the present study, signaling both corrective information and information being elicited by means of nuclear stress placement was challenging. This difficulty in placing nuclear stress may compromise the way these speakers' intent is interpreted when holding interactions in English.

ABSTRACT: Pesquisadores têm se dedicado a investigar quais aspectos da pronúncia são importantes para garantir um discurso inteligível por parte de falantes de inglês como língua internacional. Tendo em vista a importância do acento nuclear para garantir uma comunicação de sucesso entre os falantes da comunidade internacional (JENKINS, 2000), este artigo reporta os resultados de um estudo piloto que objetivou investigar o modo como o acento nuclear é produzido por quatro falantes brasileiros de inglês em nível intermediário de proficiência durante interações com outros falantes brasileiros de inglês (ouvintes). Os participantes se encontram em duplas e se engajaram em uma atividade oral controlada, que propiciou a produção de 160 assertivas gravadas em áudio. Com o uso do *Praat*, foi feita uma análise acústica e auditiva dos dados a fim de verificar se o acento nuclear fora alocado no local esperado, de acordo com os contextos discursivos estabelecidos. A análise revelou que os falantes tiveram dificuldades em colocar o acento nuclear na posição esperada, tanto no início, no meio e, surpreendentemente, no final das assertivas. Ainda, os dados demonstraram que, para os participantes deste estudo, sinalizar informações corretivas e elicitadas através do acento nuclear foi desafiador. Essa dificuldade na alocação esperada do acento nuclear pode comprometer o modo como esses falantes são interpretados em suas interações em inglês.

* Doutoranda no Programa de Pós-Graduação em Inglês da Universidade Federal de Santa Catarina (UFSC) e professora na Escola de Aprendizagem-Marinheiros de Santa Catarina (EAMSC) Florianópolis, Santa Catarina, Brasil (leonicepr@gmail.com).

** Professora Adjunta da Universidade Federal de Santa Catarina (UFSC) Florianópolis, Santa Catarina, Brasil. Atua no Departamento de Língua e Literatura Estrangeiras e na Pós-Graduação em Inglês. Bolsista de Produtividade em Pesquisa do CNPq - Nível 2 (rosanesilveira@hotmail.com).

PALAVRAS-CHAVE: International Language. Nuclear stress. Communication. Pronunciation.

KEYWORDS: Língua Internacional. Acento nuclear. Comunicação. Pronúncia.

1. Introduction

I know you think you understand what you thought I said but I'm not sure you realize that what you heard is not what I meant.
(Alan Greenspan)

Accent¹ is ubiquitous, present in every language irrespective of its status as being first (L1), second (L2), or foreign (FL), and may interfere with the choices people make in their interactions. The communication process itself is fragile. Everyone may already have gone through communication misunderstandings even in their L1, given that listener's expectations may filter the speech uttered and may apply to it an unintended meaning, that is, meaning lies in the ear of the beholder as much as in the mouth of the speaker. One may choose to communicate an idea by using specific linguistic structures based on past experiences, but the hearer listens to what someone says using the hearer's own past experiences, both hearer and speaker currently undergoing a unique situation (that might resemble others, but is unique) (BECKNER *et al.*, 2009).

Listeners from distinct L1 backgrounds experience greater frangibility in communication, given that they have to deal with foreign accents, which vary according to speakers' L1, among other individual and sociocultural factors. Foreign accents may have a great deal of undesirable consequences for international language² (IL) speakers given that they may (1) make IL speakers' speech difficult to understand, (2) cause listeners to misjudge an IL speaker affective state, and (3) cause negative personal evaluations (FLEGE, 1995; MUNRO; BOHN, 2007; DERWING; MUNRO, 2005; MOYER, 2013).

Having in mind that the current goal of IL pedagogy in terms of pronunciation is improving communication (DERWING; MUNRO, 2005), research on IL intelligibility for IL pedagogy is strikingly relevant since it sets out to establish those aspects of IL speech that affect

¹ Accent refers to the pronunciation of speech sounds (segments) and suprasegmental features (intonation, rhythm, pitch, segmental length, tempo, loudness) of a given language variety (Moyer, 2013).

² English as an International Language in the present study is an umbrella term that permeates three notions: (1) that of number (additional, second, third, and so forth in opposition to first language); (2) that of use (international – not favoring a specific English as an L1 variety; and (3) that of acquisition (interlanguage: it is shaped according to the L1 attractors and language use).

IL speakers' success in communication. Jenkins has conducted research in order to inform pedagogy on the aspects of pronunciation that would hinder communication among speakers of English from different L1 backgrounds, putting together the Lingua Franca Core (LFC), which she believes to subsume the phonetic and phonological features for successful communication in English worldwide (JENKINS, 2002). The LFC includes mainly (1) consonantal segments and their allophones (expect /θ/, /ð/), (2) phonetic details such as aspiration, voicing, and duration, (3) appropriate consonant cluster simplification, and (4) vowel contrasts. Processes in connected speech (e.g., assimilation and elision) and rhythm-related issues (e.g., such as lexical stress and tunes) are disregarded because they are considered either irrelevant or unteachable. However, some importance is given to nuclear stress and thought groups in the level of utterances (JENKINS, 2000). Research has questioned her choices for two main reasons. Firstly, Derwing (2008) states that the LFC is based on "a small sample of communication breakdowns across very few learners", and thus lacks evidence for international use. Secondly, relevant features for successful communication are L1 dependent and should be considered in a specific community of speakers (BERNS, 2008). For certain L1 backgrounds (e.g., Brazilian Portuguese), some features left aside in the Lingua Franca Core are relevant both in segmental (e.g., SCHADECH; SILVEIRA, 2013) and suprasegmental terms (e.g., PASSARELLA-REIS; GONÇALVES; SILVEIRA, in press).

The results reported here are from a pilot study investigating the intelligibility and interpretability of BP users of English, by looking into lexical stress and nuclear stress placement. This article reports the results of one of the aspects seen as important in the LFC, and which, to the best of our knowledge, has not been formally investigated in Brazil, namely, the allocation of nuclear stress³ by BP users of English⁴. Research has shown that the placement of a nuclear stress on an unexpected portion of the utterance frequently hinders communication (e.g., Jenkins, 1997). The following paragraphs review some studies dealing with intelligibility, comprehensibility and/or interpretability⁵ associated with nuclear stress placement.

³ Nuclear stress refers to the most prominent material in an utterance, and thus, holding the most important piece of information.

⁴ The results reported here are from a pilot study investigating the intelligibility and interpretability of BP users of English, by looking into both lexical stress and nuclear stress placement.

⁵ In the present study, interpretability refers to the grasp of speaker's intention and embraces the concept of intelligibility, related to word recognition, and that of comprehensibility, related to the meaning of the word.

Tiffen (1974) investigated the intelligibility of educated Nigerian speakers of IL to British IL-L1 listeners at both segmental and suprasegmental levels. For the latter, word stress, nuclear stress, and intonational patterns (tunes) were analyzed. For the nuclear stress investigation, speakers had to stress different portions of a sentence according to the stimuli given by the researcher via a context question as illustrated in (1).

(1)

Interviewer: Did BILL motor to London?

Speaker: No, JOHN motored to London. (i.e., not BILL)

Interviewer: Did John CYCLE to London?

Speaker: No, John MOTORED to London. (i.e., not CYCLED)

Interviewer: Did John motor to MANCHESTER?

Speaker: No, John motored to LONDON. (i.e., not MANCHESTER)

Production results showed that all Nigerian participants had difficulty with this feature of English pronunciation (M = 40.4%), placing nuclear stress in the rightmost portion of the utterance, irrespective of the contrast being made. Some of the productions were randomly selected to compose the listening task taken by the British IL-L1 users. Due to the unexpected productions, listeners failed to interpret which contrast was being made, leading to unsuccessful interpretations.

Lanham (1984) investigated the consequences of misallocation of stress at both lexical and sentence levels by a South African Black English speaker. Listeners were 13 L1 (white) users of English and two (black⁶) South African users of English. Participants listened to the recorded passage and answered a comprehension quiz assessing the comprehension of the passage read by the South African. Results for nuclear stress placement showed that the misallocation of nuclear stress (in order to establish focus on new information which was being elicited) hindered syntactic coherence and thus posed difficulties for the listener in “making sense” of the message being conveyed.

In line with Tiffen (1974), Atechi (2004) investigated mutual intelligibility of Cameroon English speakers with British and American speakers at both segmental and suprasegmental levels. For the latter, both word stress and nuclear stress were examined. The nuclear stress

⁶ Being “black” or “white” was important information highlighted by that scholar.

procedures and instruments resemble the ones applied in Tiffen's (1974) study. It differs only in that stimuli were produced by the three groups of speakers, and heard by listeners from the same three languages, viz. Cameroon, British, and American. Production results showed that Cameroon speakers failed to place the nuclear stress in order to highlight the contrast intended, and, accordingly, listeners (both British and American) failed to interpret the contrasts.

In her study, Jenkins (1997) investigated the production and perception of nuclear stress by IL-L1 and IL users of English. In the production task, speakers were supposed to produce one of four sets of questions by reading them. Each question had multiple endings, such as the one illustrated in (2).

- (2)
- (A) Did you buy a tennis racket at the sports center this morning, or
 - (B) was it a squash racket?
 - (C) did you buy it yesterday?
 - (D) did you only borrow one?
 - (E) was it your girlfriend who bought it?
 - (F) at the tennis club?

For the interpretation task, the second halves of the questions were removed and the first halves were played to IL-L1 and IL users of English. The list of options for the second halves was made available to the listeners, who were also the speakers. The listeners had to listen and predict the second half in each case. Results showed that IL-L1 users both produced and interpreted well all the second halves to the IL-L1 questions, while the IL users interpreted well about two-thirds of the IL-L1 speakers' second halves. The IL speakers misplaced nuclear stress in most of their first halves. Consequently, both IL-L1 and IL users misinterpreted the intended meaning in the second halves produced by the IL speakers. Her results support her hypothesis that IL learners acquire nuclear stress receptively faster than productively. This finding corroborates the need for overtly teaching nuclear stress placement in IL classes.

Zoghbor (2010) investigated the effectiveness of a pronunciation syllabus based on the LFC in improving the intelligibility and comprehensibility of Arab learners. Participants were divided into two groups: experimental (receiving the LFC pronunciation syllabus) and control (receiving traditional pronunciation syllabus). The experimental group gain scores were higher than were those of the control group, but differences did not reach statistical significance. As to the placement of nuclear stress, the scholar found that it is narrowly related to the comprehension of the message rather than to the recognition of words and that it is important

not only to facilitate comprehension and intelligibility, but also to trigger a positive judgment over a speaker's speech. Additionally, speakers allocating nuclear stress as expected were seen as more connected to the message being conveyed, and thus as more interesting to be listened.

In Brazil, intelligibility and comprehensibility⁷ have been investigated to inform pedagogy mostly at the segmental level (e.g., BECKER, 2011; 2013; CRUZ, 2003, 2004, 2008a; 2008b; SCHADECH, 2013; SCHADECH; SILVEIRA, 2013). A few other studies have investigated the effect of non-target production of suprasegments on the intelligibility of BP users of English (GOMES, BRAWERMAN-ALBINI; ENGELBERT, 2014; MARTA, 2011; PASSARELLA-REIS; GONÇALVES; SILVEIRA, in press)⁸, but, to the best of our knowledge, research on the nuclear stress associated to intelligibility, comprehensibility, and/or interpretability of BP-IL users' speech is inexistent. The present article reports on a study that seeks to promote some investigation in this area.

2. Nuclear Stress and unexpected allocation

Nuclear stress has received many names: Nuclear or tonic stress (JENKINS, 2000), prominence (CELCE-MURCIA; BRINTON; GOODWIN, 1996), phrasal stress (PIERREHUMBERT; HIRSHBERG, 1992), accent (SLUIJTER; van HEUVEN, 1996), and nuclear accent (ORTIZ-LIRA, 1998). These terms are used to refer to the placement of more prominence on a specific syllable/word/phrase in a thought group⁹, in order to highlight it and convey (1) meaning, (2) the context in which the utterance is placed, and (3) the speaker's intent (CELCE-MURCIA; BRINTON; GOODWIN, 1996). Nuclear stress, thus, has great communicative value and, as evidenced from the studies reviewed in the introduction, can mislead the way speakers are understood.

⁷ In the studies cited in this page, intelligibility refers to the extent to which listeners understand the intended message by transcribing the words. Comprehensibility refers to listeners' rating of difficulty in understanding utterances/words (MUNRO, DERWING & MORTON, 2006).

⁸ Gomes, Brawerman-Albini, and Engelbert (2014), besides investigating the intelligibility of BP-IL users when producing words ending in -ed, also investigated their intelligibility when producing suffixed words that are stressed on the fourth syllable (from the right to the left). Marta (2011) investigated BP-IL users' production and perception of intonation contours of statements indicating both surprise and disbelief and negative interrogative questions indicating both surprise and request for confirmation. Passarella-Reis, Gonçalves, and Silveira (in press) investigated the perception of three BP-IL-intonational patterns of yes-no questions and their interpretation regarding the intent of the speakers.

⁹ Also called in the literature as 'tone units', 'sense groups', 'tone groups', 'breath groups', 'intonation phrase' (PIERREHUMBERT, 1980). We chose to use the term 'thought group' because it seems to have more relation with the organization of information while 'intonation phrase' for instance, seems to bring into our minds the concepts of intonation and tunes (high and low).

The portion of the sentence where the nuclear stress falls greatly depends on the context and the intention of the speakers. In general, it occurs near the end of a thought group; however, any syllable in the thought group can be nuclear-stressed in order to express focus by placing a pitch movement on it (SLUIJTER & van HEUVEN, 1995). The rules were set long ago, “When no expressive stress disturbs a sequence of heavy stresses, the last heavy stress in an intonational unit [thought group] receives the nuclear heavy stress” (NEWMAN, 1946, p. 176 as cited in CHOMSKY & HALLE, 1968, p. 90) or as Liberman (1972) would say it, “put the strong element on the right in any given metrical constituent, if you have no good reason to do otherwise” (p. 244). A reason to do it otherwise would be the one investigated in the present study: highlighting information being contrasted or elicited, which is related to the use of focus.

General agreements about nuclear stress are that (1) nuclear stress signal focus (broad or narrow), (2) not all focused constituents need to take stress, and (3) unfocused constituents in a thought group do not take a nuclear stress (ORTIZ-LIRA, 1998). There are two types of focus: narrow (marked) and broad (unmarked). In the latter (broad), the utterance contains all-new information, such as when answering the question in (3).

- (3)
Speaker: What’s the matter?
Interlocutor: JOHN has MOVED to CANada.

The nuclear stress in the answer in (3) falls in the rightmost sentence stress available (Canada), given that all information provided here is regarded as new¹⁰. Such a question was probably triggered by the sadness in the interlocutors’ eyes, which is explained by the fact that ‘John has moved to Canada’, an unknown piece of information to the Speaker. The narrow focus, in turn, contains both given and elicited information, as illustrated in (4).

- (4)
Speaker: Who has moved to Canada?
Interlocutor: JOHN has MOVED to CANada.

¹⁰ The term *new* has been discussed in the literature. Lambrecht (1998) claims that if the referent of the so-called new information is activated in the addressee’s mind, then it is not new at all. We take a different stand and relate the term *new* here to the ignorance of a fact. It is not *new* in the sense that the addressee does not have it as a possibility in his mind, as claimed by Lambrecht (1998), but rather as the information missing in a given pragmatic context. Along our study, *new* will be referred to as *elicited*.

In (4), ‘John’ is being elicited while ‘has moved to Canada’ is shared information (ORTIZ-LIRA, 1998). In this context, the Speaker has probably heard that someone moved to Canada and he aims to know who the person is. In the answer, the focus is narrowed down to ‘John’ only. This article investigates the allocation of nuclear stress in assertions with narrow focus.

The principles guiding nuclear stress placement in English and BP are alike. The main difference lies on the fact that in BP, when in broad focus, nuclear stress is placed in the last word of a thought group, irrespective of being a content word or a function word. To illustrate it, consider an often-cited example provided by Baptista (2001), in (5).

- (5)
DÊ o LIVRO para MIM. (unmarked broad focus)
GIVE the BOOK to ME. (marked narrow focus)

In BP, the nuclear stress in ‘mim’ is interpreted as non-contrastive information. It is a request that can perfectly be uttered in a room occupied by both speaker and interlocutor only or by the two accompanied with more people. In English, however, if the location is directly transferred from the speaker’s L1 to the IL, it signals contrasting information. If nobody else were in the room, the interlocutor would probably wonder the reasons why such a request was being made, and if there were somebody else, the interlocutor would go suspicious that something else was going on, and thus misinterpret the speaker’s intent.

Jenkins (2000) describes an interaction among four students from different L1s (Brazilian, Swiss-French, Colombian, and Hungarian), while making posters for the classroom wall. The Hungarian student asks the other three the question in (6).

- (6)
Have you got a blue VUN?

The scholar reports that the three other interactants echoed the words ‘blue vun’ and ‘vun’ many times and got the intended meaning only after the Hungarian student, holding up a blue pen, explained ‘Blue vun like THIS’. The intended meaning for ‘vun’ was ‘one’. The author highlights that, although interactants were acquainted with each other’s accents in English and had enough contextual cues in order to provide the listeners with clues to meaning (they were making posters and surrounded with paper and colored pens), the misallocation of

nuclear stress was a great villain to the lack of success in communication in that interaction. Jenkins argues that the mispronunciation of 'one' would have caused no problems if the nuclear stress had been placed in 'blue' (i.e., as opposed to 'red'). Once it was placed in 'vun', no contrast was possible, and it signaled that 'vun' carried the most important information in the thought group, misleading the interpretation of meaning.

The inability to segment speech into meaningful thought groups is one of the most common factors contributing to problems with nuclear stress placement (JENKINS, 2000, p. 156). Thought groups (1) are set off by pauses before and after¹¹, (2) contain one prominent element (nuclear stress), (3) have an intonation contour of their own, and (4) have a grammatically coherent internal structure (CELCE-MURCIA *et. al.*, 1996, p. 175). Celce-Murcia *et al.* (1996) explain that the number of nuclear stress in a given utterance depends on the speaker. The more pauses the speaker produces, the more thought groups the speaker creates, and thus, the more nuclear stress allocations arise. They highlight that too many nuclear stresses (due to many pauses) make the overall message difficult to process and understand.

3. Method

3.1 Participants, procedures, and instruments

Reading aloud is one of the most used methods of data collection in the history of pronunciation research due to its manifold advantages: control for pronunciation features and other elements such as vocabulary choice or grammar usage (LEVIS, 2011). However, as Levis (2011) highlights, reading aloud is a reading skill not a speaking skill, promotes different performances (better or worse than in free speech) depending on the speaker who reads/speaks and there is an agreement that individuals feel weird when reading aloud because it is not a common activity. However, our focus is on interpretability linked to pronunciation rather than overall interpretability (linked to choice of words and grammar usage, for example) and thus the only way of trying to avoid these interfering factors is by using more controlled tasks, made possible through reading aloud activities. In this study, the interpretability assessment involved activities related to identifying speakers' intention (NELSON, 2011) based on the placement of nuclear stress while reading aloud sentences with focus on certain portions of the utterance. In

¹¹ A thought-group-final syllable lengthening associated or not with intonation extra elements (e.g., a high tone) may also indicate the limits of a thought group (PIERREHUMBERT & HIRSHBERG, 1992; PIERREHUMBERT, 1980).

the utterances tested, 22.5% had focused material in initial position (N = 36), 62.5% had focused material in medial position (N = 100), and 15% had focused material in final position (N = 24).

Participants were eight: four intermediate-level-BP-IL users (hereafter speakers), two advanced-level-BP-IL users (advanced listeners), and two BP-IL teachers (teacher listeners). The speakers were females attending English classes (level 6 – intermediate) at *Cursos Extracurriculares/UFSC*¹². They reported having an intermediate English proficiency level. Their age ranged from 18 to 52 (M = 29.5). The teacher listeners (hereafter TL) were two teachers of English and Master and PhD candidates at UFSC. They were a male and a female, and their ages at the time of data collection were 28 and 29. The advanced listeners (hereafter AL) were two male BP-IL users of English with an advanced level of proficiency¹³: (1) A graphic designer and (2) a laboratory technician who is also a master candidate at UFSC with a major in French. They were 23 and 29 years old at the time of data collection (M = 28.5).

For data collection, participants met in pairs (N = 4) only once as follows:

- 1) a TL and a speaker; and
- 2) an AL and a speaker.

The four pairs met the researcher separately at a suitable time for the participants. Each member of a pair had never met before. Prior to starting data collection, speakers performed some training with the recording equipment in a soundproof booth. Data were audio recorded by means of a *C 520 L* professional head-worn condenser microphone, connected to a hybrid audio interface called *MOTU Ultra Lite mk3*, and of an audio editor software called *OceanAudio*. Data were video recorded by using a Nikon camera full HD. Data collection followed then the steps below:

¹² *Cursos Extracurriculares* is a program at UFSC that offers language courses to faculty, students, staff, and other members of the local community.

¹³ Although Interlocutor 4 reported having a post-intermediate level of proficiency, his scores on the Oxford Proficiency Test indicated he had an advanced level. Another researcher administered this test for the purposes of her study.

- 1) Participants filled up a questionnaire to gather information on their language experience and signed a consent form¹⁴ informing the procedures for data collection and that the speaker would be audio-recorded and the interlocutor video-recorded.
- 2) They were introduced to the context of the data collection: They were both candidates for a job position in an International Company and were at the company to go through the recruiting and selection hiring process.
- 3) For their first task (*Breaking the ice*, Appendix A), they were asked to engage in a first meeting conversation in the standards each was used to and learn three things they had in common (e.g., a dislike of sushi, following TROFIMOVICH & KENNEDY, 2014). This task had the objective to break the ice and help them feel more comfortable in the presence of each other.
- 4) At this point, participants learned that the speaker was already an employee of the company's and a member of the recruiting staff (a psychologist) and that only the listener was the actual candidate for the position. They were introduced to the nuclear stress placement task individually (see Appendices B and C) and only after the task was clear to both members of the pair the task was started.
- 5) After the tasks were over, the listeners were individually interviewed in order to clarify the possible reasons for reduced interpretability identified during the recordings. During the interview, the sheet with the interlocutors' answers as well as the video-recordings were assessed. Additionally, the interview was audio-recorded for easy future retrieval.
- 6) Participants were given a gift for their participation.

The nuclear stress task was divided into two parts. In the first part, the speaker read a context question silently and then read aloud the answer to that question. The listener, in turn, chose one of the three possible questions for that answer, according to the speaker's placement of nuclear stress. For example, the speaker read aloud 'LUCY got married in 1984' in a response to 'Who got married in 1984?'. In the sheet of paper, the listener found three options: (A) 'Who got married in 1984?', (B) 'What happened to Lucy in 1984?', and (C) 'When did Lucy get married?'. If the speaker placed the nuclear stress on "Lucy" and the listener was able to notice

¹⁴ The study reported here has the Ethics Committee approval for research with human participants (protocol number 16125813.1.0000.0121).

it, then the listener would interpret that A is the question being answered. This part looked into how well speakers are able to signal information being elicited in an assertion.

In the second part, the speaker read silently a context. Next, the speaker read aloud another statement correcting one of the pieces of information present in the context statement. In turn, the listener had a sheet containing three possible contrasting ideas and chose one according to what the speaker had read. For example, the speaker read silently the context statement 'Peter has bought a red car' and then read aloud 'Peter has bought a YELLOW car'. In his/her sheet, the listener had three alternatives and was to check the information according to his interpretation: '(A) Peter did, not John.'; '(B) bought it, he didn't sell it.'; '(C) a yellow car, not a red one.' If the speaker placed the nuclear stress on 'yellow' and the listener was able to notice it, then the listener would interpret that C was the corrective information being conveyed. This part looks into how well speakers are able to highlight important contrasting information in an assertion.

Data were collected through the mediation of one of the present researchers, who intervened as little as possible during the tasks and was responsible for explaining all the tasks for data collection and solving any doubts. Data collection with each pair lasted approximately 90 minutes.

3.2 Research questions and Hypotheses

In the present study, nuclear stress placement to signal information being corrected or elicited was investigated. It was not the aim to discuss the differences in prominence and describe the intonational patterns (the tunes – High or Low), but rather identify the locations chosen by the speakers to place nuclear stress in order to signal elicited/corrective information. Although investigating how the utterances were interpreted was also the aim of the original study, the present article reports only the production results. The research questions and hypotheses guiding the study are the following:

- RQ1: Is nuclear stress placed as expected by BP-IL users (speakers) regardless of information being corrective or elicited?
- H1: Nuclear stress will be placed as expected more often when conveying corrective information.

RQ2: Is nuclear stress placed as expected by BP-IL users (speakers) in all utterance positions?

H2: Nuclear stress will be placed as expected when focus is on the final portion of the utterance. Nuclear stress in the initial and medial portions will be mostly misallocated.

3.3 Data Analysis

For the analysis of the dataset, acoustic and visual criteria were used. For the visual display of the F0 (fundamental frequency), the audio recording files were open with *Praat* and the F0 curves were individually collected. For each sentence, the F0 curve was displayed in windows whereby F0 values were set near the highest and the lowest limits of the curve in order to keep the curve integrity, with no flattening, a procedure that is often adopted in studies involving prosody (e.g., CARPES, 2014).

According to t'Hart (1981), in order for variation in tones to be perceptually distinct, there is a need of a three-semitone change in pitch. While analyzing the speakers' productions in the present study, it was noticed that some pitch changes did not reach the three-tone range. Moreover, sentences usually had more than one prosodic prominence. Based on the literature, one of the hypotheses was that BP speakers would place nuclear stress on its canonic position, that is, in the rightmost end. Surprisingly, a great number of sentences had some prominence both at the expected location and at the leftmost edge of the sentences. Therefore, in order to make important information emerge from the dataset and to understand these preliminary findings, the words in the sentences were segmented and labeled and the vertical scales were normalized from Hertz into semitones, a tool offered by *Praat*.

The thought groups within each utterance were identified according to the number of pauses made by each speaker. No consensus has been reached in the literature regarding how long a period of silence has to be in order to be considered a pause, which has ranged from 100 to 400 milliseconds of threshold (LEGE, 2012). Warren (2013) explains that because some segments such as /p/ have some natural silence due to articulatory reasons, researchers agree to set a duration of 200 milliseconds for silences to be considered as pauses. On the other hand, pauses are noticed even when there is no silence. Other signals such as duration of syllables and pitch change are cues to identify them. Based on this, silences longer than 200 ms were considered as pauses and silences shorter than that associated with rising intonation and/or

lengthening of syllables were also considered indicators of pauses in speech. Thought groups were considered as having a nuclear stress only if differences in the lower pitch and the higher pitch (and vice-versa) were equal to or higher than three semitones¹⁵.

4. Results and discussion

4.1 RQ1: Nuclear stress production and type of information

In all, 160 utterances and their repetitions when applicable were analyzed. Nuclear stress was considered placed as expected if only one thought group and one nuclear stress were produced. Allocations were considered unexpected in the cases described below:

- 1) One thought group with one nuclear stress allocated in an unexpected portion of the utterance;
- 2) Two or more thought groups, with two or more nuclear stress positions, including or not the expected one.

Hypothesis 1 predicted that nuclear stress would be placed as expected more often when conveying corrective information. As can be seen in Table 1, regarding the type of focus given, signaling information being elicited (utterances 1-20) and information being corrected (utterances 21-40), results showed that 8.75% (N = 7) of the elicited information focus and 7.50% (N = 6) of the corrective information focus were produced as expected. Regarding the unexpected productions, they were 91.25% for elicited information and 92.50% for corrective information. For speakers in the present study, signaling correction and elicited information by means of nuclear stress placement was equally challenging, which disconfirmed Hypothesis 1.

Table 1: Production results – elicited information VS corrective information.

	Elicited Information		Corrective Information	
	Frequency	Percent	Frequency	Percent
Expected	7	8.75	6	7.50
Unexpected	73	91.25	74	92.50
Total	80	100.00	80	100.00

¹⁵ The measurement of the semitones was carried out by checking the lowest pitch near the context of pitch change, irrespective of being within the portion in focus (CARPES, 2014).

4.2 RQ2: Nuclear stress production and sentence position

The utterances had different expected focused portions as follows: 22.5% (N = 36) in initial position, 62.5% (N = 100) in medial position, and 15% (N = 24) in final position. Table 1 shows the results for production by the four speakers according to the location of nuclear stress.

Table 2: Nuclear stress production according to their locations.

Location	Expected		Unexpected		Total	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Initial	5	13.9	31	86.1	36	22.5
Medial	5	5.0	95	95.0	100	62.5
Final	3	12.5	21	87.5	24	15.0
Total	13	8.1	147	90.1	160	100.0

Hypothesis 2 predicted that nuclear stress would be placed as expected when focus was on the final portion of the utterance, and misallocated when in the medial and initial portions, based on the tendency BP have of placing nuclear stress in the last word in their L1. However, this hypothesis was not supported as results (Table 2) showed that nuclear stress was produced as expected only 8.1% of the times (N = 13) and the difference of the expected production according to location per se was not considerable.

Altogether, unexpected placement occurred 90.1% of the times (N = 147). When only one nuclear stress was produced, misallocation tended towards the end of the utterance. A zoom in at the speakers' production (Table 3) showed that the unexpected allocation of nuclear stress when there was the production of only one thought group was always produced at final position (N = 4). These results go in line with the BP tendency of placing nuclear stress at sentence final position and the General Phrasal Accent Principle, which says that nuclear stress falls in the final position (LAMBRECHT, 1998). Figure1 illustrates one of these productions.

Table 3: A zoom in at the unexpected stress allocations.

	Details	Frequency	Percent
	1 TG, Final position	4	2.71
	+ TG, + nuclear stress positions	81	55.1
+ TG, + nuclear stress position, syllable lengthening and/or greater pitch range in the expected one		62	42.2
	Total	147	100.0

Figure 1 shows an illustration of a production with only one thought group having the nuclear stress placed in the last word of the utterance. The first tier displays the utterance while the second tier displays the pitch variation, calculated in semitones. Note that the second tier includes the highest and the lowest pitch values which are subtracted to yield the pitch variation (placed after an “=” signal). It is important to highlight that, for Speaker 1, the Praat window vertical settings were from 130 to 250 Hz, a somewhat narrow range, while the other speakers would have a window ranging from 100 up to, at times, 550 Hz. The sentence begins at a higher pitch and gains a slight descending movement (pitch change of 2 semitones maximum). It only has some syllable lengthening with some perceptually noticeable pitch change in the last word, namely, ‘first’. Notice that this is a production which is closer to broad focus than to narrow focus. By doing this, the intent of the speaker here would be to inform that Woody Allen was born on that specific day rather than correcting someone on inaccurate information previously provided, viz. that Woody Allen *had gotten married* on December first. So, as a consequence, Listener 1 misinterpreted the intent of the speaker.

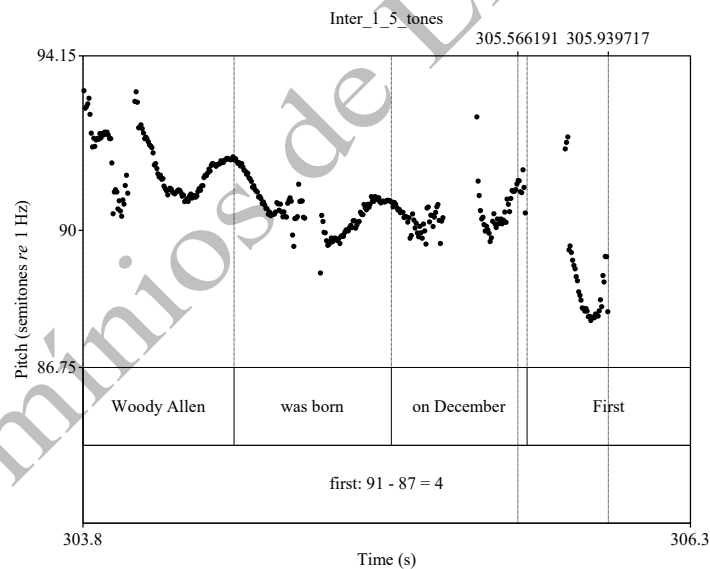


Figure 1: ‘Woody Allen was BORN on December first.’ produced by Speaker 1.

Allocations were also considered unexpected when (1) more than one thought group was produced and thus (2) nuclear stress was placed in more than one position. As displayed in Table 3, 55.1% of unexpected allocations were produced with no cues of greater stress in one specific portion, while 42.2% showed to have one specific portion with a greater pitch range or syllable lengthening. The separation of one thought group from another was sometimes

evidenced by silence gaps, intonation and/or syllable lengthening. Figures 2 and 3 illustrate such separations.

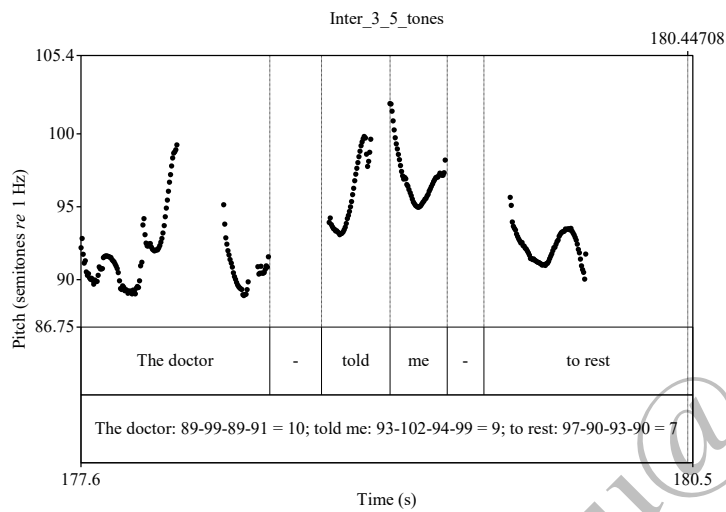


Figure 2: ‘The doctor told me to REST’ produced by Speaker 3.

In ‘The doctor told me to REST’, Speaker 3 seems to break the sentence into three thought groups: 1) the DOCTOR, 2) TOLD ME, and 3) TO REST. The second tier in Figure 2 shows the pitch variation. Observe that this utterance production has many pitch variations, which were annotated. After the = signal, the number displayed corresponds to the subtraction of semitones in two neighbor highest/lowest and lowest/highest pitch (e.g., in the case of “told me”, 93-102 yields the biggest change in pitch). Note that it is possible to observe changes in pitch in the three parts, separated by pauses. The pause between ‘doctor’ and ‘told’ is of 242ms and between ‘me’ and ‘to rest’ it is of 173ms. Even though the silence gap was shorter than 200ms, “me” had some lengthening effect that led to the feeling of a pause being placed in between. This utterance would be interpreted as if every piece of information is being corrected: the doctor, not you; told me, not you; to rest, not to exercise. Nevertheless, even having these pauses, Listener 3 was able to interpret the intended message, that is, informing the listener what the doctor had said¹⁶.

¹⁶ The context question was ‘What did the doctor tell you?’.

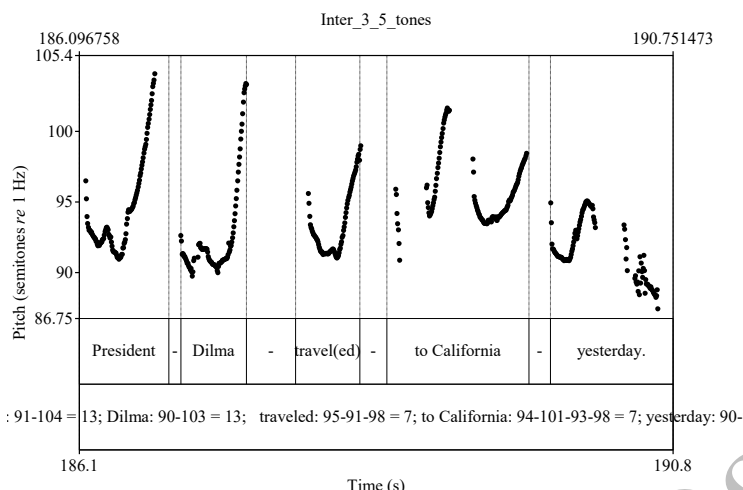


Figure 3: 'President Dilma traveled to California YESTERDAY' produced by Speaker 3.

Breaks of speech flow were also made through the combination of pauses and rising intonation. Note that in Figure 3 every word up to 'California' has a final rise followed by pauses (94ms, 385ms, 212ms, and 169ms, respectively). This pitch movement imposes somewhat a listing effect, such as that of a shopping list reading. 'Yesterday' is the only word without this trait and is the place for the expected focus. The sentence is then produced as if the speaker is trying to say every word in a suspense for the information under focus to be revealed at the end of the utterance. Even though this breaking up may make information more difficult to process, Listener 3 was able to interpret that the information being provided was when the president had traveled to California¹⁷.

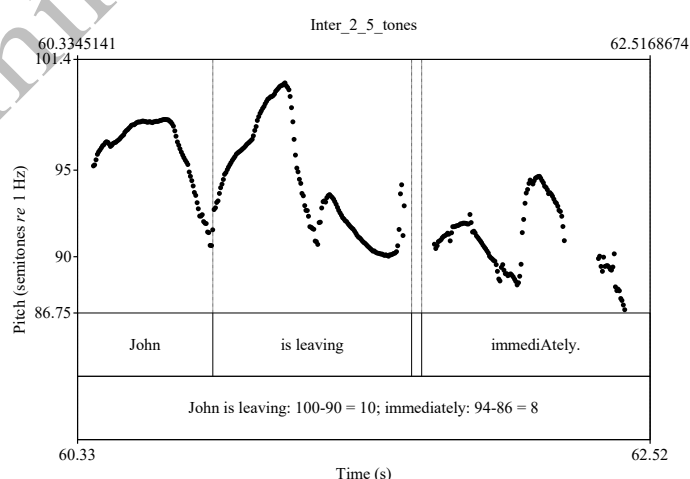


Figure 4: 'John is LEAVING immediately' produced by Speaker 2.

¹⁷ The context question was 'When did President Dilma travel to California?'.

Figure 4 shows an example of an utterance with more than one thought group and nuclear stress, with more prominence in the expected position. The thought groups were ‘John is leaving’ and ‘immediately’. It is possible to observe pitch changes in the two groups, separated by syllable lengthening and a short pause of 38ms. Nevertheless, pitch change is greater in the first portion and that was the expected nuclear stress position.

Disentangling the reasons why such thought groups were produced goes beyond the scope of this study. However, it was noticed that longer sentences were more difficult to produce with only one thought group and that long words or words which seemed to be difficult for these speakers to produce distracted them from producing the expected focus. There was a tendency for placing some prominence on difficult words to pronounce. To illustrate this tendency, let us take the utterance ‘Pearls melt in vinegar’ (Figure 5) and ‘Maria loves you’ (Figure 6). The two utterances have similar length (5 and 6 syllables, respectively) but showed different levels of difficulty in their production. While the former was produced with pauses and thus with more than one thought group, the latter was produced with no pauses in a single thought group. Let us consider each individually.

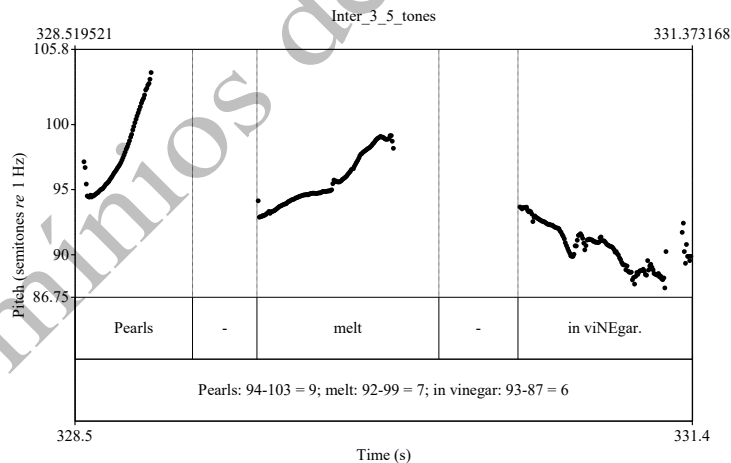


Figure 5: ‘Pearls MELT in vinegar’ produced by Speaker 3.

The context question for ‘Pearls MELT in vinegar’ (Figure 5) was ‘What happens to pearls when in vinegar?’. Hence, the expected portion to receive nuclear stress was ‘melt’. Note, however, that there is pitch movement somewhere else too. Besides having undesired pauses (297ms and 365ms respectively), there were two words which had vowel change + consonant deletion + final devoicing (‘Pearls’ [piərs]) and also word stress misallocation + vowel change

(‘vinegar’ [viˈnegər]). It is possible that the difficulty found to pronounce these words distracted Speaker 3 from making ‘melt’ the most prominent portion of the sentence and the effort to pronounce ‘pearls’ and ‘vinegar’ made the speaker stress these two words in the utterance. The combination of these factors might have contributed to the unexpected interpretation of the intent behind this sentence.

Contrastively, ‘Maria loves you’ seemed to be easier for speakers to utter since none of the words showed to be difficult to pronounce.

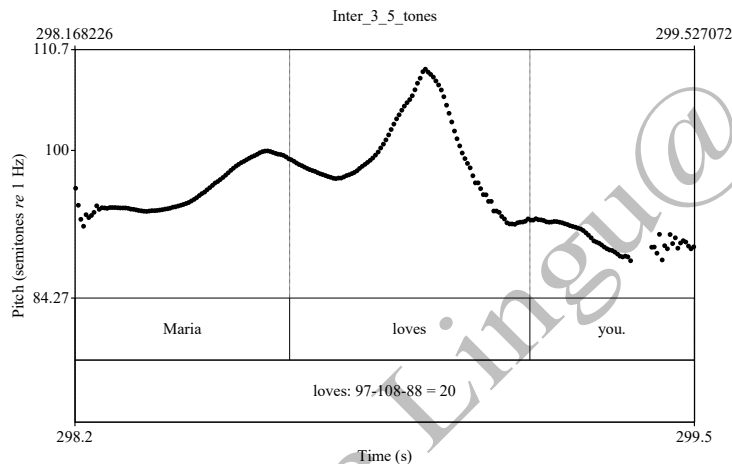


Figure 6: ‘Maria LOVES you’ produced by Speaker 3.

The context question for ‘Maria LOVES you’ was ‘Why doesn’t Maria love me?’. Thus, the expected portion to receive nuclear stress was the medial portion and it was executed as expected. No interruptions were made, no silence gaps were present and the nuclear stress is completely clear with a pitch change of 20 semitones towards the end of the utterance. Regarding the production of long sentences, let us examine the production in Figure 7.

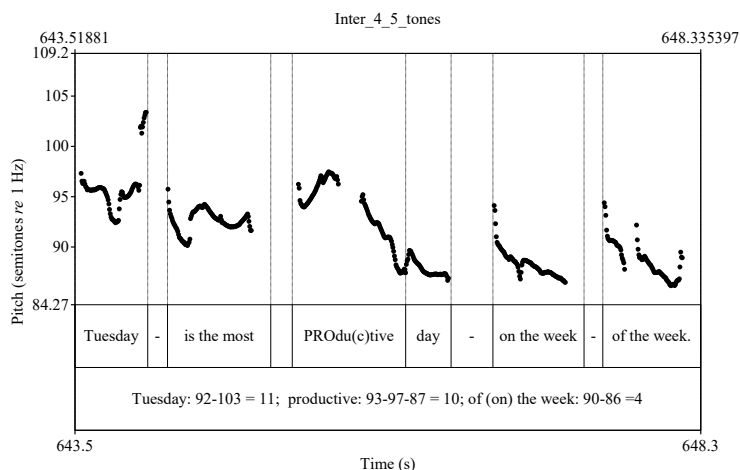


Figure 7: 'TUESDAY is the most productive day of the week' by Speaker 4.

The context sentence for the production in Figure 7 was 'THURSDAY is the most productive day of the week'. Therefore, the expected focus was on 'Tuesday'. Notice that, in this long sentence, there were many short pauses that broke it down into different thought groups, which allowed for the placement of more than one nuclear stress. Observe that the two most prominent words have similar pitch range: 11 semitones in 'Tuesday' and 10 semitones in 'productive'. This production also includes some instances of mispronunciation and word stress misplacement ('productive' as [ˈprɒdʊtɪv]) and hesitation + correction ('on the week' VS 'of the week'). These aspects taken together made this utterance uninterpretable for Listener 4.

5. Conclusion

The present article reported the production results of a pilot study which set out to investigate the interpretability of Brazilian Portuguese users of English as an International Language when word stress and nuclear stress were misallocated. Speakers were four intermediate learners of IL, listeners were two IL teachers (TL group) and two advanced BP-IL users (AL group). Speakers and listeners met in pairs and engaged in face-to-face interactions, which were audio and video-recorded.

Hypothesis one predicted that nuclear stress would be placed as expected more often when conveying corrective information and it was not confirmed. Both types of information showed to be challenging and most productions were produced under an unexpected fashion.

Hypothesis two predicted that speakers would have difficulties in placing nuclear stress in medial and initial position, while the placement in final position would pose no difficulties. This hypothesis was not confirmed as unexpected productions were more abundant in the three

positions than were the expected ones, difficulties being greater for nuclear stress allocation in medial position.

Participants in the study reported here had difficulties to segment speech into meaningful thought groups, irrespective of the type of information and the location of nuclear stress in the utterance. Unraveling the reasons why the misallocations took place go beyond the scope of the study. However, the analysis showed that the presence of challenging words in terms of pronunciation distracted the participants and might have contributed to the great number of pauses, rising intonation, and/or syllable lengthening during the production of these utterances. These issues made more thought groups arise and thus, as reviewed in Section 2 (CELCE-MURCIA *et al.*, 1996), might have made their overall message difficult to process and understand.

Limitations

The present study had several limitations. One of the limitations was posed by the instrument for the nuclear stress placement task. Utterances that had two-word phrases as subjects or as objects were difficult to produce. Additionally, three options in the listeners' sheet would not cover all the possibilities for nuclear stress placement. A better choice would be to include less complex sentences with only three possible positions for nuclear stress placement (e.g., 'Maria loves you'; 'John lives downtown'), so that all possible choices are covered in the options.

Two variables that were not controlled here were the presence of difficult words to pronounce and nuclear stress positions. Future studies should control for these variables combining an equal number of nuclear stress positions and of utterances with harder words to pronounce, such as polysyllabic words with the word stress in the fourth syllable from the right to the left (BRAWERMAN-ALBINI, 2012).

One of the speakers in this study was not that involved with the task. It may be due to a trait of that speakers' personality or a result of the task design. Although literature has supported that one context sentence (question or statement) is enough in order to establish the desired context (e.g., ATECHI, 2004), it is possible that short context narratives (CARPES, 2014) have a stronger setting of the situation to trigger the focus on the expected portion to signal a given piece of information.

Finally, only four listeners participated in the present study due to its complex design of face-to-face interactions. More participants would have yielded more reliable results.

Pedagogical implications

IL users of English take advantage of studying, learning, and practicing to distinguish the subtle shades of meaning that are conveyed by means of prosodic cues from both the receptive and productive stands. Having a good command of the reception and production of prosodic cues, more specifically for this study that of nuclear stress, is essential for a speaker (productively), who is at the same time a listener (receptively), to have his or her intent interpreted as expected. The present study has shown that at this stage of learning, its participants still have difficulties in producing nuclear stress in accordance with the discursive context. Although not scrutinized in this article, participants were misinterpreted with regard to their intent to some extent. Participants reported that they had never been explicitly taught this important prosodic cue. Likewise, the teacher participants also reported not explicitly teaching nuclear stress placement in their English classes. Based on the results of the present study, teaching materials and teaching practice of nuclear stress should then find a place in Brazilian classrooms in order to provide our Brazilian speakers with the resources in order to have the option of improving their nuclear stress placement and augment their chances of avoiding miscommunication in international language use.

Moreover, before teaching nuclear stress placement, it is important to teach how to divide speech into thought groups, because as highlighted by Jenkins (2000) and confirmed in the present study, it does affect nuclear stress placement. It is not only important to teach how to highlight, but what to highlight. In the corpus of the present study, we found problems not only of not knowing how to place nuclear stress (in the case of Speaker 1) but also of the production of too many thought groups and of additional nuclear stress.

Acknowledgements

We would like to express our deepest gratitude to all participants of the present study, who made all the efforts to participate. Our gratitude is also extended to FONAPLI for making the room available for data collection, to João Paulo Acosta Luz for the assistance with the recording equipment, to Daise Fabiana Ribeiro Pereira Carpes for the brainstorming on how to

analyze this dataset, and to the members of NUPFALLE Research Project for all the support given.

References

ATECHI, S. N. **The intelligibility of native and non-native English speech: A comparative analysis of Cameroon English and American and British English.** 2004. 263 f. Dissertation (Dr.phil). Technischen Universität Chemnitz, Germany, 2004.

BAPTISTA, B. Frequent pronunciation errors of Brazilian learners of English. In: FORTKAMP, M.; XAVIER, R. (Eds.). **EFL teaching and learning in Brazil: Theory and Practice.** Florianópolis: Insular, 2001. p. 223-230.

BECKER, M. R. A questão da inteligibilidade do inglês como língua franca. In: VII CONGRESSO INTERNACIONAL DA ABRALIN, 2011. Curitiba. **Anais do VII Congresso Internacional da ABRALIN.** Curitiba: UTFPR, 2011. p. 2789–2800.

_____. **Inteligibilidade da Língua Inglesa sob o paradigma de Língua Franca: Percepção de discursos de falantes de diferentes L1s por brasileiros.** 2013. 257 f. Doctoral Dissertation. Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2013.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M. H.; CROFT, W.; ELLIS, N. C.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language Is a Complex Adaptive System: Position Paper. **Language Learning**, Malden, v. 59, s. 1, p. 1-26, December 2009.

BERNS, M. World Englishes, English as a lingua franca, and intelligibility. **World Englishes**, Malden, 27(3), p. 327-334, August/November 2008.

BRAWERMAN-ALBINI, A. **Os efeitos de um treinamento de percepção na aquisição do padrão acentual pré-propoxítono da língua inglesa por estudantes brasileiros.** 2012. 333 f. PhD dissertation. Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2012.

CARPES, D. F. R. P. **Um estudo prosódico-semântico da não exaustividade no português brasileiro.** 2014, 107 f. Master Thesis. Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2014.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages.** New York: CUP, 1996. 447 p.

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

CRUZ, N. C. An exploratory study of pronunciation intelligibility in the Brazilian Learner's English. **the ESPECIALIST**, São Paulo, 24(2), p. 155–175, 2003.

. **Pronunciation Intelligibility in Spontaneous Speech of Brazilian Learners' English**. 2004. 258 f. PhD Dissertation. Departamento de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2004.

_____. Familiaridade do ouvinte e inteligibilidade da pronúncia de aprendizes brasileiros de inglês. **Revista Horizontes de Linguística Aplicada**, Brasília, 7(1), p. 88–103, 2008a.

_____. Vowel insertion in the speech of Brazilian learners of English: a source of unintelligibility? **Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies**, Florianópolis, (55), p. 133–152, Jul./Dez. 2008b. Retrieved from <http://www.journal.ufsc.br/index.php/desterro/article/viewArticle/16307>. Acesso em 18 de abril de 2013.

DERWING, T. M.; MUNRO, M. J. Second Language Accent and Pronunciation Teaching: A Research-Based Approach. **TESOL Quarterly**, v. 39, n. 3, p. 379-397, Setembro 2005. **crossref** <http://dx.doi.org/10.2307/3588486>

FLEGE, J. E. Second language speech learning: theory, findings, and problems. In: Strange, W. (ed.). **Speech perception and linguistic experience: Issues in cross-language research**. Timonium: York Press, 1995, p. 233-277.

GOMES, M. L. de C.; BRAWERMAN-ALBINI, A.; ENGELBERT, A. P. P. F. The perception of Vowel Epenthesis and Word Stress in an English as a Lingua Franca Context. **Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics**, Concordia, 5, p. 185-202, 2014.

JENKINS, J. Teaching intonation for English as an International Language: teachability, learnability and intelligibility. **Speak Out!**, Faversham, 21, p. 15-26, Dezembro, 1997.

_____. **The phonology of English as an International Language**. Oxford: Oxford University Press, 2000. 258 p.

_____. A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. **Applied Linguistics**, 23(8), p. 83-103, 2002. **crossref** <http://dx.doi.org/10.1093/applin/23.1.83>

LAMBRECHT, K. **Information structure and sentence form: topic, focus and the mental representations of discourse referents**. Cambridge: Cambridge University Press, 1998. 388 p.

LANHAM, L. W. Stress and Intonation and the intelligibility of Sound African Black English. **African Studies**, Johannesburg, 43:2, p. 217-230, Janeiro, 1984.

LEGE, R. F. **The Effect of Pause Duration on Intelligibility of Non-Native Spontaneous Oral Discourse**. 2012. 76 f. Master Thesis. Department of Linguistics and English Language, Brigham Young University, Provo, 2012.

LEVIS, J. Assessing speech intelligibility: Experts listen to two students. **Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference**, 2010. Ames. Iowa: Iowa State University, 2011. p. 56-69.

LIBERMAN, M. Y. **The intonational system of English**. 1972. 319 f. PhD Dissertation, Massachusetts Institute of Technology, Cambridge, 1972.

MARTA, E. de S. **Avaliação da produção e percepção da prosódia na expressão de modalidades por falantes de inglês como L1 e aprendizes de inglês como LE**. 2011. 225 f. Tese de doutorado, Programa de Pós-Graduação em Linguística, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

MOYER, A. **Foreign Accent: The phenomenon of Non-Native Speech**. New York: Cambridge University Press, 2013. 232 p. **crossref** <http://dx.doi.org/10.1017/CBO9780511794407>

MUNRO, M. J.; BOHN, O.-S. The study of second language speech: A brief overview. In BOHN, O.-S.. **Language Experience in Second Language Speech Learning: In Honor of James Emil Flege**. Amsterdam: John Benjamins Publishing Company, 2007, p. 3-12. **crossref** <http://dx.doi.org/10.1075/llt.17.06mun>

MUNRO, M. J.; DERWING, T. M.; MORTON, S. L. The Mutual Intelligibility of L2 Speech. **SSLA - Studies in Second Language Acquisition**, New York, 28, p. 111 – 131, Março, 2006.

NELSON, C. L. **Intelligibility in World English: Theory and Application**. New York: Routledge, 2011. 134 p.

ORTIZ-LIRA, H. Word stress and sentence accent. **Cuadernos de la Facultad, Monografías Temáticas**, Santiago, 16, 81 p. Faculdade de Historia, Geografía y Letras, Universidad Metropolitana de Ciencias de La Educación, Santiago de Chile, maio, 1998.

PASSARELLA-REIS, L.; GONÇALVES, A. R.; SILVEIRA, R. Perception of Intonational Patterns and Speaker's Intentionality in English Yes-No Questions Produced by Brazilians. **RELIN**, Belo Horizonte, no prelo.

PIERREHUMBERT, J. B. **The phonology and phonetics of English Intonation**. 1980. 401 f. PhD Dissertation, Department of Linguistics and Philosophy, MIT, Cambridge, 1980.

PIERREHUMBERT, J.; HIRSCHBERG, J. The Meaning of Intonational Contours in the Interpretation of Discourse. In: COHEN, P. R.; MORGAN, J; POLLACK, M. E. (editors). **Intentions in Communication**. Cambridge: MIT Press, 1992, p. 271-311.

SCHADECH, T.S. **The production of word-initial /ɪ/ by Brazilian learners of English and the issues of comprehensibility and intelligibility**. 2013. 98 f. Master Thesis, Programa de Pós-Graduação em Inglês, Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2013.

SCHADECH, T. S.; SILVEIRA, R. How do the non-target pronunciations of the consonants /θ/ and /ð/ by Brazilian learners of English affect comprehensibility? **Cadernos do IL**, Porto Alegre, (1), p. 4-23, junho, 2013.

SLUIJTER, A. M. C.; van HEUVEN, V. J. Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch. **Phonetica**, Basel (Switzerland), 52, p. 71–89, 1995. **crossref** <http://dx.doi.org/10.1159/000262061>

_____. Acoustic correlates of linguistic stress and accent in Dutch and American English. International Conference on Spoken Language - ICSLP. **Proceedings – ICSLP**, Philadelphia, PA: IEEE, 1996. p. 630-633.

T'HART, J. Differential sensitivity to pitch distance, particularly in speech. **Journal of Acoustic Society of America**, New York, 69(3), p. 811-821, Março, 1981.

TIFFEN, B. **The Intelligibility of Nigerian English**. 1974. 384 f. Tese de Doutorado, University of London, London, 1974.

TROFIMOVICH, P.; KENNEDY, S. Interactive alignment between bilingual interlocutors: Evidence from two information-exchange tasks. **Bilingualism: Language and Cognition**, Cambridge, v. 17, p. 1-15, Outubro, 2014. **crossref** <http://dx.doi.org/10.1017/s1366728913000801>

WARREN, P. **Introducing Psycholinguistics**. New York : CUP, 2013. 286 p.

ZOGHBOR, W. S. **The Effectiveness of the Lingua Franca Core (LFC) in Improving the Perceived Intelligibility and Perceived Comprehensibility of Arab Learners at Post-Secondary Level**. 2010. 227 f. Doctoral Dissertation. School of Education, University of Leicester, Leicester, 2010.

Appendix A – Breaking the ice activity



This activity is the first part of a recruiting and selection hiring process of an International Company.

For this activity, you will meet a person for the first time. You don't know the person's name, nationality, age or any further general information.

Try to get to know this person. Ask questions that you would normally ask when meeting someone. Try to sound as natural as possible. Additionally, try to find out three things that you have in common. For example, a dislike

of jilo, Brazilian country music, and sushi.
Remember: You need to use English.

Three things we have in common:

- 1) _____
- 2) _____
- 3) _____



Appendix B – What I am saying?

Finally, you are about to finish the interviews. This time you will help verify how clever the candidate is.

This task is divided into two parts.

In Part I, you will read a question **silently**. Next, you will read **aloud** the **answer** to that question as if you were really answering the question. See an example below:

Read silently: Who got married in 1984?

Read aloud: **Lucy got married in 1984.**

In Part II, you will read a statement silently. Then, you will read aloud a statement to contradict the previous statement. See an example below.

Read silently: Peter has bought a red car.

Read aloud: **Peter has bought a yellow car.**

In the sentence you read aloud you contradict the color. In fact, it is a **YELLOW** car that he has bought, not a **RED** car.

START

Appendix C – What are you saying?

This is the last part of the recruiting and selection hiring process. This time your ability to interpret the intention of speakers will be verified.

This task is divided into two parts. In Part I, you will check the best question for an answer read by the interviewer. See an example below:

You hear: LUCY got married in 1984. (emphasis in “Lucy”)

Your choices are:

- (A) Who got married in 1984?
- (B) What happened to Lucy in 1984?
- (C) When did Lucy get married?

The best alternative is (A), because “Lucy” is the new information highlighted by the interviewer.

In Part II, you will check the alternative that shows an idea that the interviewer is probably contradicting. See an example below:

You hear: Peter has bought a YELLOW car. (emphasis in “yellow”)

Your choices are:

- (A) PETER did, not JOHN.
- (B) BOUGHT it, not SOLD it.
- (C) YELLOW car, not RED.

The best alternative is (C), because by putting emphasis in “yellow” the interviewer is making it clear that the car Peter had bought wasn’t red.

Artigo recebido em: 19.09.2015

Artigo aprovado em: 15.03.2016

Adaptações fonológicas na pronúncia de nomes comerciais com elementos do inglês no Brasil

Phonological adjustments in pronunciation of trade names with elements of English in Brazil

Natália Cristine Prado*

RESUMO: Este estudo tem como objetivo analisar, a partir das teorias fonológicas não-lineares, as adaptações fonológicas que ocorrem nos nomes comerciais formados com elementos do inglês em contexto de Português Brasileiro, especificamente os processos de epêntese e de apagamento. Para realizar este estudo, primeiramente, coletamos nomes comerciais com elementos do inglês no interior do estado de São Paulo (Brasil) e, em seguida, submetemos uma amostra desses nomes comerciais que mantêm sua grafia inglesa à leitura por falantes do Português Brasileiro, com a intenção de comparar as realizações desses sujeitos com a pronúncia dessas palavras em Inglês Norte-Americano. Com esta pesquisa, notamos que a ocorrência de epêntese é mais comum que a ocorrência de apagamento e que ambos os processos ocorreram para adequar as sílabas inglesas ao padrão silábico do português.

PALAVRAS-CHAVE: Nomes comerciais. Português Brasileiro. Inglês Norte-Americano. Epêntese. Apagamento.

ABSTRACT: This study aims to observe, from non-linear phonological theories, phonological adjustments in pronunciation of trade names formed with elements of English in the context of Brazilian Portuguese, specifically the epenthesis and the deletion. For this study, we first collected trade names with elements of English in São Paulo countryside (Brazil). Using a sample of these trade names that keep its English spelling, recordings were carried out with Brazilian Portuguese speakers to compare with the pronunciation of these words in North American English. We verify that the occurrence of epenthesis is more common than the occurrence of deletion and both processes occurred to adapt the English syllable to Portuguese syllable pattern.

KEYWORDS: Trade names. Brazilian Portuguese. North American English. Epenthesis. Deletion.

1. Introdução

O objetivo principal desta pesquisa é analisar a formação de nomes comerciais com elementos do inglês em Português Brasileiro (doravante PB) observando principalmente como os empréstimos se comportam fonologicamente nessa variedade da Língua Portuguesa. Pretendemos, com nossas análises, avaliar como falantes do PB pronunciam nomes comerciais

* Mestre e doutora em Linguística e Língua Portuguesa pela UNESP/Araraquara. Docente do Departamento de Línguas Vernáculas (DLV) da Universidade Federal de Rondônia (UNIR).

criados a partir de palavras do inglês – enfocando especificamente o estudo da epêntese e do apagamento.

O empréstimo de palavras entre idiomas é um fenômeno natural e antigo. De acordo com Paiva (1991, p. 109), “assim como importamos mercadorias de outros países, palavras estrangeiras também terão ampla acolhida no mercado linguístico brasileiro”. No entanto, esse assunto vem ganhando destaque no meio acadêmico, na mídia e até mesmo na política¹ por conta do número de palavras emprestadas do inglês: os chamados anglicismos.

Trask (2004, p. 164) observa que

nas últimas décadas os sociolinguistas estão conscientes de que conferir uma identidade à pessoa como indivíduo e membro de um grupo é uma das mais importantes funções da língua. Considerar que a língua tem essa função de identificação é crucial para entender muitos tipos de comportamento social e linguístico.

Essas observações de Trask são interessantes para os estudos sobre os anglicismos, afinal notamos que as palavras emprestadas de outras línguas trazem “uma suspeita de identidade alienígena, carregada de valores simbólicos relacionados aos falantes da língua que origina o empréstimo” (GARCEZ; ZILLES, 2004, p. 15). Os autores lembram que os valores associados aos empréstimos podem ser conflitantes (e não raro provocam reações negativas por parte de puristas e de falantes “comuns”) uma vez que “diferentes grupos em uma comunidade podem atribuir valores diversos às identidades ligadas aos falantes de outras línguas” (GARCEZ; ZILLES, 2004, p. 15-16). Assim, torna-se interessante verificar como se comportam falantes do português diante do fenômeno dos empréstimos de palavras do Inglês Americano (doravante IA).

Para Paiva (1991, p. 22), nos últimos tempos, aprender inglês começou a se tornar muito importante e são diversos os motivos que levam as pessoas a estudar este idioma. O inglês, segundo a autora, é visto inclusive como uma forma de ascender socialmente e melhorar de vida. É interessante notar, em alguns contextos, que o fato de não saber inglês pode, até mesmo,

¹Não podemos deixar de nos referir aqui ao Projeto de lei n.º 1676 de 1999 do então deputado Aldo Rebelo. Este Projeto dispunha sobre “a proteção, a defesa e o uso da Língua Portuguesa” e visava à proibição do uso de palavras estrangeiras, sobretudo os anglicismos, no PB. Na época, o assunto gerou grande revolta na comunidade linguística, que se manifestou contrária ao Projeto (FARACO, 2004; FIORIN, 2004; GARCEZ; ZILLES, 2004; MASSINI-CAGLIARI, 2004a, 2004b; ZILLES, 2004; CARVALHO, 2009). Posteriormente, este Projeto foi reformulado e aprovado pelo Senado em uma nova versão que foi proposta pelo senador Amir Lando, em 28 de maio de 2003.

representar algo negativo. De acordo com Carvalho (2009, p. 27-28), no princípio os anglicismos como *bife*, *rosbife* e *lanche* vieram diretamente da Inglaterra. No entanto, atualmente, com o deslocamento do centro do poder político para os EUA, os anglicismos passaram a ter origem, sobretudo, no IA. Exemplos de vocábulos emprestados recentemente, já num momento de maior influência do IA, são palavras como *pen drive*, *skate* e *design* (CARVALHO, 2009, p. 28).

No contexto comercial, é muito importante que o produto a ser vendido alcance cada vez mais consumidores, ou seja, “fale” com o maior número de pessoas. Nesse sentido, a questão fundamental a ser pensada pelos grandes comerciantes e publicitários diante de um mundo “globalizado” é como negociar de modo eficiente em nível global e, para isso, muitos investem não só em publicidade em outros idiomas, mas nomeiam seu produto com palavras de uma língua que eles acreditam que muitas pessoas conheçam (ou que tenha prestígio), isto é, um idioma que tenha um papel de língua global.

Atualmente, enquanto o inglês é reconhecido por muitas pessoas como uma língua de alcance global (CRYSTAL, 2003) – e muitos comerciantes escolhem palavras deste idioma para nomear seus produtos e negócios – muitas indústrias, lojas e marcas importadas, às vezes dos E.U.A., também entram na competição pelo mercado consumidor de diversos países. Isso mostra que, em qualquer lugar do globo, é possível comer no *McDonald's* ou comprar um tênis *Nike*. Como produtos importados geralmente não são acessíveis a todos os consumidores por conta do seu valor monetário muitas vezes mais alto que os produtos nacionais, tradicionalmente eles ficaram restritos à parcela da população economicamente privilegiada, adquirindo, assim, um *status* sofisticado.

Assim, podemos dizer que, no uso de elementos da língua inglesa na formação de nomes comerciais, há muito mais uma necessidade simbólica de identificação social e cultural do que propriamente a de nomear um novo objeto ou produto. Afinal, como lembram Garcez e Zilles (2004, p. 22-23):

O apelo da máquina capitalista globalizante é forte demais para que a mídia da informação, do entretenimento e, principalmente, da publicidade possa ou queira deixar de explorar as associações semióticas entre a língua inglesa e o enorme repositório de recursos simbólicos, econômicos e sociais por ela mediados.

Devido ao grande número de palavras da língua inglesa circulando de várias formas na Língua Portuguesa, sempre surgem pesquisas sobre os anglicismos. Estes estudos, como afirma Paiva (1991, p.127), vêm seguindo duas posturas diferentes: parte dos autores defende, de modo intransigente, o idioma nacional de uma influência estrangeira, enquanto a outra parte, com postura mais liberal, considera esse fenômeno uma consequência normal do desenvolvimento de qualquer língua. Entretanto, ainda são poucos os pesquisadores que têm se dedicado mais especificamente aos estudos da adaptação de estrangeirismos à fonologia do português. Dentre os trabalhos que observam este fenômeno, podemos destacar os estudos de Freitas (1984, 1992), Freitas et al. (2003), Assis (2007), Massini-Cagliari (2010, 2011a, 2011b, 2013), Souza (2011), Prado (2014) e Macedo (2015).

Portanto, assim como os trabalhos citados, este estudo tem o intuito de contribuir para as análises das adaptações fonológicas que podem acontecer quando um termo de uma língua estrangeira é inserido no contexto linguístico do português. Deste modo, pretendemos estudar os processos de epêntese e apagamento que ocorrem na pronúncia dos anglicismos presentes no meio comercial em PB.

2. Estudos sobre nomes de estabelecimentos comerciais

Ao longo de nossa pesquisa, encontramos poucos trabalhos que têm como foco o estudo de nomes de estabelecimentos comerciais. Neves (1971, p. 30), por exemplo, pesquisou nomes próprios comerciais na cidade de Belo Horizonte e observou que, do mesmo modo que o pai batiza um filho com o nome da moda ou nomes de origem familiar/afetiva, o comerciante também nomeia seu estabelecimento seguindo esses mesmos critérios e pensando que o nome deve ser eficaz para representar seu negócio junto ao público consumidor.

Para a autora (NEVES, 1971, p. 38), é de extrema mobilidade o léxico da propaganda comercial, pois, além de refletir a agitada vida do comércio, acompanha o progresso da ciência e tecnologia. A estudiosa reflete sobre a questão da transparência das palavras comerciais em oposição à sua opacidade, pois, ao contrário do que ocorre no léxico comum, muitas vezes imotivado, o nome comercial normalmente tem uma forte relação com o estabelecimento que nomeia. Neves (1971, p. 41) acredita que “é o nome próprio comercial que, a nosso ver, justifica a volta da antiga polêmica entre os linguistas, em torno da teoria da arbitrariedade do signo linguístico”.

Segundo Neves (1971, p.43), apenas pessoas com pouca imaginação não se preocupam em batizar, em individualizar o seu negócio dentre os congêneres, sendo que o mais frequente e inteligente é procurar o nome mais apropriado para o estabelecimento comercial. Para essa tarefa, os comerciantes têm à disposição um material rico e variado tanto nacional quanto estrangeiro e podem usar ou não palavras denotativas do ramo de atividades a que se dedica (*bar, mercearia, loja* etc). A partir de seus estudos a autora afirma que as denominações usuais de estabelecimentos comerciais de Belo Horizonte, e possivelmente de todo o Brasil, “oferecem ao estudioso da língua um material farto e praticamente inexplorado” (NEVES, 1971, p. 58-59).

Os nomes próprios criados para estabelecimentos comerciais, como qualquer outro neologismo vocabular, dividem-se em formações vernáculas e empréstimos. No entanto, de acordo com Neves (1971, p. 63), nada impede que os nomes de formação vernácula apresentem elementos mórficos de línguas estrangeiras, clássicas ou modernas, por exemplo, *Chez Bastião*, um bar presente no *corpus* de sua pesquisa, que é um nome formado em português, mas com elementos do francês. Em nosso *corpus*, também encontramos vários nomes comerciais que misturam palavras do português com palavras do inglês, como *Academia Centro Aquático Shark*.

A autora também cita como exemplo a palavra *Drugstore* que, em seu *corpus*, é um nome de um bar enquanto na língua inglesa significa *farmácia*, para mostrar que os comerciantes se apropriam de palavras estrangeiras podendo lhes dar novos sentidos. Apesar de encontrarmos, no Brasil, nomes comerciais que são conhecidos no mundo todo (como *Chevrolet, Ford*), nem toda firma que tem um nome comercial escrito com palavras estrangeiras é necessariamente estrangeira ou está pensando em internacionalização. De acordo com a opinião da autora, “o uso de estrangeirismos, nesse caso, é um recurso de expressividade”.

Em seu trabalho, a pesquisadora mostra que os comerciantes vão não só em busca de vocábulos em línguas clássicas, como o grego (como, por exemplo, a palavra *Delta* – presente em vários nomes comerciais) e latim (como, por exemplo, o nome *Fiat Lux* – conhecida marca de fósforos), mas também bebem nas fontes das línguas modernas como francês, inglês, italiano etc. para encontrar o nome perfeito para seu negócio. Embora existam nomes comerciais com elementos de várias línguas, a autora acredita que “nos últimos anos, porém, tem se sobressaído o inglês” (NEVES, 1971, p. 81).

Para a estudiosa, o uso do inglês na moderna publicidade brasileira é bastante insinuante e se manifesta até mesmo na estruturação frasal, como nos casos de nomes comerciais com o caso genitivo, como, nos exemplos citados por ela (NEVES, 1971, p. 87), *Juka's* (restaurante) e *Mangueiras's* (drive-in). Na época em que realizou sua pesquisa, Neves (1971, p. 90) acreditava que a influência do inglês no meio comercial tendia “a diminuir, haja vista a reação nacionalista que está sacudindo o Brasil e que já se faz sentir inclusive nas denominações vernáculas de estabelecimentos comerciais²”; entretanto, ainda hoje, a maior parte dos empréstimos comerciais (e de nomes comuns) vem do inglês, mas, assim como no trabalho de Neves (1971), a maioria dos nomes comerciais consultados para nossa pesquisa são formados por palavras da língua portuguesa.

Neves (1971, p. 105) ainda comenta alguns processos especiais que são utilizados na formação de nomes comerciais. O primeiro deles é chamado por ela de *neologismo de grafia estilizada*. A autora explica que, embora a grafia estilizada das palavras seja um recurso usual de expressividade, é nos nomes de marcas de produtos que ela se torna mais requintada. Deste modo, para a autora, “muitos neologismos comerciais baseiam-se quase que exclusivamente numa maneira bizarra, extravagante, de escrever a palavra”.

Dentre os nomes com grafias estilizadas coletados pela autora em seu trabalho estão alguns bastante interessantes e criativos (NEVES, 1971, p. 106), como *JAI000* (loteria esportiva), *Pal Pit's* (loteria esportiva), *BE-B* (bar) e *Dom Quixope* (bar). Esses nomes fazem trocadilhos com a língua portuguesa e também se valem da inspiração da língua inglesa, como no caso de *Pal Pit's*, para estilizar as palavras e mostram que, no momento da criação de um nome comercial, as pessoas se valem de toda a criatividade e imaginação possíveis.

Em suas observações finais, a autora nota que os nomes comerciais caracterizam-se pelo predomínio de palavras do tipo transparente sobre as do tipo opaco; além disso, para atender às necessidades sempre crescentes da publicidade, os nomes comerciais importam material linguístico do estrangeiro – também recorrendo ao português numa utilização às vezes agramatical, porém bastante expressiva. Especificamente sobre os estrangeirismos no comércio, a autora acredita que isso mostra não a cultura ou o domínio de línguas estrangeiras por parte dos comerciantes, mas que essas línguas podem ser utilizadas como um dos mais eficientes recursos linguísticos de expressividade publicitária.

² Neste momento, o Brasil vivia o chamado “milagre econômico”, período por um crescimento acelerado, decorrente em grande parte das reformas ocorridas no período anterior e das condições internacionais favoráveis.

De modo geral, segundo a autora, “a escolha dos nomes para a firma ou para o estabelecimento comercial demonstra, por parte de nossos comerciantes, um pleno reconhecimento da importância publicitária da imagem desse mesmo nome junto à opinião pública” (NEVES, 1971, p. 140). Para a estudiosa, a linguagem da propaganda comercial brasileira é um importante elemento para se conhecer a cultura, a psicologia e os interesses de nosso povo, em um dado momento da história.

Para Giacomini Filho e Borba (2010, p. 218), assim como as pessoas e os objetos, os produtos e empresas precisam adotar nomes que os representem no contexto social e de mercado. Para os autores, a correta atribuição de um nome de marca pode implicar no êxito ou no fracasso dos intentos de marketing, além das implicações legais e de responsabilidade social organizacional no uso e registro da marca. Os estudiosos fizeram uma pesquisa com abordagem mercadológica e onomástica para 16 nomes de marcas de empresas que fabricam chocolates artesanais e semi-artesanais e observaram que o “uso intenso de palavras com etimologia estrangeira predomina, talvez pela intenção dos fabricantes de associar o chocolate artesanal a algo especial, diferenciado, exótico, incomum”, o que pode ser obtido com uso de expressões inspiradas em regiões e outros países para nomear os estabelecimentos que vendem chocolates artesanais e os próprios produtos, como *Chocolates Di Siena*, *Casa de Chocolate Nobre Windsor*, *Chocolates Viermon Ltda*, *Chocolates Genebra* e *Chocolates Delicatto Ltda*, (GIACOMINO FILHO; BORBA, 2010, p. 226). De acordo com o estudo dos autores, foi observado que os nomes de marca “são fundamentais para os processos de marketing que, por sua vez, destacam instrumentos investigatórios e de gestão de marcas na construção das expressões que nomearão produtos e empresas na mente do consumidor e dos públicos de interesse” (GIACOMINO FILHO; BORBA, 2010, p. 227).

Serra e Nodari (2011) investigaram a motivação para o uso de estrangeirismos nas fachadas de lojas encontradas no centro comercial de São Luís – Maranhão, especificamente na Rua Grande – via de comércio bastante popular na cidade. As autoras observaram que as fachadas dos estabelecimentos comerciais presentes nessa rua empregam tanto estrangeirismos com a grafia original do inglês quanto termos já adaptados. No entanto, para refletir sobre o assunto, as pesquisadoras entrevistaram donos, gerentes e clientes de apenas nove lojas, dando prioridade para aquelas que tinham os nomes totalmente em inglês (sem nenhuma palavra do português). Quando perguntados sobre o motivo de terem colocado o nome do estabelecimento em inglês, cerca de 56% dos proprietários entrevistados “ênfatizaram que o uso do inglês em

detrimento do português é uma forma de chamar atenção e diferenciar o produto”, e 44% dos entrevistados “ressaltaram o uso da denominação em inglês como uma forma de ‘marketing’ do empreendimento” (SERRA; NODARI, 2011, p. 233).

Segundo as autoras, as respostas que obtiveram no questionário aplicado apontam o prestígio que as lojas com nomes em inglês têm entre os clientes, já que 52% dos fregueses entrevistados afirmaram que o nome estrangeiro os influenciava a comprar mais. Os proprietários também afirmaram que ter um estabelecimento com nome em inglês interfere na venda, induzindo o cliente a comprar mais. A partir dessas respostas, as pesquisadoras observaram a intenção explícita dos proprietários de dar um ar mais “chique” aos seus produtos e seus estabelecimentos.

Outro dado interessante da pesquisa é que a maioria dos clientes, cerca de 73% deles, não sabiam o significado do nome da loja em português; além disso, 22% dos responsáveis pelas lojas também não sabiam. Ao serem questionados acerca da compreensão do significado do nome comercial por parte dos clientes, todos os donos e gerentes responderam que achavam improvável que as pessoas entendessem o significado do nome da loja. Entretanto, esse não pareceu ser um fator de preocupação para os empresários, pois eles não viam necessidade de uma tradução para os termos estrangeiros empregados.

A pesquisa das autoras também mostra que o inglês pode transferir ao produto a ideia de maior qualidade e sofisticação, pois 55% dos clientes afirmaram que uma loja que tem um nome com termos em inglês oferece melhores produtos aos clientes. Um dado curioso levantado pela investigação com os clientes das lojas populares é o fato de 55% deles acreditarem que as lojas da *Rua Grande* que possuem anglicismos no nome poderiam perfeitamente ser comparadas com as do shopping da cidade (*São Luís Shopping*). As investigadoras observaram ainda que a maioria dos lojistas insiste no argumento de que os estrangeirismos repercutem melhor e atraem de forma mais eficaz a atenção do consumidor e também acreditam que seus estabelecimentos comerciais realmente oferecem produtos diferenciados e de boa qualidade.

A partir da investigação de Serra e Nodari (2011), podemos inferir que a inclusão de termos da língua inglesa no léxico comercial dos nomes observados traz consigo certos valores simbólicos. Isso fica claro no momento em que os clientes das lojas populares estabelecem comparação dos produtos dessas lojas com os vendidos em lojas de grife, lojas que carregam o simbolismo da sofisticação e da qualidade normalmente associados a produtos importados, baseando-se unicamente na presença de estrangeirismos no nome da casa comercial.

O estudo dos nomes comerciais pode interessar a diversas áreas do conhecimento, além dos estudos linguísticos. Com relação aos estudos linguísticos, o estudo dos nomes comerciais encontra-se entre um dos tópicos de interesse da onomástica, da morfologia, da fonologia, da neologia, da semântica, da estilística e de muitas outras disciplinas; no entanto, ainda se nota que poucos autores debruçam-se sobre este tema dentro da linguística.

3. Reflexões sobre a epêntese no PB

Cagliari (1999a, p. 129) afirma que, na fonologia, há vários processos que dizem respeito à inserção de um segmento dentro de palavras. O autor explica que, além do termo *inserção*, o termo *epêntese* refere-se, de um modo geral, ao acréscimo de um segmento a uma palavra. Segundo o linguista, “quando a adição de um segmento acontece no início de uma palavra, o termo mais específico para esse processo é *prótese*, como ocorre em *stress* [is'tres]. Se ocorrer o acréscimo de um segmento no final de uma palavra, o processo de inserção tem o nome específico de *paragoge*”. Entretanto, para Massini-Cagliari (2005, p. 275), epêntese e paragoge são processos fonológicos distintos,

uma vez que existem dois tipos de inserção de vogais que atuam em final de palavra: um primeiro, motivado pela busca de estruturas silábicas possíveis dentro da língua (em relação ao qual será mantido o rótulo de ‘epêntese’), e um outro, de motivação rítmica (para o qual será reservado o rótulo de ‘paragoge’). Nesse sentido, o termo *paragoge* pode ser sucintamente definido como ‘epêntese rítmica’.

Como se pode notar, esse tipo de inserção afeta as sílabas que já são bem formadas, como em *falar* > *falare*.

Cagliari (1999, p. 132) explica que a inserção ou epêntese de uma vogal acontece para tornar uma sílaba canônica, desfazendo uma estrutura mal formada, sendo que essa inserção pode aparecer no início, no meio ou no final de palavras.

De acordo com Hogg e McCully (1999 [1987], p. 35), o inglês apresenta 12 padrões silábicos para monossílabos tônicos e 3 para monossílabos átonos, como se pode ver no quadro 1:

Quadro 1: Molde silábico do inglês (adaptado de HOGG; McCULLY, 1999 [1987], p. 35).

Monossílabos Tônicos				Monossílabos Átonos	
Id	VC	isle	VV	a	V
Bad	CVC	bye	VVC	an	VC
Brad	CCVC	bide	CVV	the	CV
Band	CVCC	bind	CVVCC		
Brand	CCVCC	brid	CCVVC		
I	V	grind	CCVVCC		

No entanto, o quadro 1 não dá conta de sílabas do tipo CCCVCCCC, como em *strengths*, por exemplo, e de algumas outras sílabas da língua. Diante desse problema, concordamos com Roach (1998, p. 71-74), que afirma que essa é a maior estrutura silábica do inglês: três consoantes em *onset* e até quatro consoantes em coda silábica. Collischonn (2005 [1996], p. 117) explica que não há consenso entre os estudiosos em relação ao molde silábico do PB, que determina o número máximo e mínimo de elementos permitidos em uma sílaba desta língua. A autora define os seguintes padrões silábicos para o PB:

Quadro 2: Padrões silábicos do PB (COLLISCHONN, 2005 [1996], p. 117).

V	É
VC	Ar
VCC	<u>inst</u> ante
CV	Cá
CVC	Lar
CVCC	<u>mon</u> stro
CCV	Tri
CCVC	três
CCVCC	<u>trans</u> porte
VV	<u>A</u> ula
CVV	Lei
CCVV	Grau
CCVVC	<u>clau</u> stro

Ainda em relação ao português, Cagliari (2007, p. 116) propõe os moldes silábicos³ apresentados no quadro 3⁴:

³ Optamos por mostrar a apresentação dos moldes silábicos de dois autores diferentes – Collischonn (2005) e Cagliari (2007). já que que Collischonn apresenta os moldes silábicos a partir de uma perspectiva da forma subjacente, enquanto Cagliari parece estar preocupado com a forma de superfície.

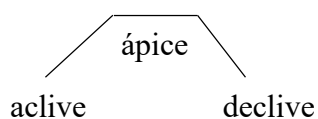
⁴ No quadro 3, M significa monotongo, D significa ditongo e T significa tritongo.

Quadro 3: Sílabas do português (CAGLIARI, 2007, p.116).

Padrão Silábico	Exemplo	Forma ortográfica	
V	[ɛ]	É	M
	[eɔ]	Eu	D
CV	[pɛ]	Pé	M
CV	[teɔ]	Teu	D
	[kɔaɔ]	Qual	T
CCV	[kru]	Cru	M
	[kreɔ]	Creu	D
VC	[ɛs]	És	M
	[eis]	Eis	D
VCC	[xúʔɲs]	Ruins	M
	[léõʔɲs]	Leões	D
CVCC	[pɛs]	Pés	M
	[teɔs]	Teus	D
	[kɔas]	Quais	T
CVCC	[perspektiva]	Perspectiva	M
	[mẽʔɲs]	Mães	D
	[sáɣõõʔɲs]	Saguões	T
CCVC	[plastikɔ]	Plástico	M
	[ũmbrais]	Umbrais	D
CCVCC	[trẽʔɲs]	Trens	M
	[kɔbrõõʔɲs]	Cobrões	D

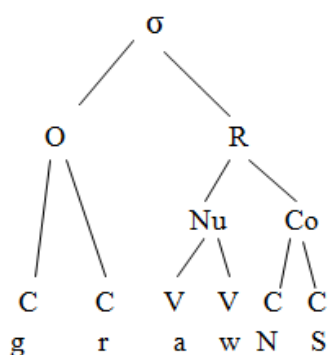
Collischonn comenta os trabalhos de Câmara Jr. (2010[1969]), Lopez (1979) e Bisol (1989), que refletem sobre a questão dos moldes silábicos. Embora Câmara Jr. (2010[1969]) não tenha feito um estudo dos moldes silábicos, podemos deduzi-lo a partir de sua análise. Para ele, a sílaba é formada de um active, de um ápice e de um declive, como no exemplo (2):

(2)



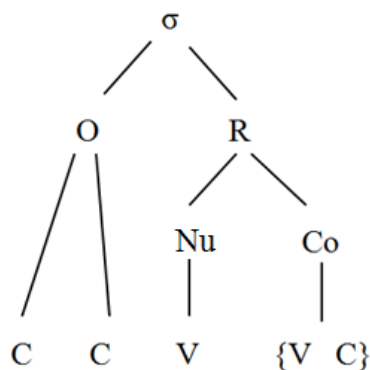
Nesta representação, o ápice é constituído por uma vogal, o aclave, por uma ou duas consoantes e, por fim, o declive, por uma das seguintes consoantes: /S/, /r/, /l/. Além dessas possibilidades, o declive pode apresentar uma consoante nasal – já que o autor interpreta as vogais nasais como sendo fonologicamente vogais fechadas por consoante nasal. Essa análise admitiria seis segmentos na sílaba, como, por exemplo, na palavra *grãos*, que, fonologicamente, teria a representação /grawNS/, como no exemplo (3), em que a ideia de Câmara Jr. aparece representada a partir da estrutura arbórea adotada posteriormente pelos modelos não-lineares.

(3)



Segundo Collischonn (2005 [1996], p. 118), esse molde é inadequado, pois não há, em português, sequências de ditongo e duas consoantes (como **cairs* e **peuls*); dessa forma, seria necessário que a esse molde fossem acrescentadas algumas restrições, a fim de evitar essas sequências. Já Lopez (1979) propõe dois moldes: um para sílaba subjacente e outro para sílaba de superfície. Na sílaba subjacente, a autora considera até quatro elementos, sendo dois no *onset* e dois na rima. Segundo Collischonn, a proposta de Lopez pode ser traduzida na representação (4), na qual as chaves indicam diferentes alternativas de preencher um mesmo constituinte.

(4)



Esse molde, significativamente reduzido em relação ao de Câmara Jr., exclui as sílabas inexistentes, mas também exclui as existentes, como *deus*, *dois* e *cáustico*. No entanto, na forma fonética, esse molde muda, admitindo três elementos no *onset* (como em *criança* e *prior*). Diferentemente de Câmara Jr., Lopez acredita que ditongos nasais têm a forma subjacente VnV, assim, *grão* tem a forma subjacente /granu/. Dessa forma, na ótica de Lopez, os ditongos nasais derivam de duas sílabas, das quais a segunda é iniciada por consoante nasal, que é apagada, passando a VV na superfície. A esta sílaba, podemos acrescentar o morfema de plural, o que faria com que a rima ficasse com três elementos e não quatro, como na proposta de Câmara Jr. Todavia, para Bisol (1989), o ditongo nasal deriva de uma sequência de VC subjacente, em que C se torna flutuante e uma vogal temática é acrescida. A associação posterior do traço nasal à rima cria o ditongo nasal. Para Collischonn (2005 [1996], p. 120), esses dados indicam que a sequência de segmentos se ajusta ao padrão silábico CCVCC, cobrindo as sílabas máximas, como *fausto* e *monstro*.

No entanto, de acordo com Câmara Jr. (2004[1970], p. 54), as semivogais funcionam como consoante, ou seja, ocupam posições na sílaba normalmente reservadas a consoantes, mas são de natureza vocálica, ou seja, foneticamente são articuladas e compreendidas como vogais. Desse modo, surge uma dúvida acerca da representação das sílabas que contêm ditongos: seriam CVC ou CVV? O molde CVC pressupõe uma sílaba travada, enquanto que CVV, uma sílaba aberta. Embora esta seja uma discussão importante para o levantamento dos tipos silábicos do português e para outras questões que envolvem a determinação do peso e da estrutura silábica (como o posicionamento do acento lexical), neste trabalho, admitiremos que o *glide* está posicionado no núcleo, seguindo a opinião de Câmara Jr. (2004[1970]), uma vez que a diferenciação de posicionamento do *glide* no núcleo ou na coda não afeta as análises

desenvolvidas neste artigo⁵. Assim, considerando-se o *glide* no núcleo da sílaba, uma palavra como “grãos” teria a estrutura CCVVCC, conforme propõe Câmara Jr.. Pode-se assumir que a sílaba máxima do PB é, então, CCVVCC e não CCVCC, como observado nos quadros anteriores.

Freitas e Neiva (2006, p. 6) explicam que, em relação ao *onset* do inglês, quando as três posições possíveis são preenchidas, C1 deve ser obrigatoriamente a fricativa alveolar surda [s]; em posição de C2 só podem ocorrer oclusivas surdas, ou seja [p,t,k]; e C3 é ocupado apenas por líquidas ou *glides*, ou seja, [l, ɹ, j, w]. Ademais, as combinações possíveis destas consoantes também são restritas. Quando todas essas posições são preenchidas, a consoante mais próxima ao núcleo (C3) é uma soante contínua e pode vir precedida imediatamente por uma obstruinte não sonora (C2). Já a posição C1 só pode ser ocupada, neste caso, pelo segmento [s]. No entanto, quando as consoantes do *onset* são apenas duas, há um número maior de possibilidades e, se a consoante em C1 é uma fricativa alveolar surda, [s], ampliam-se ainda mais os tipos de combinações de segmentos. Neste caso, a posição C2 pode ser ocupada por uma soante (exceto a palatal [j] e a velar [ŋ]), por uma obstruinte surda não contínua, ou ainda por [f], esta última uma sequência rara na língua. Quando qualquer outra obstruinte diferente de [s] preenche C1, a posição C2 fica restrita a *glides* e líquidas.

Segundo Cagliari (2007, p. 117), nas sílabas CV em PB, o *onset* pode ser qualquer consoante, porém não ocorre o tepe em início de palavras, e a lateral e a nasal palatais, no mesmo contexto, ocorrem só em poucas palavras do português. Já nas sílabas CCV (contendo C1 e C2), C2 será ou um tepe ou uma lateral alveolodental e, quando C2 for tepe, C1 pode ser qualquer oclusiva ou fricativa labiodental. Já quando C2 for uma lateral alveolodental, C1 poderá ser qualquer oclusiva, exceto oclusiva dental sonora, ou ainda, poderá ser uma fricativa labiodental desvozeada, como podemos observar no quadro 4:

Quadro 4 – Possibilidades do *onset* em português (CAGLIARI, 2007, p. 117)

C1	C2	V
p, b, t, d, k, g, f, v	ɹ	V
p, b, t k, g, f	l	V

⁵ Para conhecer os argumentos contrários ao posicionamento do *glide* no núcleo silábico em ditongos decrescentes do PB, remetemos o leitor ao trabalho de Bisol (1989) e Zucarelli (2002).

Collischonn (2005 [1996], p. 120) acredita que, em língua portuguesa, determinadas sequências de segmentos não são permitidas no *onset* complexo. Nesta posição, as sequências permitidas são plosiva + líquida; entretanto, os grupos /dl/, /tl/ e /vl/, apesar de permitidos pelo molde, são restritos a nomes próprios de origem estrangeira, como *Vladimir*. A autora também considera que o grupo /tl/ ocorre na palavra *atlas* e seus derivados. No entanto, podemos encontrar o grupo /dl/ em siglas, como em DLA (Departamento de Linguística Aplicada). Além disso, em posição inicial de palavra, não ocorre /ɲ/ e /ʎ/, de modo que só encontramos estes segmentos nesta posição em palavras emprestadas de outras línguas, como, por exemplo, *Thama* e *nhoque*.

Sobre o núcleo do IA, que pode ser simples ou ramificado, Wells (2008) apresenta os seguintes ditongos /eɪ/ (**day**), /aɪ/ (**try**), /ɔɪ/ (**boy**), /oʊ/ (**show**) e /aʊ/ (**now**). Em relação ao núcleo do PB, que também pode ser simples ou ramificado, os ditongos podem ser classificados em ditongos crescentes e decrescentes (CÂMARA Jr., 2004[1970], p. 56). Os ditongos decrescentes são aqueles em que a proeminência silábica ocorre na primeira vogal e os ditongos crescentes são aqueles em que a proeminência silábica ocorre na segunda vogal, sendo que os ditongos nasais são sempre decrescentes. Para o autor, há um único caso de ditongo crescente em PB que nunca se realiza como hiato: a vogal assilábica /u/ depois de plosiva labial /k, g/ antes de vogal silábica, como na palavra *qual*, que produz o que o autor chama de “tritongo”. Collischonn (2005 [1996], p. 121) também acredita nessa ideia, dizendo que “há um tipo de ditongo crescente que não alterna com hiato. Trata-se de kw/gw, seguidos de a/o”. Já Bisol (1989) defende a não ocorrência de ditongos crescentes subjacentes e interpreta estes casos como fonemas velares labializados: /kw/ e /gw/.

Freiras e Neiva (2006) comentam que, como a língua inglesa permite um agrupamento de até quatro consoantes em coda de final de palavra, considerando-se a sílaba do ponto de vista fonético, esta particularidade da língua inglesa impõe grandes dificuldades para os falantes brasileiros, cuja língua materna permite a ocorrência de no máximo dois segmentos consonânticos ao final de sílabas, com restrições muito limitadas quanto à classe de consoantes possíveis em tal posição.

Nas sílabas VC, conforme Cagliari (2007), a coda do PB poderá ser uma fricativa alveodental ou palatoalveolar, dependendo do dialeto, uma nasal qualquer ou um dos sons do R. Já nas sílabas VCC (contendo C3 e C4), C3 poderá ser uma nasal palatal ou velar ou um dos sons do R, RR. Já C4 será sempre uma fricativa alveolar ou palatoalveolar, variando de acordo

com o dialeto. Este último caso, lembra o linguista, ocorre somente em poucas palavras da língua, como nas palavras *perspectiva* e *perspicaz*. Cagliari (2007, p. 117) afirma que “no português brasileiro, algumas palavras variam foneticamente, podendo ter uma sílaba a mais ou a menos, dependendo de uma vogal breve e átona, em geral [ɪ]⁶, entre uma oclusiva, uma nasal bilabial ou uma fricativa alveolar surda, por um lado, e uma outra consoante por outro lado”. Já, para Lee (1993), a vogal epentética do PB é sempre, fonologicamente, /e/, já que este é o único segmento não-especificado na representação de base, que pode assumir as formas fonéticas de [e] ou [i]. Quanto aos estrangeirismos em específico, Lee (1993, p. 848) diz que “o falante nativo de português insere /e/ quando pronuncia palavras estrangeiras e siglas que têm os sons de [-soa] na posição final da sílaba”.

Como vimos, o padrão silábico da língua inglesa admite até três consoantes na posição de *onset*, diferentemente do português, que, para esta posição, admite apenas duas, obedecendo à fonotática da língua, que impõe restrições de agrupamento. Assim, algumas combinações que são possíveis de realização no IA não o são no PB (como é o caso da fricativa desvozada [s] seguida de uma ou mais consoantes (<sp>, <st>, <str>, <spr>) em início de sílaba). Por esse motivo, o falante do PB tem a propensão de acrescentar a vogal anterior alta [i] cada vez que se depara com *clusters* consonantais inexistentes em seu sistema fonológico.

Em pesquisa variacionista para verificar incidência de vogal epentética (para desfazer encontros consonantais em palavras como *admitir*, *tecnologia*, *opção*, entre outras) no português, Collischonn (2003) observou algumas cidades do sul do Brasil e diferentes variáveis linguísticas a partir dos dados do Banco Varsul. Com relação às capitais, Curitiba, Florianópolis e Porto Alegre, a pesquisadora notou que o aumento da escolaridade implicou em taxa menor de epêntese. Os dados apresentados por Collischonn (2003) evidenciam que a consciência da forma escrita deve ter alguma influência sobre a variação. Para a autora, “esse é um caso em que não é a prescritividade escolar que está agindo, mas sim, o fato de que a forma escrita tem um impacto direto sobre a competência oral do falante que usa a escrita” (Collischonn, 2003, p. 291). Nas cidades de Panambi e Blumenau, a autora constatou uma mudança em curso, em direção a maior taxa de realização de epêntese. Como são cidades predominantemente bilíngues (português/alemão), uma das evidências a favor da posição da pesquisadora foi o fato de que as formas nativas da língua alemã foram realizadas, na maior parte das vezes, sem a epêntese. De

⁶ A vogal representada por [ɪ], em Cagliari (2007), corresponde a [i], no padrão atual do IPA.

acordo com Collischonn (2003, p. 291), Isto foi observado através de uma variável linguística que considerava se a palavra-alvo era uma palavra nativa ou um empréstimo. O resultado demonstrou que o falante bilíngue tem consciência da variação em estudo e é capaz de usar conscientemente as variantes dependendo da origem da palavra empregada. Na última cidade estudada pela autora, Flores da Cunha, que também tem falantes bilíngues (português/italiano), observou-se um maior índice de epêntese entre as mulheres.

Outro trabalho que observou o fenômeno de epêntese foi Schneider e Schwindt (2010). Neste estudo, os autores pesquisaram a realização da epêntese medial no PB e em inglês. Para a pesquisa, os autores levaram em consideração o grau de transparência dos prefixos, na intenção de testar a hipótese de Collischonn (2002) de que o número de ocorrências de epêntese após prefixos, como *sub-*, seja maior do que o número de ocorrências de epêntese em encontros consonantais localizados no interior de vocábulos não prefixados. A partir desta investigação, os autores confirmaram que, em PB, prefixos com grau elevado de transparência parem estar mesmo mais sujeitos à aplicação de epêntese, o que não pôde ser confirmado para o inglês, já que o contexto morfológico não foi uma variável estatisticamente relevante nas rodadas do trabalho. Com este trabalho, os autores chegaram à interessante conclusão de que há uma relação próxima entre os comportamentos da epêntese medial em PB e em língua inglesa, assim, para eles, há uma transferência linguística que opera na direção da L1 para a interlíngua.

4. Procedimentos metodológicos

Inicialmente, preocupamo-nos com a coleta dos dados que virão a ser descritos e analisados no presente estudo. Nosso *corpus* é constituído de nomes comerciais de empresas formados com palavras da língua inglesa concentradas no interior de São Paulo. A coleta dos dados foi realizada a partir do site Guia Mais⁷, que contém a informação comercial dos 26 estados do Brasil e do Distrito Federal. Entretanto, mesmo sendo possível pesquisar nomes de estabelecimentos comerciais existentes no país inteiro, optamos por fazer um recorte e realizar a pesquisa apenas no interior de São Paulo. A partir desse site, pudemos observar os nomes comerciais do interior de São Paulo⁸; além disso, as empresas estão separadas por categorias, o que facilitou a busca, a separação, a categorização e a quantificação dos dados. Consultamos

⁷ www.guiamais.com.br.

⁸ Observamos as cidades de Araraquara, Campinas, Jaú, Ribeirão Preto, São Carlos, Rio Claro e seus arredores.

um total de 7271 estabelecimentos cadastrados no interior de São Paulo, sendo que 862 deles têm algum elemento da língua inglesa, conforme podemos observar na tabela 1.

Tabela 1: Total dos nomes comerciais com elementos do inglês no PB separados por tipo de categoria comercial.

Total dos nomes com elementos do inglês no PB		
Categorias	n° de nomes coletados	% dos nomes coletados
Academias Desportivas	84	10%
Automóveis Peças e Serviços	67	8%
Cabeleireiros e Institutos de Beleza	207	24%
Informática – Equipamentos e Assistência	85	10%
Hotéis e Motéis	64	7%
Lavanderias	26	3%
Móveis	34	4%
Padarias e Confeitarias	17	2%
Restaurantes e Bares	104	12%
Roupas	174	20%
Total	862	100%

Consideramos nomes com elementos do inglês qualquer nome comercial que fosse totalmente em inglês, como no caso da academia *Lofty Sport*, ou que misturasse elementos ingleses com palavras do português, como *Aplausos Studio's Hair*. Como podemos ver através da tabela 1, consultamos nomes comerciais em dez setores do comércio: *Academias Desportivas*, *Automóveis Peças e Serviços*, *Cabeleireiros e Institutos de Beleza*, *Informática – Equipamentos e Assistência*, *Hotéis e Motéis*, *Lavanderias*, *Móveis*, *Padarias e confeitarias*, *Restaurantes e Bares* e, por fim, *Roupas*, sendo que encontramos nomes comerciais com elementos do inglês em todos os setores, conforme podemos observar na tabela 1.

A maior parte desses nomes comerciais no Brasil é ligada, respectivamente, ao setor de Cabeleireiros e Institutos de beleza (24% dos nomes comerciais – o que corresponde a 207 nomes); de Roupas (20% dos nomes comerciais – o que corresponde a 174 nomes) e, finalmente, de Restaurantes e bares (12% dos nomes comerciais – o que corresponde a 104 nomes comerciais). Nomes como *Libertway Motel*, *Click Computadores*, *Fast Printer*, *Baby Shop*, *Big Mix* e *Clean Plus* são alguns exemplos que se encontram em nosso *corpus*.

Além desses nomes, quantificamos também um total de 145 nomes comerciais com grafia estilizada. Como vimos na seção anterior, segundo Neves (1971, p. 105), alguns nomes comerciais podem se basear numa maneira bizarra, extravagante de escrever a palavra. No caso desta pesquisa, esta definição será usada para classificar nomes comerciais que fogem à grafia

da língua portuguesa, mas também não são ingleses, como, por exemplo, *Kool Dream* e *Autentyúnica*. Os nomes com grafias estilizadas foram quantificados, mas não foram aproveitados para as análises fonológicas, pois nosso interesse é investigar que tipos de adaptações fonológicas acontecem quando falantes do PB pronunciam palavras inglesas e não palavras cuja ortografia é inspirada em modelos estrangeiros. Mesmo tendo encontrado um bom número de nomes comerciais com elementos do inglês e também com grafia estilizada, constatamos que 6266 nomes, num total de 7271 estabelecimentos consultados, são formados, sobretudo, por palavras do português⁹ (como podemos ver na tabela 2).

Tabela 2: Total dos nomes comerciais coletados no PB separados por elementos linguísticos.

<i>Categorias</i>	<i>Nomes com grafia estilizada</i>	<i>Nomes com elementos do inglês</i>	<i>Outros nomes</i>	<i>Total dos estabelecimentos cadastrados</i>
Academias Desportivas	11	84	211	306
Automóveis - Peças e Serviços	9	67	611	687
Cabeleireiros e Institutos de Beleza	33	207	1115	1355
Informática - Equipamentos e Assistência	4	85	215	304
Hotéis e Motéis	2	64	386	452
Lavanderias	12	26	164	202
Móveis	8	34	614	656
Padarias e Confeitarias	8	17	882	907
Restaurantes e Bares	19	104	926	1049
Roupas	37	174	1142	1353
<i>Total</i>	143	862	6266	7271

Nossos dados do PB mostram que apenas uma parcela pequena dos nomes comerciais é composta por elementos da língua inglesa – 12% do total dos nomes consultados – e uma parcela ainda menor – apenas 2% dos nomes consultados – são formados por grafia estilizada. Desse modo podemos concluir que a maioria dos nomes comerciais do interior de São Paulo – 86% dos nomes comerciais consultados – é formada principalmente pela língua portuguesa, como podemos observar no gráfico 1.

⁹A categoria “Outros nomes” inclui todos os nomes que não são alvo desta pesquisa, isto é, nomes formados por palavras do português e formados com elementos de outras línguas. Como nesta pesquisa serão enfocados apenas os nomes comerciais com elementos da língua inglesa, não foi feita a quantificação dos nomes comerciais com elementos de outras línguas estrangeiras, entretanto, como pudemos observar ao longo da coleta de dados, estes nomes apareciam raramente, sendo mais comum o aparecimento de nomes comerciais formados somente com palavras do português.

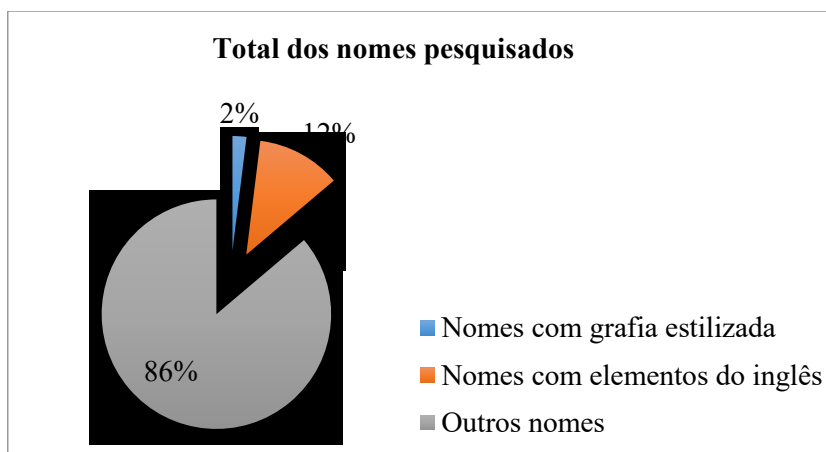


Gráfico 1: Porcentagem do total dos nomes comerciais pesquisados separados por tipos de elementos linguísticos.

Posteriormente à coleta dos dados, realizamos a preparação de um experimento de leitura que tem como intenção observar a pronúncia dos nomes comerciais por dois grupos de falantes de PB, compostos de homens e mulheres do interior de São Paulo com idades entre 18 e 38 anos: um grupo de cinco pessoas que nunca havia estudado inglês em escolas de idiomas e um grupo de cinco pessoas que estudavam inglês na mesma escola de idiomas há dois anos. Para facilitar a identificação da pronúncia de cada informante, utilizamos as abreviações PB-S-X para nos referirmos aos informantes do PB que não estudam inglês, PB-C-X para os informantes do PB que estudam inglês.

Como a produtividade do empréstimo de palavras do inglês em nomes comerciais é relativamente alta, para esse experimento, foi escolhida apenas uma amostra do *corpus*. Essa amostra é composta de 184 nomes comerciais com palavras da língua inglesa que mantêm sua grafia original e que pertencem a diferentes setores do comércio e serviços. Os nomes comerciais foram apresentados em contexto de língua portuguesa em frases criadas para este experimento, como “Eu frequento a *Academia Power Muscle Totonho*” e “Comprei na *All Car Peças*”.

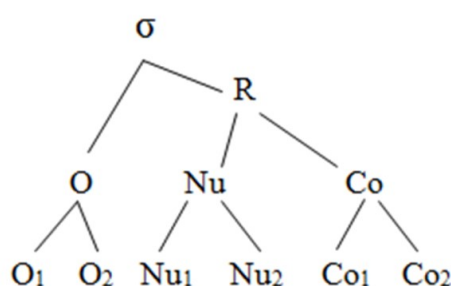
Como havia nomes comerciais com palavras repetidas, foram selecionados, para as análises dos processos fonológicos, apenas 40 nomes que apresentavam contextos que poderiam desencadear esses processos na pronúncia dos falantes de português – como, por exemplo, a palavra *resort*, do nome comercial *Quality Resort Centro de Convenções* em que, por conta da presença da consoante /t/ em posição final de sílaba, observamos um contexto para

a ocorrência da epêntese. Em seguida, foram feitas, de oitiva, as transcrições fonéticas dos nomes comerciais, de acordo com a pronúncia realmente realizada pelos falantes.

Os dados coletados foram transcritos – utilizando-se o Alfabeto Fonético Internacional (IPA) – e as pronúncias dos informantes do PB foram comparadas às pronúncias de um informante norte-americano gravado para esta pesquisa e às pronúncias registradas para o IA no dicionário *Longman Pronunciation Dictionary*. A partir dessa comparação, encontramos alguns casos de epêntese e apagamento¹⁰, que foram analisados a partir das teorias fonológicas não-lineares.

Com o desenvolvimento da fonologia não-linear, nas últimas décadas do século XX, surgiu a ideia de que o componente fonológico é caracterizado por um conjunto de sistemas hierarquicamente organizados que interagem, sendo cada um governado por seus próprios princípios. Dessa forma, a fonologia não-linear trata o componente fonológico como sendo um sistema heterogêneo. Nos modelos fonológicos não-lineares, a sílaba é o ponto central das discussões, o que é muito importante na observação dos estrangeirismos já que uma das principais adaptações que vem sendo observada na passagem do IA para o português é o processo de ressilabação. Para Blevins (1995, p. 209-210), “*in a number of languages, native speakers have clear intuitions regarding the number of syllables in a word or utterance, and in some of these, generally clear intuitions as to where syllable breaks occur*”¹¹. Assim, para o desenvolvimento deste trabalho, a estruturação das sílabas será representada em forma de árvore, como se pode ver a seguir (adaptada de Selkirk, 1980, p. 6 e Hogg e McCully, 1987, p.36):

(1)



¹⁰ É importante mencionar que também foram encontrados outros fenômenos fonético-fonológicos que não são alvo desta pesquisa, como, por exemplo, a palatalização, a vocalização e o deslocamento de acento.

¹¹ “Em várias línguas, os falantes nativos têm intuições claras sobre o número de sílabas que existem em uma palavra ou enunciado e, em alguns deles, os falantes nativos geralmente têm intuições claras quanto ao local onde ocorrem as quebras silábicas” (tradução nossa).

Neste esquema, o símbolo σ representa a própria sílaba enquanto unidade; **O** significa “onset” (ou “ataque”), que é o início da sílaba; **R** é a “rima”, a parte da sílaba que constitui as vogais da sílaba e todas as consoantes que a(s) segue(m); **Nu** é o “núcleo”, ou seja, o ponto mais forte, mais proeminente, normalmente é a vogal (ou as vogais, no caso de ditongos); **Co** significa “coda”, formada pelas consoantes que vêm depois da(s) vogal(ais) do núcleo, ou seja, aquelas consoantes que não pertencem ao “onset” da sílaba.

5. Análise dos casos de epêntese

Primeiramente, analisaremos os casos de epêntese que ocorreram na pronúncia dos nomes comerciais com elementos do inglês no contexto de língua portuguesa. Neste trabalho, identificamos a epêntese em duas formas de realização, inicial e final. Não observamos casos de epêntese medial. Em PB, os casos de prótese foram motivados pelo não licenciamento de *onsets* compostos pelas sequências <sp> (palavras *sport*¹² e *speed*) e <str> (palavra *street*) e foram realizados na maioria das pronúncias dos informantes (apenas uma realização no PB-S e duas no PB-C na palavra *street* não apresentaram esse fenômeno).

Já os casos de paragoge (no sentido de epêntese em final de palavra, não rítmica) em PB aconteceram em virtude do não licenciamento de oclusivas (/t/ - *sport*, /d/ - *speed*, /p/ - *stop*, /g/ - *bag*, /k/ - *tech*), fricativas (/f/ - *beef*¹³, /ʃ/ - *brush*) e nasal alveodental (/n/ - *design*)¹⁴ na posição de coda no português. A vogal [i], acrescentada por epêntese, desencadeou um processo de ressilabação das palavras, uma vez que a vogal epentética passa a ser núcleo de uma nova sílaba, com podemos ver nos exemplos (5), (6), (7) e (8).

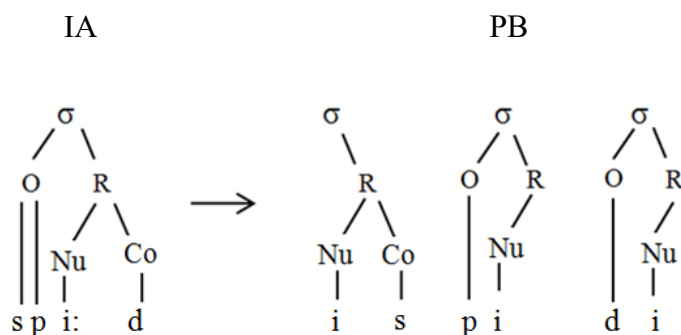
¹² A palavra *sport*, enquanto nome comum, já está ortograficamente adaptada ao português *esporte*. Neste trabalho, observamos apenas anglicismos não adaptados ortograficamente ao PB.

¹³ A palavra *beef*, enquanto nome comum, já está ortograficamente adaptada ao português *bife*.

¹⁴ Câmara Jr (2004[1970], p.52) considera como sendo apenas quatro as consoantes portuguesas possíveis em posição posvocálica, a saber /S/, /N/, /l/ e /R/.

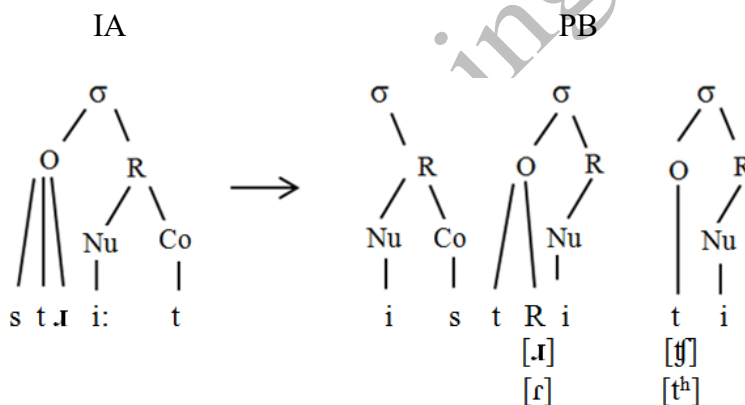
(5)

Speed



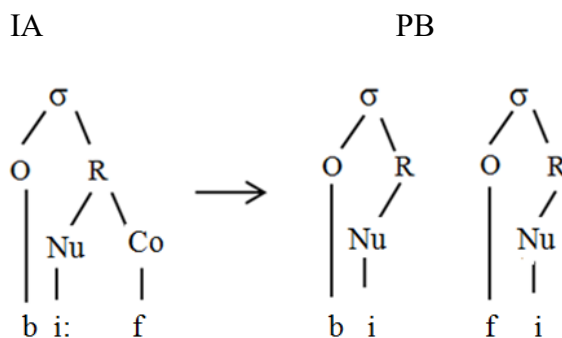
(6)

Street

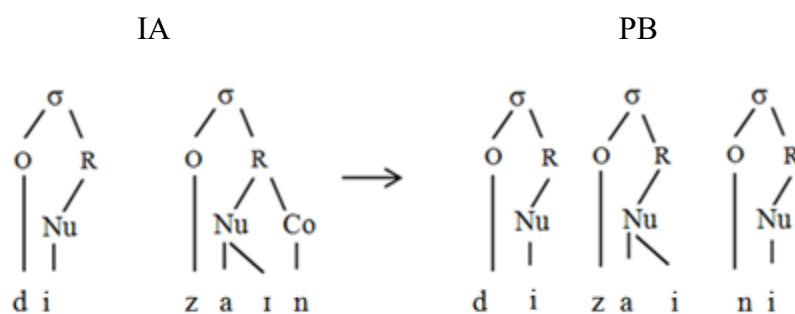


(7)

Beef



(8)

Design

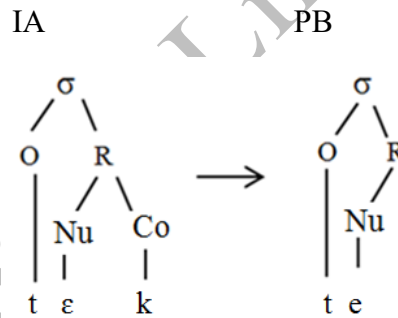
Assim, as palavras *speed* e *street*, que tinham apenas uma sílaba em IA, passaram a ter três sílabas em português, já que houve contexto para duas ocorrências de epêntese em cada palavra; a palavra *beef*, que também tinha uma sílaba em IA, passou a duas sílabas em português; e, por fim, a palavra *design*, que tinha duas sílabas em IA, passou a ter três em português. Com relação à palavra *design*, só há contexto para a epêntese porque o /n/ final não nasalizou as vogais anteriores na pronúncia de alguns sujeitos da pesquisa. Entretanto, em algumas realizações de informantes do PB-C, não ocorre epêntese no caso da palavra *design*, que passa a ser pronunciada com o /i/ final nasalizado \tilde{i} .

Notamos que os casos observados para o PB na pronúncia de nomes comerciais são semelhantes aos casos de epêntese mapeados na adaptação de nomes comuns, estudados por Assis (2007), e de nomes próprios, observados por Massini-Cagliari (2010, 2011a,b, 2013) e Souza (2011). Em relação aos nomes comuns, Assis (2007, p. 154) constatou a ocorrência da epêntese para “resolver” sílabas não canônicas em PB, como no caso da palavra *spam*, um monossílabo em IA que passa a ter duas sílabas em PB após a epêntese (/spæm/ em IA e /iS.paN/ em PB). O mesmo foi constatado em relação aos nomes próprios; assim, o nome *Robert* pode ser adaptado para [hɔ.bɛr.tʃi], com três sílabas e uma vogal epentética no fim da palavra (MASSINI-CAGLIARI, 2011a, p. 62), e o nome *Steve* que, ao ser adaptado ao PB, também fica com três sílabas após duas epênteses (/iS.ti.vi/) (SOUZA, 2011, p. 195).

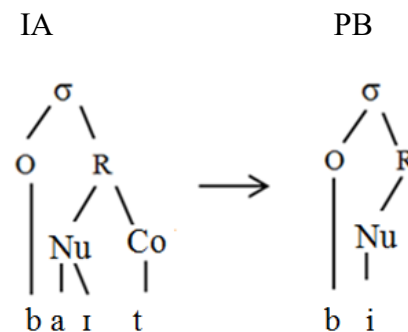
6. Análise dos casos de apagamento

Já os casos de apagamento (processo que também é conhecido como queda, eliminação ou truncamento) ocorrem quando “há a supressão de um segmento da forma básica de um morfema” (CAGLIARI, 2002b, p.101). Assim como a epêntese, o apagamento ocorre para “resolver” sílabas anômalas. Entretanto, esses dois processos atuam em direções contrárias, uma vez que a epêntese acrescenta elementos e o apagamento os suprime. Com relação às pronúncias dos nossos informantes do PB, encontramos dois casos de apagamento¹⁵ nas realizações dos informantes PB-S, como podemos observar nos exemplos a seguir. Os informantes PB-C não realizaram apagamentos, talvez já por conhecerem um pouco a estrutura da língua inglesa.

(9)

Tech

(10)

Bite

¹⁵ Houve outros fenômenos, relativos à realização das vogais, embora o trabalho não tenha por objetivo analisá-los.

Nos exemplos (9) e (10), vemos que houve o apagamento do /k/ final da palavra *tech*, do nome comercial *Car Tech*, ([ka:ɪ'tɛk] no IA e ['kaɪ'te] no PB) e do /t/ final da palavra *bite*, do nome comercial *Bytecenter* ([ˈbaɪt, sɛnrɪ] no IA e ['biːsẽ̃tɛɪ] no PB, pois o português não permite esses sons em posição de coda silábica.

Nesta pesquisa, essas realizações não foram comuns, sendo mais corriqueira¹⁶ a pronúncia desses nomes com a ocorrência de epêntese, ([ˈbaɪt h̥ɪˈsẽ̃tɛɪ] e ['kaɪ'tɛkɪ], respectivamente). Além desses exemplos, observamos também um único caso de apagamento da consoante nasal [n] no fim da palavra *seventeen* (em IA [sɛvən'ti:n]) que foi realizada [ˈsiʋvẽ̃t̪h̥ɪ] por um informante PB-S.

Nas pesquisas de Assis (2007, p. 181), Souza (2011, p. 210) e Massini-Cagliari (2010, p. 168), as autoras também observaram casos de apagamento. Com relação aos nomes comuns, Assis observou o apagamento do /t/ final, como na palavra *impeachment*, de modo que a sílaba final desta palavra, que em IA era CVCC (um padrão possível no PB, mas não com os segmentos /Nt/ na posição de coda), é re-estruturada, tornando-se CVC, com a nasal /N/ na coda.

Em relação aos nomes próprios, Souza comenta três casos de apagamento: o apagamento do /k/ do antropônimo *Victor*, dessa forma, a sílaba [vɪk], do IA, tornou-se [vi], assim, com o apagamento do segmento /k/ da coda, houve a transformação de uma sílaba marcada, CVC, em uma não marcada, CV, que é o padrão do PB; e dois apagamentos em *Herbert*, que foi pronunciado [ɛbeɦt̪ɪ] pelo informante de sua pesquisa, evidenciando o apagamento do [h] inicial (compreensível já que o <h> é sempre zero fonético nesse contexto em PB) e o [ɪ] da coda da sílaba <her>, pois sílabas marcadas por consoante retroflexa na posição de coda não obedecem ao padrão fonético da variedade estudada pela autora (PB de São Luís – MA).

Por fim, dois exemplos de apagamentos analisados por Massini-Cagliari (2010) ocorrem no antropônimo *Irso* (de Wilson), em que, na sílaba inicial, se verifica o apagamento da semivogal do *onset* e, na sílaba final, ocorre também o apagamento da nasal.

¹⁶ Apenas um informante pronunciou essa palavra com apagamento, os outros fizeram uma epêntese.

7. Considerações finais

Como vimos, o estudo dos nomes comerciais pode ser de interesse para várias disciplinas, além da linguística, e ainda desperta a curiosidade das pessoas de modo geral. Neste trabalho, vimos que é importante ter um nome comercial que chame a atenção do público consumidor, assim, no meio comercial, o inglês acaba sendo visto como uma forma interessante de valorização do produto que está sendo vendido. Desse modo, os comerciantes acham mais atrativa uma padaria chamada *Big Pão* do que, simplesmente, *Grande Pão* – talvez numa tentativa de associar aquele estabelecimento a um público-alvo formado de pessoas de maior poder aquisitivo.

A pesquisa realizada neste estudo observa as adaptações fonológicas que ocorrem na pronúncia de nomes comerciais com elementos do inglês por falantes de PB, focando especificamente nos casos de epêntese e apagamento. Notamos com esse estudo que o PB tende a realizar epênteses e apagamentos para “resolver” as sílabas inglesas que não são possíveis nessa língua. Entretanto, observamos que o apagamento ocorreu mais raramente (apenas duas vezes), sendo mais corriqueira a pronúncia dos nomes comerciais do nosso *corpus* com a ocorrência de epêntese.

De modo geral, as análises das adaptações fonológicas de estrangeirismos presentes no léxico comercial corroboram os resultados alcançados por outros autores em pesquisas anteriores (FREITAS, 1984,1992; FREITAS et al. 2003; FREITAS; NEIVA, 2006; ASSIS, 2007; SOUZA, 2011) de que as primeiras adaptações que acontecem com anglicismos, quando realizados por falantes de português, ocorrem no nível fonético-fonológico, sendo influenciadas, sobretudo, pelas distinções existentes entre os inventários fonético-fonológicos dessas duas línguas.

Referências Bibliográficas

ASSIS, A. B. G. de. **Adaptações fonológicas na pronúncia de estrangeirismos do inglês por falantes de Português Brasileiro**. Dissertação (Mestrado em Linguística e Língua Portuguesa) - Faculdade de Ciências e Letras, Universidade Estadual Paulista Júlio de Mesquita Filho, Unesp, Araraquara, 2007.

BISOL, L. O ditongo na perspectiva da fonologia atual. **D.E.L.T.A.**, São Paulo, v. 5, n. 2, 1989. p. 185-224.

CAGLIARI, L. C. **Elementos de fonética do português brasileiro**. São Paulo: Paulistana, 2007.

CAGLIARI, L. C.. **Fonologia do Português - Análise pela Geometria de Traço e pela fonologia lexical (Parte II)**. Campinas: edição do autor, 1999.

CÂMARA JR., J. M. **Problemas de linguística descritiva**. Petrópolis: Vozes. 2010 [1ª ed. 1969].

CARVALHO, N. **Empréstimos linguísticos na língua portuguesa**. São Paulo: Cortez, 2009.

COLLISCHONN, G. A epêntese vocálica no português do sul do Brasil: análise variacionista e tratamento pela teoria da otimidade. **Letras de Hoje**, v. 35, n. 1, p. 285-318, 2000.

COLLISCHONN, G. Epêntese vocálica no português do sul do Brasil: variáveis extralinguísticas. **Revista Letras**, nº. 61, especial, p. 285-297, 2003.

COLLISCHONN, G. A sílaba em português. In: BISOL, L. (Org.). **Introdução a estudos de fonologia do português brasileiro**. 4ª. ed. Porto Alegre: EDIPURCS, 2005. p. 101-133. [1ª. ed. em 1996].

CRYSTAL, D. **English as a Global Language**. Cambridge: Cambridge University Press, 2003.
crossref <http://dx.doi.org/10.1017/CBO9780511486999>

FARACO, C. A. Guerras em torno da língua – Questões de política linguística. In FARACO, C. A. (Org.) **Estrangeirismos: guerras em torno da língua**. 3ª. ed. São Paulo: Parábola, 2004. p. 9-14.

FIORIN, J. L. Considerações em torno do Projeto de Lei nº 1676/99. In: FARACO, C. A. (Org.) **Estrangeirismos: guerras em torno da língua**. 3. ed. São Paulo: Parábola, 2004. p. 107-125.

FREITAS, M. A. de. Empréstimos, teoria auto-segmental e abertura vocálica. **Cadernos de Estudos Linguísticos**. Campinas: 1992, p. 71-81.

FREITAS, M. A. de. **Os Empréstimos do Inglês no Português do Brasil**. Dissertação (Mestrado em Linguística) - Faculdade de Letras da Universidade Federal do Rio de Janeiro, UFRJ, Rio de Janeiro, 1984.

FREITAS, M. A. de; NEIVA, A. M. S. Estruturação silábica e processos fonológicos no inglês e no português: empréstimos e aquisição. **Revista Virtual de Estudos da Linguagem (ReVEL)**. Ano 4, n. 7, agosto de 2006. Disponível em: <http://paginas.terra.com.br/educacao/revel/>. Acesso em 30 jan. 2010. p. 1-27.

GARCEZ, P. M.; ZILLES, A. M. S. Estrangeirismos: desejos e ameaças. In: FARACO, C. A. (Org.) **Estrangeirismos: guerras em torno da língua**. 3. ed. São Paulo: Parábola, 2004. p. 15-36.

GARCEZ, P. M.; ZILLES, A. M. S. Estrangeirismos: desejos e ameaças. In: FARACO, C. A. (Org.) **Estrangeirismos: guerras em torno da língua**. 3ª. ed. São Paulo: Parábola, 2004. p. 15-36.

GIACOMINI FILHO, G.; BORBA, O. R. Nomes de Marca: aspectos mercadológicos e onomásticos. In: **Revista Eletrônica Signos do Consumo**. Vol. 2. nº. 2. São Paulo: Universidade de São Paulo, 2010. p. 217-229. **crossref** <http://dx.doi.org/10.11606/issn.1984-5057.v2i2p217-229>

HOGG, R.; McCULLY, C. **Metrical Phonology: a coursebook**. Cambridge: Cambridge University Press, 1999. [1ª ed. 1987]

LEE, S.-H.. Epêntese no Português. In.: **Estudos Linguísticos XXII – Anais dos Seminários do GEL**. Ribeirão Preto: Instituição Moura Lacerda, 1993. Vol. II, p.847-854.

LOPEZ, B. **The sound pattern of Brazilian Portuguese (Cariocan dialect)**. Tese (doutorado, PhD) – Los Angeles: University of California, 1979.

MACEDO, N. Z. **Análise fonológica de nomes próprios de origem estrangeira e novas criações em Português Brasileiro**. 171 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras (Campus de Araraquara), Araraquara, 2015.

MASSINI-CAGLIARI, G. Changing Attitudes: Ways of Phonologically Adapting Proper Names in Archaic Brazilian and European Portuguese. In: KOLOKONTE, M.; JANKE, V. (Org.). **Interfaces in Language 3**. 1ª ed. v. único. Newcastle upon Tyne: Cambridge Scholars, 2013, p. 159-178.

MASSINI-CAGLIARI, G. Loans and foreign first names as clues to Phonological Identity in Brazilian Portuguese. In: HORNSBY, D. (Org.). **Interfaces in language 2**. 1ª ed. v. 1. Newcastle upon Tyne: Cambridge Scholars, 2011a. p. 53-67

MASSINI-CAGLIARI, G. Adaptação fonológica de nomes próprios de origem estrangeira: comparação entre Português Arcaico e Português Brasileiro. **Estudos Linguísticos**, São Paulo, 40 (2): p. 795, mai.-ago. de 2011b.

MASSINI-CAGLIARI, G. Discutindo questões de identidade a partir da (não) adaptação fonológica de nomes próprios de origem estrangeira no Brasil. In: NEVES, M. H. de M. **As interfaces da gramática**. Araraquara: FCL-UNESP Laboratório Editorial; São Paulo: Cultura Acadêmica, 2010. (Série Trilhas Linguísticas 18). p.73-90

NEVES, V. L. H. **Nomes próprios comerciais e industriais no português: um aspecto da nomenclatura do comércio e da indústria em Belo Horizonte**. Tese (Doutorado em Letras) - Faculdade de Letras da Universidade Federal de Minas Gerais, UFMG, Belo Horizonte: 1971.

PAIVA, V. L. M. de O. e. **A língua inglesa enquanto signo na cultura brasileira**. Tese (Doutorado em Linguística) – Faculdade de letras da Universidade Federal do Rio de Janeiro, UFRJ, Rio de Janeiro: 1991.

PRADO, N. C. **A influência da língua inglesa na formação de nomes comerciais: questões de identidade linguística e cultural**. Tese (Doutorado em Linguística e Língua Portuguesa) -

Faculdade de Ciências e Letras, Universidade Estadual Paulista Júlio de Mesquita Filho, Unesp, Araraquara, 2014.

ROACH, P. **English Phonetics and Phonology: a practical course**. Cambridge: Cambridge University Press. 1998.

SCHNEIDER, André; SCHWINDT, Luiz Carlos. A epêntese vocálica medial em PB e na aquisição de inglês como LE: uma análise morfofonológica. **Letras de Hoje**, v. 45, n. 1, p. 1626, 2010.

SERRA, M.; NODARI, J. I. Os Anglicismos nas fachadas de estabelecimentos comerciais da Rua Grande no centro de São Luís – MA: ampliação do léxico ou ameaça à hegemonia da Língua Portuguesa? In. **Littera on line**. Vol. 2. nº. 4. São Luís: Universidade Federal do Maranhão, 2011. Disponível em <http://www.periodicoseletronicos.ufma.br/index.php/littera/article/view/769/484>. Acesso em 10 de abr. de 2014.

SOUZA, S. M L. S. de. **Antropônimos de origem inglesa: adaptações ortográficas e fonético-fonológicas realizadas por falantes do português brasileiro de São Luís-MA**. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Letras da Universidade Estadual Paulista Júlio de Mesquita Filho, FCL/UNESP, Araraquara: 2011.

TRASK, R. L. **Dicionário de linguagem e linguística**. Tradução e adaptação de Rodolfo Ilari. São Paulo: Contexto, 2004.

WELLS, J. C. **Longman pronunciation dictionary**. 3rd ed. Edinburg Gate: Pearson, 2008.

ZUCARELLI, F. E. **Ditongos e hiatos nas cantigas medievais galego-portuguesas**. Dissertação (Mestrado em Linguística e Língua Portuguesa) – FCL/UNESP, Araraquara, 2002.

Artigo recebido em: 29.09.2015

Artigo aprovado em: 26.03.2016

Uma discussão acerca da aplicação do *Perceptual Assimilation Model-L2* à percepção fônica de língua estrangeira: questões de pesquisa e desafios teóricos

Discussing the application of the *Perceptual Assimilation Model-L2* to foreign language speech perception: research enquiries and theoretical challenges

Reiner Vinicius Perozzo*
Ubiratã Kichhöfel Alves**

RESUMO: Este trabalho tem o objetivo de apontar possíveis alterações ao *Perceptual Assimilation Model-L2* [*PAM-L2* (BEST; TYLER, 2007)], modelo de percepção da fala em segunda língua, para que o mesmo dê conta da percepção fônica em contexto de línguas estrangeiras. Para tanto, apresentamos o *PAM-L2* e suas diretrizes, visitamos alguns estudos nacionais que se utilizaram do modelo e problematizamos suas bases fonológica e filosófica. Neste sentido, advogamos pela instauração do gesto acústico-articulatório concebido por Albano (2001) como unidade perceptual e defendemos o realismo indireto como teoria filosófica para acomodar a percepção fônica no cenário estrangeiro da aprendizagem perceptual. Concluímos nossa exposição levantando alguns pontos teóricos e metodológicos que merecem atenção quando tratamos da percepção fônica em língua estrangeira, fundamentais ao modelo em sua totalidade.

PALAVRAS-CHAVE: Ecologia da percepção. Gesto articulatório. *PAM-L2*. Percepção fônica estrangeira.

ABSTRACT: This article aims to suggest some changes to the *Perceptual Assimilation Model-L2* (*PAM-L2* (BEST; TYLER, 2007)), a model of second language speech perception, so that it may account for speech perception of foreign languages. In order to do so, at first we present *PAM-L2* and its guidelines, then we list some research studies carried out in Brazil that have utilized the model, and finally we problematize its phonological and philosophical bases. In this regard, we claim for the use of the articulatory gesture conceived by Albano (2001) as the perceptual unit of speech and we suggest that indirect realism may be suitable as a philosophical theory to encompass speech perception in foreign settings of perceptual learning. We conclude our discussion raising some other relevant theoretical and methodological issues, fundamental to the model as a whole.

KEYWORDS: Articulatory gesture. Foreign speech perception. *PAM-L2*. Perceptual ecology.

1. Introdução

Os estudos de percepção dos sons da fala, especialmente no que diz respeito àqueles que veiculam contrastes fonológicos, tiveram seu início nos Laboratórios *Haskins*, nos Estados

* Mestre em Letras. Instituto de Letras. UFRGS.

** Doutor em Letras. Instituto de Letras. UFRGS.

Unidos, por volta de 1950 (NISHIDA, 2012). Estudiosos como Alvin Liberman, Pierre Delattre e Franklin Cooper (DELATTRE; LIBERMAN; COOPER, 1955) se debruçavam sobre experimentos que continham sequências sonoras representativas de sílabas CV do inglês e, através de métodos físicos, alteravam suas configurações formânticas, a fim de promover diferenças acústicas cujos resultados implicavam distinções de ponto de articulação para segmentos oclusivos.

Com a evolução das pesquisas sobre percepção fônica, cresciam, também, as discussões sobre quais eram as bases das informações sonoras que refletiam as diferenças lexicais, causando uma oposição entre a sua natureza acústica e a sua natureza articulatória, conforme indica Nishida (2012). Neste cenário, encontram-se, de um lado, evidências fonológicas com base em propriedades acústicas do sinal de fala, observadas por Ohala (1996), e, de outro, fatos articulatórios veiculados através da percepção e da produção da fala, sustentados por Fowler (1996)¹. Além disso, aliada à busca de compreender as propriedades que marcariam fronteiras perceptuais dos sons da fala, havia a necessidade de modelar formalmente como os indivíduos, ao adquirir seu idioma materno, percebiam auditivamente seus elementos sonoros. Tal modelagem, alavancada pela busca dos primitivos referentes à percepção da fala em língua materna, se estendeu, posteriormente, a fenômenos em línguas não nativas.

Ao longo dos últimos anos, verificamos a preponderância de três modelos perceptuais da fala não nativa no âmbito das pesquisas nacionais: o *Speech Learning Model* [*SLM*, Modelo de Aprendizagem da Fala (FLEGE, 1995)], o *Perceptual Assimilation Model* [*PAM*, Modelo de Assimilação Perceptual (BEST, 1995)] e o *Perceptual Assimilation Model of Second Language Speech Learning* [*PAM-L2*, Modelo de Assimilação Perceptual da Aprendizagem da Fala em Segunda Língua (BEST; TYLER, 2007)]². O *SLM* trata tanto da percepção como da produção de sons não nativos, habilidades balizadas pela experiência linguística do aprendiz, além de variáveis relacionadas à idade de aprendizagem do sistema sonoro alvo e ao período de chegada à comunidade geográfica em que o idioma alvo é falado. O *PAM* volta-se aos chamados “*naïve listeners*”, isto é, ouvintes sem experiência com outro idioma, e parte do

¹ Citam-se, aqui, os trabalhos do ano de 1996, especialmente, por terem sido publicados no mesmo periódico científico, representando, portanto, o confronto direto entre ambos os estudiosos. Todavia, cabe lembrar que tal embate teórico iniciou anos antes da referida publicação, a exemplo da proposta de Fowler (1986).

² A partir deste ponto, no presente texto, quando mencionarmos os modelos perceptuais elencados, iremos suprimir sua referência autoral para fins de delimitação. Esclarecemos, portanto, que, ao tratarmos do *SLM*, do *PAM* e do *PAM-L2*, estamos nos referindo a Flege (1995) para o primeiro, a Best (1995) para o segundo e a Best e Tyler (2007) para o último.

pressuposto de que a língua materna influencia a percepção da fala não nativa, além de evidenciar as possibilidades de assimilação perceptual de elementos sonoros que não fazem parte do sistema materno dos ouvintes. O foco do *PAM-L2* (que é, por sua vez, uma extensão do *PAM*) recai sobre aprendizes de uma segunda língua (L2) que estão adquirindo o sistema fônico alvo, e seu postulado fundamental é o de que a aprendizagem perceptual é determinada por princípios diferentes daqueles do idioma materno.

No presente artigo, daremos especial atenção ao *PAM-L2* em virtude de sua crescente utilização nos trabalhos em contexto de L2, de seu compromisso com o contexto da percepção de sons não nativos, de seu robusto referencial teórico (tanto no que se refere ao entendimento filosófico da percepção da fala como no que tange à adoção clara de um primitivo fônico, o qual se mostra condizente com nossa perspectiva de aquisição de linguagem) e de suas predições quanto às possibilidades de discriminação categórica acerca das unidades sonoras veiculadas na fala. Entendemos, nesta direção, que seja necessário trazer à tona alguns pontos intrigantes a seu respeito, os quais se instanciam no âmago de sua proposta e, necessariamente, advêm da postura teórica de seu predecessor, o *PAM*. Por conseguinte, trataremos de um questionamento pertinente aos estudos que adotam o *PAM-L2* como fundamentação teórica, seja para a condução da pesquisa em si, seja para a análise e a interpretação dos resultados que obtêm, a saber: *quais são as modificações formais necessárias no modelo para sua aplicação a contextos de percepção fônica de uma língua estrangeira (em detrimento de L2)?*

Para responder a tal questionamento, julgamos relevante resenhar o modelo de percepção fônica que estamos discutindo, o *PAM-L2*, elencar alguns estudos que o utilizaram no âmbito nacional e, por fim, problematizar suas assunções teóricas na medida em que o contexto de percepção em língua estrangeira (LE) se instaura.

2. Percepção fônica de elementos sonoros não nativos: o *PAM-L2*

Primeiramente, podemos dizer que o *PAM-L2* (assim como o *PAM*, seu predecessor) advoga pelo gesto articulatório como primitivo fônico e pelo realismo direto como base filosófica. No arcabouço teórico de Best (1995) e de Best e Tyler (2007), tanto a fonética como a fonologia são fundamentadas no domínio dos gestos articulatórios, ou seja, existe um domínio gestual comum tanto para o material fonético como para a representação fonológica dos sons da fala. Duas constatações importantes acerca do modelo são as seguintes: a primeira delas é a de que *fonética* e *fonologia* não são interpretadas como dois domínios informacionais

separados, e a segunda prevê que propriedades fonéticas se encontram alocadas na estrutura fonológica.

Em 2007, Catherine Best e Michael Tyler lançaram uma versão expandida do *PAM*, a qual se ocupa da percepção de sons não nativos por aprendizes de L2. A esta versão estendida do modelo foi atribuído o nome de *PAM-L2*. A tarefa básica do *PAM-L2* é estabelecer que o aprendizado perceptual da L2 seja determinado por princípios não nativos de percepção de fala, considerando elementos comuns e complementares entre ouvintes sem experiência com a segunda língua e aprendizes de tal idioma. O modelo ainda se volta para a relação entre o aprendizado de uma língua e sua informação fonética e fonológica, e sugere brevemente como a experiência linguística – sendo monolíngue ou multilíngue – influencia a percepção fônica.

Na concepção de Best e Tyler (2007), os usuários da L2 estão sempre aprendendo o sistema alvo, sendo que existem diferenças qualitativas entre aprendizes que estão ativamente adquirindo a L2 e aprendizes em processo mais estável. Sob a ótica de Antoniou, Tyler e Best (2012), o principal foco do *PAM-L2* são os aprendizes que estão constantemente aprendendo a L2, além de fatores fundamentais para a formação das categorias fonético-fonológicas desse sistema, como a idade com que o aprendiz inicia sua aquisição, o tempo de residência no país da L2 e o uso da L2 (PEROZZO, 2013).

Destacamos que, no *PAM-L2*, os ouvintes estão sujeitos a perceber diferenças linguísticas tanto em relação ao detalhe fonético de caráter contrastivo (em termos de distintividade entre dois itens lexicais), como também no que tange aos aspectos variáveis (não distintivos) dentro das categorias das vogais e das consoantes. Convém, ainda, ressaltar que os ouvintes podem mostrar sensibilidade perceptual quanto ao aspecto fônico variável tanto na fala nativa como na fala em L2. Portanto, a percepção de elementos fônicos variáveis não só é prevista pelo *PAM-L2*, como também é consistente com seus princípios.

O *PAM-L2* prevê que os fones da L2 são assimilados, primeiramente, em categorias da L1 (língua materna) já existentes, e, então, tornam-se estabelecidos como novas categorias. Tal processo ocorre, em um primeiro momento, no nível fonético, e, na medida em que o vocabulário da L2 se expande, os aprendizes se adequam à fonologia do idioma-alvo, possibilitando que os fones sejam discriminados com base em diferenças de categorias que são lexicalmente relevantes (ANTONIOU; TYLER; BEST, 2012).

Algo interessante e que merece atenção é o fato de haver no *PAM-L2* a existência de uma categoria fonológica e uma categoria fonética, as quais se alinham aos aspectos funcionais

no sistema, assumidos pelos gestos fônicos. A categoria fonológica “diz respeito à informação da fala que é relevante para diferenças lexicais mínimas em uma determinada língua”, e a categoria fonética representa as “relações gestuais invariantes que são sistemática e potencialmente perceptíveis aos ouvintes com percepção mais fina, como alofones posicionais ou diferentes realizações de uma categoria fonológica entre dialetos ou línguas” (BEST; TYLER, 2007, p. 25). Assim sendo, a estrutura fonética serve de base para a organização fonológica, sendo que fonética e fonologia fazem parte de um único sistema, o qual possui um nível mais baixo e um nível mais alto, mas se instanciam, representacionalmente, através do gesto articulatório. Desta maneira, o modelo reconhece tanto a sensibilidade aos detalhes gradientes e físicos do nível fonético como a categorização fonológica, de caráter abstrato, referente à percepção de distinções lexicais, aspectos esses que são abarcados pelo primitivo gestual.

Segundo Best e Tyler (2007), o *PAM-L2* se sustenta sobre quatro possíveis casos de contrastes dos sons da L2 que os aprendizes percebem como segmentos da fala:

- (a) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe uma delas como sendo equivalente a determinada categoria fonológica da L1. Quanto à outra categoria a ser assimilada, é possível que ela seja alocada de maneira adequada no sistema da L1, ou seja, ela representa um bom exemplar para outra categoria da L1, ou, diferentemente, é possível que ela não seja categorizada. Dessa forma, o aprendiz praticamente não teria dificuldades em discriminar, minimamente, os segmentos-alvo que contrastam em pares mínimos.
- (b) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe as duas como pertencentes à mesma categoria fonológica da L1. No entanto, uma das categorias é percebida como mais desviante do que a outra. Espera-se, portanto, que os aprendizes sejam capazes de discriminar relativamente bem estes dois fones da L2, mas não tão bem como a discriminação que poderia haver caso eles conseguissem classificar os segmentos em categorias diferentes.
- (c) Frente a um contraste entre duas categorias fonológicas da L2, o aprendiz percebe as duas como pertencentes à mesma categoria fonológica da L1. Todavia, ambas são classificadas como exemplares igualmente bons ou igualmente desviantes daquela categoria. O aprendiz terá problemas ao discriminar estes dois fones, que seriam

assimilados tanto fonética como fonologicamente a uma única categoria da L1, e duas palavras da L2 minimamente contrastantes seriam percebidas como homófonas.

- (d) Se o ouvinte não conseguir encontrar uma categoria da L1 para alocar dois segmentos contrastantes da L2, então não haverá assimilação fonológica da L2 para a L1. Neste sentido, o ouvinte reconhece uma mistura de similaridades entre os segmentos e não é capaz de categorizá-los devidamente.

Best (1995) e Best e Tyler (2007), a fim de situar filosoficamente seus modelos de percepção da fala, lançam mão de uma teoria perceptual defendida por James Gibson (um psicólogo experimentalista americano), denominada *Realismo Direto*. Basicamente, o livro de Gibson (1986), *The Ecological Approach to Visual Perception*, é um tratado sobre a percepção visual (em detrimento de outros sentidos), e o autor se preocupa em descrever a realidade circundante ao percebedor, utilizando-se de uma nomenclatura própria e procurando evidenciar supostas lacunas nas teorias mais tradicionais da ótica.

Nas primeiras páginas de sua obra, Gibson (1986) deixa claro que seu objeto de estudo refere-se ao nível ecológico, ao habitat dos animais e dos homens, cujos sistemas perceptuais não são capazes de detectar extremos, como átomos ou galáxias. Para o psicólogo, devemos considerar como percebemos o ambiente, ou seja, como apreendemos as mesmas coisas que nossos ancestrais humanos apreenderam antes mesmo de saberem sobre partículas atômicas e sistemas planetários. Por conseguinte, Gibson (1986) preocupa-se com a percepção direta e não com aquela indireta, a qual é obtida através de microscópios, telescópios, fotografias e figuras, e muito menos com o tipo de apreensão (de conhecimento) obtida pela fala e pela escrita (p. 10). Esta é a primeira acepção do termo “direto” na obra de Gibson, que *não prevê qualquer aparato entre o percebedor e o ambiente*. A sua segunda acepção diz respeito ao fato de os percebedores não necessitarem de uma quantidade excessiva de aprendizagem (p. 143), ou seja, *eventos cognitivos que se relacionam a inferências e a representações mentais são ignorados na perspectiva direta da percepção*.

Um conceito notável na obra de Gibson (1986) é o de “*affordances*”, ou concessões³, que se refere, *grosso modo*, àquilo que o meio nos concede, ou seja, nos oferece. Para o

³ Tradução nossa.

estudioso, uma concessão é invariante e se faz constante desde o início da evolução animal, além de implicar a complementaridade entre animal e ambiente. O fogo, por exemplo, concede calor em uma noite fria, mas também oferece a possibilidade de alguém se queimar; um objeto que se aproxima concede a possibilidade de colisão. Nessa direção, segundo Gibson (1986), para nossos ancestrais, um coelho que se aproximasse concederia a possibilidade de se alimentarem, enquanto que um tigre que se aproximasse concederia a possibilidade de serem devorados. Conforme explica o psicólogo, as concessões não são, assim, boas ou ruins, positivas ou negativas, elas existem e se prestam às ações que permeiam o animal e seu ambiente.

Gibson (1986) assume que nós não percebemos o tempo, mas, sim, processos, mudanças e sequências. Segundo o estudioso, os eventos (episódios alinhados ao tempo) encontram-se dentro de eventos e formas encontram-se dentro de formas. As unidades naturais do ambiente terrestre e dos eventos não devem ser confundidas com as unidades métricas de espaço e tempo, que são, por excelência, convencionalizadas a partir de uma construção humana. Ao se aproximar de sua defesa à ecologia da percepção, Gibson (1986) elenca que o ambiente normalmente manifesta coisas que persistem e coisas que não persistem, além de pontuar que algumas características são invariantes e outras são variantes. Um espaço totalmente invariante, imutável nas suas partes e imóvel, seria completamente rígido e, obviamente, deixaria de ser um ambiente (de fato, não haveria nem animais e nem plantas). Por outro lado, “um espaço inteiramente variante e mutável quanto a todas as suas partes não seria um ambiente” (GIBSON, 1986, p. 14), uma vez que elementos de um ambiente precisam se perpetuar ao longo das cadeias temporal e espacial.

A essência de um ambiente ecológico, segundo os preceitos gibsonianos, é a de que ele circunda um indivíduo e, em um ambiente, há diversos pontos de observação. Estes, por sua vez, são constituídos pelos caminhos disponíveis para locomoção em um meio, e todos os habitantes de um ambiente têm igualmente a chance de explorá-lo. Nesse sentido, o ambiente circunda todos os observadores do mesmo modo que circunda um único observador. É justamente a ideia de um sujeito exploratório, capaz de acessar os objetos reais do ambiente, que dá sustentação aos modelos *PAM* (BEST, 1995) e *PAM-L2* (BEST; TYLER, 2007). Segundo Best (1995), a estimulação do ambiente fornece recursos informacionais diretos, ricos e confiáveis quanto ao mundo, sendo completamente dispensáveis os mecanismos cognitivos de representação mental e inferência. Apoiando-se sobre o aporte gibsoniano, Best (1995)

entende que a aprendizagem perceptual envolve uma sintonização do percebedor com o sistema-alvo de sons para a detecção de invariantes de alta ordem (as concessões), implicando diferenças funcionais no componente fonológico em desenvolvimento.

Em linhas gerais, segundo a visão de Best (1995) acerca da abordagem gibsoniana, “o percebedor apreende o objeto perceptual diretamente e não meramente um representante ou ‘substituto’ a partir do qual o objeto deva ser inferido” (BEST, 1995, p. 173). Sob o prisma de Best (1995) e de Best e Tyler (2007), perceber é extrair informação direta do mundo acerca de seus componentes – objetos, pessoas, relações, etc. – no tempo e no espaço, sem a mediação de um conhecimento inato ou associações mentais adquiridas. Nesse quesito, a pesquisadora faz a seguinte colocação:

Mecanismos cognitivos especiais para lidar com representações mentais e inferências indiretas não são necessários, porque os estímulos fornecem uma fonte de informação direta, rica e confiável sobre o mundo, e porque os percebedores se ocupam de seus sistemas perceptuais integrados na exploração ativa do mundo. Nesta visão, a aprendizagem perceptual envolve um maior ajuste para se detectarem invariantes de alta ordem disponíveis na cadeia de informação, em vez de mudanças em representações mentais e processos inferenciais. (BEST, 1995, p. 175)

No entanto, não são somente Best (1995) e Best e Tyler (2007) que se apropriam do construto gibsoniano para designar o que eles entendem por percepção⁴. Antes mesmo de os modelos *PAM* e *PAM-L2* de fato existirem, Fowler (1986) já fazia uma adaptação da teoria de Gibson (1986) para a sua teoria de percepção da fala nativa. Ou seja, verificamos a transposição de um objeto *somente visual* para um objeto que é visual, mas, em larga medida, também é *acústico*, de modo que estes dois meios remetam ao aparato articulatório. Ressaltamos, nesse âmbito, que a Fonologia Articulatória (ou Fonologia Gestual) de Browman e Goldstein (1989, 1992) e de Fowler (1986, 1996) prevê o aporte articulatório para a concepção da unidade gestual – além disso, o posicionamento dos referidos autores vai ao encontro de uma abordagem multimodal para a fala, em que variáveis de diversas naturezas (acústica, visual, etc.) atuam para a apropriação do sistema fônico. De acordo com Fowler (1996), sua teoria de percepção é uma teoria em que os gestos fonológicos são as ações públicas do trato vocal que causam a

⁴ Best (1995) afirma explicitamente que, em função de ela seguir Gibson (1986), inferências e representações mentais são desnecessárias ao evento perceptual. No entanto, a autora prevê a existência de segmentos fonológicos e fonéticos, que estão intimamente ligados à noção representacional. Além disso, sua referência massiva ao segmento, dentro de uma proposta gestual, sugere certa inconsistência com seu posicionamento.

estrutura (elemento que condiciona a percepção) nos sinais acústicos da fala. Logo, a perspectiva adotada por Fowler (1986) – e, por conseguinte, a de Best (1995), principalmente – já é uma releitura da abordagem ecológica de Gibson (1986), adaptada ao âmbito dos sons da fala.

Aos moldes de Gibson (1986) e de Fowler (1986), Best (1995) e Best e Tyler (2007) notam que a relação estabelecida entre o sujeito percebedor e o ambiente se torna a condição necessária e suficiente para que o conhecimento, que ocorre via percepção, seja instaurado a partir da detecção das concessões que o ambiente fornece. Tal constatação leva os autores, portanto, a fazer a ressalva de que o empreendimento teórico que realizam opera primordialmente em situações de percepção de L2 em detrimento de LE. Para os autores, o cenário de LE encontra-se nitidamente abaixo do ideal no que se refere à ecologia da aprendizagem de uma língua, pois oferece um contexto empobrecido para a aprendizagem da língua alvo. Além disso, ocorre em um ambiente em que a L1 é muito influente sobre a língua alvo, a qual, geralmente, não se estende para fora da sala de aula. Nesse sentido, o quadro em que se insere a LE frequentemente emprega instrução formal e informações gramaticais, enquanto que as situações de conversação são muito minimizadas. Os autores também mencionam que os momentos de aprendizagem de uma LE constituem um espaço em que a língua alvo é frequentemente veiculada por professores com sotaque materno, apresentando um modelo variável ou incorreto dos detalhes fonéticos a serem adquiridos. Um ambiente de aprendizagem de L2, para Best e Tyler (2007), não contemplaria tais características e seria suficientemente robusto para garantir a aprendizagem perceptual da língua alvo.

Convém deixar claro que, conforme os próprios Best e Tyler (2007) salientam, a percepção fônica em um contexto em que a língua alvo é estrangeira – e não uma segunda língua⁵ – ainda necessita de muita investigação, e o modelo por eles desenvolvido aplica-se fundamentalmente ao cenário de L2. Esta restrição não somente é explicitada pelos autores, mas também se caracteriza como condição imprescindível para a adoção do modelo. Oportunamente, a próxima seção irá tratar dos estudos realizados no Brasil que se utilizaram do aporte teórico do *PAM-L2*, ainda que para explicar a percepção fônica em contexto estrangeiro.

⁵ Diferenças entre L2 e LE serão explicitadas na seção 3.

3. A utilização do *PAM-L2* em contextos de percepção fônica de LE

Os estudos sobre percepção dos sons de L2 no Brasil envolvem, principalmente, o inglês como língua alvo e se utilizam, basicamente, dos modelos *SLM*, *PAM*, e *PAM-L2*. Na alçada do *SLM*, apontamos os estudos de Koerich (2002), Kluge (2004, 2009), Rauber (2006), Reis (2006) e Moore (2008). No que tange ao *PAM*, destacamos a pesquisa de Reis (2010)⁶ e, quanto aos trabalhos que utilizaram o *PAM-L2*, elencamos os de Bettoni-Techio (2008), Gutierrez e Battisti (2012), Perozzo (2013), Feiden, Alves e Finger (2014) e Feiden *et al.* (2016, no prelo).

Os estudos acima citados foram encontrados a partir de dois diferentes mecanismos de busca. O primeiro deles, de caráter documental, caracterizou-se por um levantamento realizado por Silveira (2010), a partir do qual selecionamos trabalhos que envolviam a utilização dos três modelos mencionados para explicar ou situar os resultados obtidos. Como tal repertório contou com a descrição de estudos conduzidos até 2009, procedemos ao exame dos currículos Lattes daqueles pesquisadores que, além de terem sido contemplados no levantamento de Silveira (2010), haviam começado a orientar trabalhos, em nível de pós-graduação, que também tivessem o compromisso de tratar seus dados à luz dos modelos perceptuais referidos. No entanto, não fomos bem sucedidos nesta segunda parte, pois não constavam, na lista de trabalhos orientados, quaisquer estudos que tivessem relação com os modelos perceptuais em que tínhamos interesse. Dessa forma, passamos ao segundo mecanismo de busca, de cunho exploratório, em que rastreamos estudos de pesquisadores que tiveram contato com nosso grupo de pesquisa, através de disciplinas oferecidas no Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul ou por meio de trabalhos realizados em conjunto.

Focaremos, nesta seção, apenas nos trabalhos que se utilizaram do *PAM-L2* enquanto referencial teórico para a percepção de sons não nativos e, também, que se valeram de suas previsões de assimilação para interpretar os resultados obtidos. Nossa linha de argumentação tem por finalidade apresentar cinco investigações interfonológicas que fazem uso do *PAM-L2* para compará-las entre si e evidenciar suas características comuns. Veremos que, apesar de ser um construto teórico voltado à L2, o modelo está sendo utilizado em pesquisas cujo foco de investigação é a percepção fônica de uma LE, o que exige dele futuras adaptações em nível estrutural, tarefa essa cujo primeiro passo damos no presente artigo. Além disso, resenharemos

⁶ Reis (2010) menciona que os resultados encontrados em sua pesquisa foram analisados também sob a ótica do *PAM-L2* (BEST; TYLER, 2007).

os estudos mencionados porque eles serão, também, motivação para justificar a proposta de alterações ao *PAM-L2*, o que será feito ainda nesta seção.

O primeiro estudo a ser citado é o de Bettoni-Techio (2008), o qual buscou compreender como brasileiros viriam a perceber e produzir⁷ a fricativa alveolar /s/ presente em [s]-clusters em posição de ataque inicial relativos a palavras do inglês, quando tal segmento era precedido por palavras cuja consoante em coda era vozeada e não vozeada. Além disso, a pesquisadora se interessava em verificar os efeitos de treinamento na percepção e na produção de tais sequências sonoras. Sua amostra de sujeitos foi composta por 23 alunos de inglês, adultos e crianças com minimamente 200 horas de instrução prévia no idioma, e a eles foram administradas tarefas de identificação, discriminação, fala semiespontânea e leitura. Os resultados analisados apontaram maiores índices de acuidade na identificação, na discriminação e na produção das sequências sonoras, tanto para as formas testadas quanto para novas formas. Adicionalmente, altos índices de acuidade nas tarefas foram observados na sessão de retenção, que ocorreu oito meses após o treinamento.

O segundo trabalho de que tratamos é o de Gutierrez e Battisti (2012), cujo objetivo era analisar a percepção da consoante nasal velar do inglês em posição de coda silábica final por aprendizes brasileiros do idioma. Para tanto, participaram da pesquisa 9 acadêmicos de uma instituição de ensino superior no sul do Brasil, com idade entre 19 e 25 anos, que estudavam a língua inglesa há aproximadamente 6 anos. Os alunos foram testados quanto ao seu nível de proficiência em inglês e, segundo o *Oxford Placement Test* (ALLAN, 2004), foram classificados como aprendizes de nível avançado. Para a condução do estudo, as pesquisadoras aplicaram aos acadêmicos um teste de discriminação categórica envolvendo as palavras alvo. As autoras apontam que os resultados obtidos indicam um bom nível de acuidade na discriminação perceptual entre a consoante nasal velar e a consoante nasal alveolar. Neste sentido, as pesquisadoras ressaltam que, na maioria das respostas fornecidas pelos alunos, a consoante nasal alveolar foi percebida, de fato, como sendo alveolar, e a nasal velar percebida, de fato, como velar.

O terceiro estudo que mencionamos é o de Perozzo (2013), que investigou como brasileiros percebiam o ponto de articulação de oclusivas não vozeadas sem soltura audível (a partir da informação coarticulatória com a vogal precedente) em posição de coda final de sílabas

⁷ O modelo de Best e Tyler (2007), ao contrário do de Flege (1995), é um modelo perceptual, não de percepção-produção. Extensões à produção, assim como fez Bettoni-Techio (2008), ficam a critério da autora.

da língua inglesa. A pesquisa contou com 17 acadêmicos do curso de graduação em Letras de uma universidade no sul do Brasil, com nível de proficiência estipulado entre básico e intermediário, de acordo com o *Oxford Placement Test* (ALLAN, 2004). Para medir a acurácia quanto ao ponto de articulação das consoantes propostas, foram aplicados um teste de identificação e um teste de discriminação. Também com o interesse de verificar se a instrução explícita sobre o fenômeno da não soltura de oclusivas teria papel sobre a acuidade na percepção dos pontos de articulação das consoantes finais, os acadêmicos foram alocados em dois grupos, experimental e de controle. Os resultados sugeriram maiores índices de acuidade perceptual quanto aos pontos de articulação labial e velar. Além disso, o grupo experimental obteve maiores índices de acuidade quando comparado ao grupo controle em ambas as tarefas, e os acadêmicos de nível intermediário apresentaram maiores índices de acuidade do que os acadêmicos de nível básico.

O quarto estudo a que nos remetemos é o de Feiden, Alves e Finger (2014), cujo objetivo foi constatar os efeitos da anterioridade e da altura da língua sobre a percepção de argentinos quanto às vogais médias-altas e médias-baixas do português brasileiro. Os participantes da pesquisa foram 16 falantes de espanhol rioplatense, com idade entre 16 e 18 anos, estudantes de ensino médio que cursavam uma disciplina de língua portuguesa há, pelo menos, 2 anos em uma escola na Argentina⁸. Os alunos foram submetidos a um teste de identificação perceptual contendo pares mínimos do português, os quais eram dissilábicos e apresentavam padrão acentual trocaico. Os resultados mostraram que, tanto nas vogais anteriores quanto nas posteriores, os participantes obtiveram índice mais elevado de acerto nos itens que apresentavam vogais médias altas em comparação com os itens que apresentavam vogais médias baixas. Como conclusão, os pesquisadores observaram que os participantes tiveram dificuldades na identificação das vogais médias do português, porém advertem que os referidos aprendizes se encontram em um estágio de formação das novas categorias fonético-fonológicas da língua alvo, apesar da dificuldade em delimitar as fronteiras acústico-articulatórias que caracterizam tais categorias.

⁸ Embora tais aprendizes tenham sido testados no Brasil, ou seja, em um ambiente de LE, eles aprenderam português na Argentina. Entretanto, os participantes eram intercambistas aqui no país, e estudaram português durante a sua estada. Tal fato nos impede de classificá-los, com propriedade, como aprendizes de português como LE ou como L2. Entretanto, inclinamo-nos a referi-los como aprendizes de português como LE em função de terem iniciado o estudo do português em seu país de origem e não no contexto nacional, e por terem recebido um maior número de horas de instrução no cenário estrangeiro.

Por fim, o quinto trabalho que referimos é o de Feiden *et al.* (2016, no prelo), que se propôs a analisar os índices de discriminação categórica entre as vogais médias-altas e médias-baixas do português brasileiro por falantes nativos de espanhol argentino. Este estudo é semelhante ao de Feiden, Alves e Finger (2014) por incorporar os mesmos estímulos auditivos e integrar os mesmos participantes, porém difere drasticamente no que concerne à natureza da tarefa: Feiden, Alves e Finger (2014) propuseram um teste de identificação perceptual, enquanto que Feiden *et al.* (2016, no prelo) administraram um teste de discriminação categórica, já mencionado. Neste estudo, os pesquisadores observaram que os alunos testados se mostraram capazes de discriminar o segmento anterior médio-alto do anterior médio-baixo do português, mas o mesmo não ocorreu para o contraste de altura dos segmentos análogos posteriores, indicando clara conjunção de tais elementos em uma única categoria fônica.

Para além de empregarem o mesmo modelo perceptual à análise e à interpretação dos resultados e se enquadrarem em uma metodologia experimental, os trabalhos destacados compartilham de uma situação comum extremamente importante para o propósito deste artigo: *a amostra utilizada em todos eles se caracteriza por indivíduos que percebem elementos fônicos de uma LE*. Nesse mérito, notamos que a motivação para alocá-los sob o rótulo de aprendizes de língua *estrangeira* (em detrimento de aprendizes de *segunda* língua) resulta de fatores cognitivo-ambientais, referentes ao domínio de uso do idioma, e pedagógicos, os quais se encontram interconectados.

Como toda a atribuição de “rótulos” pode ser escorregadia (ainda mais no que tange a aspectos humanísticos), justificaremos, brevemente, com base em Gass e Selinker (2008), nossa opção por tratar os participantes das pesquisas enumeradas como percebedores de sons de *LE*. Segundo Gass e Selinker (2008), as situações de aprendizagem de *LE*, e, portanto, de percepção fônica de *LE*, se diferenciam daquelas referentes à *L2* na medida em que ocorrem no ambiente em que a *L1* é falada (como, por exemplo, brasileiros aprendendo inglês no Brasil). Pedagogicamente, para os autores, é bastante comum em tal cenário haver instrução formal acerca do idioma a se adquirir, enquanto que, em um ambiente de aprendizagem de *L2*, geralmente, não há apelo a meios instrucionais. Tal definição vai ao encontro das características elencadas por Best e Tyler acerca das diferenças entre a aprendizagem de uma *L2* e de uma *LE*, reiterando o contraste entre o tipo de exposição informacional no caso de aprendizes de *L2* e aprendizes de *LE*. Em nosso entendimento, se os aprendizes estão inseridos em um ambiente onde eles têm contato com a língua alvo em um amplo espectro de situações, a aprendizagem

perceptual da fala deverá, por hipótese, ser otimizada. Oportunidades mais limitadas de contato com o idioma alvo implicam menor participação em práticas situacionais que o envolvem, e disso decorre menor acesso a informações oriundas de comunicação autêntica.

Dadas essas condições, os estudos relatados estão, de fato, se utilizando do *PAM-L2* para dar sustentação às suas pesquisas, bem como para analisar, em termos de percepção, os fatos fônicos que encontram. Verificamos, pois, que tais estudos alocam-se no contexto de percepção fônica em LE, mas fazem uso de um modelo voltado para a percepção fônica de uma L2, e, conforme explanam os próprios autores, o *PAM-L2* não se presta para a percepção fônica no contexto de LE. Dessa forma, vislumbramos duas opções: abandonar o modelo, no caso da percepção fônica em LE, ou adaptá-lo. Optamos, todavia, pela última alternativa. Sugerimos, portanto, que seja elementar a necessidade de adaptar o *PAM-L2* ao contexto de percepção de LE (que configura o caso da grande maioria dos aprendizes de idiomas no cenário nacional), no sentido de levar em conta a realidade dos aprendizes de línguas estrangeiras que residem no Brasil ou em países em que não haja uma segunda língua que ofereça uma quantidade similar de *input* em comparação com o da L1. Tais alterações dizem respeito aos seguintes aspectos:

- A percepção e a aprendizagem perceptual de uma LE deve ser encarada como um processo *cognitivo*, o qual envolve o processamento da informação veiculada na fala, e não como a *apropriação de ações* sem relação com mecanismos inferenciais e representações mentais.
- A unidade fônica da percepção deve ser o *gesto acústico-articulatório* cunhado por Albano (2001), uma vez que tal primitivo se relaciona e interage com outras operações linguísticas que compõem a língua a ser compreendida, e não o *gesto articulatório* de Browman e Goldstein (1989, 1992) e de Fowler (1986, 1996), que não possui compromisso com outros níveis da gramática.
- A noção de *ambiente ecológico* deve ser desvinculada do contexto de percepção em sala de LE, em virtude de que tal espaço carece das condições necessárias para configurar um ambiente ecológico de fato.
- O realismo *indireto* perceptual (a ser tratado na seção 4.2), em detrimento de sua contraparte *direta*, deve ser a teoria filosófica que embasa a percepção dos sons de uma LE, não somente por conseguir captar as relações fônicas entre L1 e LE, mas também

por se tratar de uma unidade referente ao signo linguístico, que é construído histórica e socialmente, figurando como uma entidade convencionada.

É no tocante a tais reflexões que nos ocupamos da próxima seção, com vistas a ponderar sobre as alterações necessárias ao modelo para que ele dê conta, em alguma medida, da percepção fônica em LE.

4. *PAM-L2* e sua demanda na LE: possíveis alterações no modelo

Nosso intuito ao discorrer sobre o *PAM-L2*, nesta seção, é o de prover mais esclarecimentos sobre as possíveis alterações no modelo para que ele venha a abarcar os fatos perceptuais em contexto de percepção fônica estrangeira, de acordo com os aspectos motivadores que justificam tal exercício teórico, expostos na seção anterior. Destacamos que não seremos exaustivos quanto às sugestões, mas tentaremos contemplar seus aspectos mais fundamentais, em função de que nossas ponderações configuram um primeiro passo na tentativa de adaptação do modelo, com o objetivo de que ensejem discussões subsequentes acerca de nossa reflexão.

Consideramos que, se o *PAM-L2* se resumisse apenas às predições de assimilação perceptual descritas na seção 2, em primeira análise, não necessitaria mais do que uma reflexão teórica que resultasse em acréscimos a casos de assimilação não contemplados pelo modelo original, independentemente de diferenciarmos LE de L2. Entretanto, não é só de predições que o modelo é formado: há todo um embasamento teórico que o sustenta (no sentido de conceber o desenvolvimento de conhecimento linguístico), o qual pode ser discutido sobre o ponto de vista fonológico, em que o modelo assume como primitivo fônico o gesto articulatório, e sobre a perspectiva filosófica, que tem a ver com a abordagem ecológica da percepção e com o realismo direto de James Gibson. Problematicamos, assim, nas duas subseções a seguir, o aspecto fonológico da percepção em LE e a noção de ambiente ecológico e realismo direto que embasam o *PAM-L2*, e sugerimos aportes alternativos para tratar do fenômeno perceptual em contextos de LE.

4.1. O gesto articulatório e a gramática fônica do aprendiz/percebador de LE

A justificativa para realizarmos uma discussão sobre o aspecto fonológico da percepção em LE reside no fato de que, ao perceber os elementos sonoros da LE a que está exposto, o

percebedor/aprendiz os relaciona àqueles de sua L1 (FLEGE, 1995; BEST; TYLER, 2007). Em nossa visão, diferentemente do que concebem Best e Tyler (2007), este processo diz respeito a um empreendimento mental⁹ e motor. Desse modo, quando o conhecimento fonológico materno e o sistema fonológico alvo estão em contato, a *habilidade mental* (portanto, representacional) de categorizar os novos sons para distinções lexicais e a *habilidade motora*¹⁰ (portanto, física) do aprendiz de perceber/produzir rotinas articulatórias operam absolutamente em conjunto.

A título de esclarecimento, julgamos pertinente definir o que concebemos pelos termos *habilidade mental* e *representação mental*, pois se farão úteis ao longo de nossa exposição. *Habilidade mental* ou *atividade mental*, segundo Gazzaniga *et al.* (2012), tem a ver com a mente em ação, isto é, diz respeito ao processamento das informações. As experiências de percepção que um indivíduo detém enquanto interage com o mundo, assim como memórias, pensamentos e sentimentos, são exemplos da mente em ação. Portanto, a atividade ou a habilidade mental resulta de processos biológicos (as ações das células nervosas e suas reações químicas associadas) no cérebro. *Representação mental* refere-se a um símbolo cognitivo interno que corresponde à realidade externa e pode ser *analógica*, quando trazemos à mente a imagem de um objeto, por exemplo, ou *simbólica*, que exprime o conceito ou nosso conhecimento sobre o objeto e não sobre suas propriedades físicas¹¹ (GAZZANIGA *et al.*, 2012). Evocamos, portanto, os postulados de Gazzaniga *et al.* (2012) e, para efeitos linguísticos, acreditamos que as representações mentais dos indivíduos sejam adaptativas, complexas e não lineares (DE BOT; LOWIE; VERSPOOR, 2007; BECKNER; *et al.*, 2009), em oposição à noção estática, cartesiana e serial de representação mental, veiculada fortemente pela tradição gerativa.

No arcabouço do modelo fonológico conhecido como Fonologia Articulatória, conforme propõem Browman e Goldstein (1989, 1992), os gestos articulatorios vêm a ser as unidades básicas de contraste fonológico nas línguas, figurando como caracterizações abstratas de eventos articulatorios físicos de tempo intrínseco¹². Para os estudiosos, gestos articulatorios

⁹ Seguimos a definição de mente de Kosslyn e Koenig (1995, p. 4), de acordo com a qual “a mente é aquilo que o cérebro faz”. Dessa maneira, segundo postulam Gazzaniga *et al.* (2012, p. 5), “o cérebro físico habilita a mente”, e ambos são inseparáveis.

¹⁰ O termo *motor* está sendo usado aqui como sinônimo de *mecânico* e não está vinculado aos preceitos da Teoria Motora da Percepção da Fala (LIBERMAN; MATTINGLY, 1985).

¹¹ Para fins de argumentação, seguiremos a contraparte *simbólica* da representação mental.

¹² Por tempo intrínseco, entenda-se a incorporação da dimensão *tempo* na especificação fonológica de um segmento. Para uma explicação mais elaborada acerca de uma unidade sonora com tempo intrínseco, recomendamos a leitura de Fowler (1980).

e organizações gestuais podem captar informações tanto categóricas quanto gradientes na cadeia da fala.

Goldstein e Fowler (2003), trabalhando no desenvolvimento da Fonologia Articulatória, consideram que as formas linguísticas são os meios que as línguas oferecem para que a comunicação entre as pessoas se faça possível e, por isso, advogam que tais formas são tipos de ações públicas e não exclusivamente categorias mentais, isoladas da realidade física, como apresenta a maioria das teorias fonológicas, de percepção e de produção. Os pesquisadores entendem que as unidades fonológicas são abstratas com respeito às variáveis articulatórias e acústicas que são tipicamente medidas, mas não tão abstratas a ponto de ignorarem a área vocal e alocarem-se exclusivamente na mente. Interessantemente, sob a perspectiva gestual, “as unidades fonológicas mais básicas devem ser discretas e recombináveis” (GOLDSTEIN; FOWLER, 2003, p. 189), além de deverem se utilizar de uma moeda comum entre percepção e produção.

Algo que merece particular atenção é o fato de a Fonologia Articulatória não ter sido criada com a pretensão de explicar o processo de aquisição ou desenvolvimento do sistema sonoro de uma L2 ou LE, repousando exclusivamente sobre a L1. Conforme podemos observar nos textos iniciais do modelo (BROWMAN; GOLDSTEIN, 1989, 1992), a Fonologia Articulatória se reserva a elucidar relações gestuais veiculadas na fala e que têm um compromisso com distinções lexicais, mas nada afirma explicitamente sobre a *adição* de um novo sistema àquele já existente:

O objetivo de uma gramática fonológica é dar conta do conhecimento (implícito) *nativo* dos falantes com relação à estrutura fonológica e às regularidades em uma determinada língua, incluindo o inventário de unidades lexicalmente contrastivas, restrições quanto a formas fonológicas e alternâncias sistemáticas no que se refere a formas lexicais resultantes de combinação morfológica e adjunção em contextos prosódicos específicos. [(GOLDSTEIN; FOWLER, 2003, p. 170) grifo nosso]

Contudo, o modelo pode ser utilizado para ilustrar a percepção e a produção da fala em L2 ou LE (segundo faz, por exemplo, o *PAM-L2*), uma vez que as informações sonoras linguísticas se apoiam sobre unidades gestuais. Neste sentido, podemos fazer um paralelo entre o desenvolvimento fonológico estrangeiro e aquele que ocorre em estágios iniciais da L1, levando em conta os postulados de Goldstein e Fowler (2003). De acordo com os pesquisadores, para que a criança associe seu trato vocal e os movimentos de sua face àqueles do modelo (um

adulto), a representação da fala deve refletir o que é comum entre a ação do adulto percebida visualmente e a ação gerada pela criança, a qual é percebida proprioceptivamente. De igual maneira, acreditamos que, em um ambiente de L2 ou LE, o aprendiz (que é percebedor e produtor de sons da fala) deve ter como representação o reflexo daquilo que é percebido visual e acusticamente com base no modelo e daquilo que é percebido via propriocepção pelo próprio aprendiz.

Concordamos com Browman e Goldstein (1989, 1992) acerca das faces abstrata e concreta do gesto articulatório e concordamos, também, com Goldstein e Fowler (2003) sobre a necessidade de uma teoria fonológica levar em conta aspectos mais concretos da cadeia sonora, como a gradiência apresentada nos sons da fala. Compartilhamos, também, da postura de Albano (2001) no que compete ao fato de as variáveis do trato vocal, responsáveis por definir determinado gesto articulatório, estarem duplamente relacionadas aos articuladores: para a autora, na esfera abstrata, as variáveis encontram-se funcionalmente unidas sob o articulador envolvido mais diretamente com as restrições que executam, e, na esfera concreta, tais variáveis oportunizam o movimento do articulador (ou seja, a realização de uma tarefa mecânica) em questão.

O que nos parece muito oportuno na descrição gestual delineada por Albano (2001) é o seu compromisso de relacionar operações gestuais a operações linguísticas outras que ocorrem na gramática como um todo. Tal conexão fica clara quando a pesquisadora explana que o realinhamento e o redimensionamento dos gestos articulatórios não somente dão conta da lexicalização dialetal ou estilística daqueles processos que, em princípio, seriam esporádicos na fala corrente, mas também atuam como o meio através do qual as regularidades fônicas adentram “níveis mais profundos da gramática” (ALBANO, 2001, p. 102). Evidências de cunho morfofonológico (entre outras) trazidas por Albano (2001) compreendem parte dos exemplos que dão subsídio a tal constatação. Vemos, portanto, ao longo da exposição feita por Albano (2001), a tentativa de prever relações entre unidades de natureza fônica e outras unidades gramaticais, como o morfema, no caso da formação do plural em alguns itens lexicais do português, e como o sintagma, nos casos de sândi externo.

Embora Goldstein e Fowler (2003) admitam que haja alguma abstração no que se refere ao gesto articulatório, sua argumentação é conduzida, em boa parte, a salientar com enorme robustez o caráter físico de tal unidade fônica. É nesse sentido que julgamos mais conveniente a visão de Albano (2001), em que as faces abstrata e física do gesto articulatório parecem estar

mais equilibradas, além de tal unidade estar conectada e interagir com outros aspectos da gramática, também de cunho abstrato. Ora, em nosso ponto de vista, ao contrário do que referem os construtos de Gibson (1986), intriga imaginar que, em um ambiente de percepção de sons de uma LE, o percebido tenha que se valer com tamanho afinco de ações puramente¹³. Destacamos que, de maneira nenhuma, estamos deixando de lado a contraparte física do gesto articulatório (afinal, ela é de fundamental importância na percepção e na produção da fala), mas nos cabe acrescentar que, frente ao sistema linguístico da L1, que já está formado e internalizado, é indispensável que o percebido atue com base em inferências e representações mentais¹⁴.

Em consonância com Flege (1995) e Best (1995), parece-nos irrefutável a afirmação de que os indivíduos, enquanto percebidos de sons não nativos, recorram ao seu conhecimento prévio da L1 no momento em que decidem se determinado som da fala corresponde ou não a uma categoria nativa. No entanto, chamamos a atenção para o fato de que, no momento em que o percebido reconhece distinções lexicais veiculadas através da discriminação categórica – atribuindo, assim, significado linguístico à massa sonora a que tem acesso –, ele está resgatando o inventário lexical da língua alvo, cumprindo um exercício *cognitivo* por excelência. Perceber diferenças sonoras em uma LE, em nosso entendimento, não consiste apenas em se apropriar de uma tarefa articulatória, mas compreende a relação cognitiva entre unidades sonoras prévias (L1) e novas (LE), cujo resultado pode ser visto em rotinas motoras¹⁵.

Finalmente, reiteramos que o gesto articulatório deva ser o primitivo sonoro que permeia a percepção de novos elementos fônicos. Além disso, com base em Albano (2001), acreditamos que, de fato, existe uma gramática fônica, abstrata e representacional, a qual opera com o gesto articulatório. Este, por sua vez, corresponde a uma unidade de contrastes lexicais que se apresenta tanto de maneira abstrata (mental) como de maneira física (motora). A gramática fônica do aprendiz/percebido de LE, em nossa visão, se configura, pelo menos no que compete à discriminação categórica, como um espaço de constantes relações entre unidades sonoras

¹³ Prenunciamos um dos motivos por que discordamos da adoção da abordagem realista direta à percepção sonora em LE. Outros argumentos serão fornecidos na seção 4.2.

¹⁴ Este argumento também será utilizado na seção 4.2 para trazer à baila o construto do realismo indireto, uma perspectiva antagônica ao realismo direto.

¹⁵ Explicitamos, portanto, o fato de que os ouvintes percebem os sons da LE com base em seu inventário fônico da L1, e aproveitamos para destacar, neste âmbito, que o percebido/aprendiz deve aprender a temporalizar os gestos articulatórios que já possui, habilidade referente à orquestração destes gestos (SILVA, 2014; ZIMMER; ALVES, 2012).

prévias e novas, as quais podem ser mais ou menos estáveis a depender do domínio de uso da LE.

4.2 A percepção fônica não nativa e os limites de sua ecologia

Acreditamos que a fala, em seu sentido tradicional, é parte constitutiva da língua. Desse modo, operações fonéticas, fonológicas, morfológicas e sintáticas, bem como relações semânticas e pragmáticas, estão alocadas sob o domínio da língua e estão em constante interação. Sendo a língua, portanto, um bem cultural que funciona como instrumento de comunicação convencionado e situado social e historicamente, e que envolve alto grau de abstração, ela não está no escopo de uma teoria como a de Gibson (1986), conforme o próprio autor salienta, a qual se ancora no realismo direto. Fowler (1986, 1996) e Goldstein e Fowler (2003) parecem, de certa forma, ignorar a advertência feita por Gibson (1986) acerca de sua teoria não ser aplicável ao conhecimento linguístico, veiculado pela fala ou pela escrita (p. 10)¹⁶.

Imaginemos um contexto de percepção de LE, como é o caso da sala de aula, por exemplo¹⁷. Os percebedores/aprendizes têm contato com apenas uma porção selecionada do idioma alvo, não sendo capazes de experienciá-lo em sua totalidade. A parte sonora do idioma geralmente é acessada pelos percebedores/aprendizes através da mediação da fala do professor ou de gravações em áudio e/ou vídeo, sendo, portanto, indireta. Lembramos que, para Gibson (1986), a percepção direta implica a ausência de um aparato entre o percebedor e o objeto, portanto, recursos tecnológicos audiovisuais não se encaixam em uma visão direta da percepção. Ademais, a fala do outro, seja a do professor ou a dos colegas, tampouco pode ser considerada como sendo perceptualmente direta. Isso se dá em função de que o realismo direto despreza a percepção em *segunda mão* (utilizando o termo gibsoniano), ou seja, o resultado da percepção de outro indivíduo como fonte primária para a percepção.

¹⁶ Para Gibson (1986), a luz estruturada fornece informação sobre o ambiente visual. Para Fowler (1986, 1996) e Goldstein e Fowler (2003), analogamente, o gesto articulatório é especificado pelo sinal acústico da fala (p. 183), ou seja, o ar acusticamente estruturado fornece informação sobre o ambiente auditivo. Todavia, mantemos certa reserva quanto a tal postulado de Goldstein e Fowler (2003), pois questionamos a comensurabilidade entre a informação veiculada visualmente e aquela veiculada acusticamente, uma vez que os aparatos visual e auditivo operam de maneira distinta no que compete à detecção dos estímulos que circundam o percebedor.

¹⁷ Referimo-nos ao contexto instrucional para exemplificar o ambiente perceptual de LE apenas para fins ilustrativos. Convém destacar que, para Best e Tyler (2007), independentemente de ser instrucional, qualquer ambiente perceptual de LE é deficitário quando comparado ao de L2.

As palavras e os diálogos que constam nos manuais didáticos são gravados por locutores treinados, os quais modalizam sua voz para atingir diversos efeitos entoacionais. Aquela informação que soa como surpresa ou animação é propositalmente criada e não reflete, necessariamente, a emoção que o locutor está sentindo no momento da elocução. Ademais, as gravações ocorrem em estúdios profissionais com tratamento acústico, visando a impedir que o sinal acústico seja mascarado por qualquer barulho simultâneo à fala, o que, de fato, não acontece em situações reais de comunicação. Enfatizamos, assim, que o uso das gravações não segue os preceitos ecológicos gibsonianos e tampouco resulta em uma percepção direta, o que, a uma primeira vista, viria a corroborar a previsão de Best e Tyler (2007) de que, para a percepção em contexto de LE, seu modelo não se aplica.

Obviamente, não podemos ser ingênuos a ponto de reduzir as limitações do ambiente perceptual de LE a uma pedagogia de LE problemática: mesmo se dispuséssemos da mais adequada das pedagogias, ela não seria suficiente para garantir as concessões que um ambiente de L2 *teoricamente* garante. Reiteramos que nossa argumentação não objetiva criticar a pedagogia de línguas estrangeiras desenvolvida no país, pois falantes de LE altamente competentes, oriundos de instrução formal explícita¹⁸, abundam nas diversas interações comunicativas.

Inicialmente, sem considerar o que expusemos até aqui, poderíamos pensar que o contexto de sala de aula, no que tange à percepção fônica especialmente, seria um ambiente ecológico. Tal constatação seria baseada no fato de que a sala de aula configura o espaço que cerca o percebedor/aprendiz e lhe fornece os subsídios necessários para que ele perceba elementos de seu entorno e esteja apto a desempenhar tarefas. Essa condição estaria de acordo com a visão de Gibson (1986), uma vez que, para o psicólogo, a característica fundamental de um ambiente é a de que ele circunda o indivíduo. No entanto, uma leitura mais cuidadosa da obra de Gibson (1986) nos fornece subsídios para constatarmos que, de fato, o cenário de percepção e aprendizagem de LE não seja ecológico, e apresentaremos duas razões para não seguirmos essa taxonomia.

O primeiro motivo para não rotularmos a percepção de LE em sala de aula como ecológica, em termos gibsonianos, ancora-se na proposição de que um ambiente ecológico, nos moldes de Gibson (1986), é muito amplo e está ligado à exploração e à sobrevivência. Gibson

¹⁸ Presumimos que esta mesma condição pode gerar questionamentos pertinentes, inclusive, para contextos de aprendizagem de L2, conforme mencionamos na conclusão deste trabalho.

(1986) declara que, para que o percebedor obtenha sucesso na detecção das concessões, é primordial que o ambiente seja significativo. Um ambiente significativo consiste de uma área territorial, abrigo, água, objetos, ferramentas e outros animais, além de estar em mútua relação com o percebedor e de fornecer a ele tudo de que precisa para se desenvolver. Bem sabemos que a sala de aula de LE não oferece todos os mecanismos e ferramentas para que o aprendiz/percebedor desenvolva seu conhecimento linguístico e perceptual, e isso se torna um problema para uma abordagem ecológica da percepção. O segundo motivo para não tratarmos o contexto da percepção em LE como relacionado a um ambiente ecológico reside na constatação de que a sala de aula, isto é, o entorno do percebedor, faz parte de um cenário idealizado: tal ambiente foi pensado para fins didáticos e construído para prover algumas condições de percepção¹⁹. Em outras palavras, o espaço de sala de aula de LE é uma criação e uma adaptação de um contexto muito maior do que é capaz de expressar se comparado ao contexto de imersão *idealmente* propiciado pelo cenário de L2, e não estava disponível naturalmente ao aprendiz/percebedor desde sua gênese.

Sugerimos, portanto, que o termo *ambiente ecológico* seja desvinculado daquilo que concebemos ser o espaço referente à sala de aula de LE, local em que ocorre em larga medida a percepção fônica estrangeira, uma vez que tal cenário não cumpre os pré-requisitos básicos do que entendemos por ambiente ecológico, conforme explicitado no parágrafo anterior.

No tocante ao acesso à realidade que envolve o percebedor de LE, propomos que a forma mais adequada de realismo seja a *indireta* ao invés de direta. Porém, antes de defendermos o realismo indireto, julgamos apropriado definir o que é o realismo em si. Segundo informa o *Blackwell Dictionary of Western Philosophy* (BUNNIN; YU, 2004), os diferentes tipos de realismo dizem respeito à existência objetiva de vários objetos e propriedades, tais como o mundo externo, objetos matemáticos, universais, entidades teóricas, relações causais e propriedades estéticas e morais. A ideia central do realismo é a de que os objetos existem no universo, independentemente de o percebedor saber ou acreditar em sua existência. Em outras palavras, conforme ilustra Dancy (1985), o realismo na teoria da percepção – realismo perceptual – representa o ponto de vista de que os “objetos que apreendemos são capazes de

¹⁹ Trazemos esta discussão sobre a percepção fônica no contexto de sala de aula apenas como referência a um cenário mais prototípico de um ambiente de LE. Todavia, ressaltamos que o contexto de LE é muito mais amplo do que o espaço da sala de aula, não se reduzindo, portanto, a somente este contexto.

existir e geralmente existem de fato, e retêm pelo menos algumas das propriedades que apreendemos terem, mesmo quando são despercebidos” (p. 182).

Essencialmente no que compete à natureza da percepção humana, em um enfoque realista, há uma longa discussão que impõe uma questão-problema aos filósofos da percepção: percebemos os objetos do mundo de maneira direta ou indireta? Para Gibson (1986), Fowler (1986), Best (1995) e Best e Tyler (2007), não há dúvida: a percepção humana é realista direta. Para os filósofos da percepção, o enredo não é tão simples assim. Dancy (1985) afirma que o contraste entre o direto e o indireto é bastante escorregadio e, por vezes, difícil de estabelecer com firmeza. Mound (2003) explana que a controvérsia reside no fato de percebermos os objetos do mundo diretamente ou por meio de elementos intermediários, como ideias, imagens, impressões, sensações ou dados dos sentidos.

De modo bastante interessante, Dancy (1985) comenta que a disputa entre o realista direto e o indireto se estabelece a partir da possibilidade de estarmos ou não diretamente conscientes acerca da existência e da natureza dos objetos. Sob esta perspectiva, o autor expõe que:

Ambos, como realistas, concordam que os objetos físicos que vemos e tocamos são capazes de existir e retêm algumas das suas propriedades, (mesmo) quando não apreendidos. Mas o realista indireto afirma que nunca estamos diretamente conscientes de objetos físicos; estamos apenas indiretamente conscientes deles, em virtude de uma consciência direta de um objeto intermediário (variadamente descrito como uma ideia, dado dos sentidos, percepto ou aparência). O realista direto nega esta afirmação. [(DANCY, 1985, p. 183) observações parentéticas nossas]

O realismo indireto prevê que o mundo externo realmente existe, mas nossa percepção em relação a ele é mediada pela percepção de objetos intermediários subjetivos, como, por exemplo, sensações (BROWN, 2009). Conforme Dancy (1985), a vertente indireta do realismo presume que nunca estamos diretamente conscientes dos objetos físicos, em função de estarmos diretamente conscientes de um objeto intermediário, que podem ser ideias, imagens, impressões, sensações ou dados dos sentidos, segundo, também, aponta Mound (2003). A possibilidade de haver algo entre o percebedor e o objeto a ser percebido vai completamente ao encontro da hipótese de que *a percepção de elementos sonoros de uma LE é filtrada pelo conhecimento fônico da L1*, conjectura em favor da qual nos posicionamos.

Seguimos a perspectiva realista indireta (JACKSON, 1977, 2010), portanto, em função de três premissas elementares. A primeira delas, conforme problematizamos nesta seção, é mais transparente quanto à conotação do termo *indireto* e tem a ver com a constatação de que, no contexto de LE, a percepção e a aquisição fônicas ocorrem de maneira mediada, seja pela fala do professor, seja pelo ferramental tecnológico disponível ao aprendiz. A segunda premissa em favor da abordagem indireta, advinda das ponderações feitas na seção anterior, reside na assunção de que o conhecimento fônico da L1 baliza a percepção de elementos sonoros estrangeiros, condicionando as suas categorizações (segundo evidencia a literatura em aquisição fonológica não nativa). Por fim, a terceira premissa que motiva a noção indireta do realismo se apoia sobre a interação da unidade gestual, de natureza acústico-articulatória, com outros âmbitos da gramática, refletindo-se, inclusive, no estabelecimento da relação significante-significado.

Cabe destacar que, em nossa concepção, a aprendizagem perceptual é um processo cognitivo por excelência, que envolve o processamento da informação da fala, assim como outras funções mentais – outro ponto que nos diferencia da perspectiva de Gibson (1986), Best (1995) e Best e Tyler (2007). Desse modo, seguimos Chamot *et al.* (1993), de acordo com o qual a aprendizagem, seja ela relacionada à habilidade linguística ou a outras habilidades humanas, implica um curso cognitivo *per se*, em que os indivíduos ativamente selecionam e organizam as informações que os circundam, além de relacioná-las ao seu conhecimento prévio. Desse modo, os aprendizes internalizam aquilo que consideram ser importante e utilizam as informações armazenadas. A ação cognitiva de *selecionar as informações linguísticas dispostas no ambiente, analisar suas formas e funções, refletir sobre suas produções, antecipar os tipos de demandas possíveis e ativar o conhecimento e as habilidades prévias para a execução de novas tarefas linguísticas* configura uma capacidade cognitiva complexa (MCLAUGHLIN, 1987).

Paralelamente, Brown, Roediger e McDaniel (2014) explicam que a aprendizagem se refere ao ato de adquirir conhecimento e habilidades, tornando-os disponíveis na memória para serem utilizados em oportunidades e problemas futuros, mecanismo que, por si só, já é capaz de abalar a perspectiva realista direta acerca da percepção da fala em LE. Ao que ilustra Mildner (2008), todo o tipo de aprendizagem se apoia sobre a *plasticidade*²⁰ do nosso cérebro,

²⁰ A plasticidade desenvolvimental compreende mudanças em conexões que ocorrem após o nascimento como consequência de interações com o meio. Tais mudanças tornam possível o processo de aprender e adquirir novas

contemplada na estrutura, no desenvolvimento e no funcionamento do sistema nervoso. Vinculamo-nos, portanto, a uma perspectiva cognitivista (e representacional) acerca da aprendizagem, diferente da posição de Gibson (1986).

A título de conclusão, não podemos esquecer que nós, seres humanos, temos a habilidade de racionalizar o conhecimento que chega até nós e de filtrar cognitivamente aquilo que nos é representativo, cabendo a este processo, também, recursos inferenciais e a instituição de uma representação mental acerca dos objetos percebidos. Adicionalmente, recorreremos à ideia da existência de um conhecimento mental quanto a unidades fônicas justamente pelo respaldo na literatura em psicologia e neuropsicologia (LUND, 2001; LECLERCQ; ZIMMERMANN, 2005; STYLES, 2005; KRAMER; WIEGMANN; KIRLIK, 2007; MILDNER, 2008; GAZZANIGA; HEATHERTON; HALPERN; HEINE, 2012), a qual faz referência a mecanismos de atenção e traz evidências a favor do armazenamento de informações sonoras na memória. Nesse âmbito, pensamos que uma abordagem da percepção fônica vinculada ao realismo indireto parece ser muito mais apropriada, além de dar conta, de modo mais elegante, das estipulações que fazemos sobre o mundo que nos cerca e, conseqüentemente, sobre o conhecimento dos padrões sonoros a que temos acesso.

5. Conclusão

Os pontos que levantamos neste artigo dizem respeito a possíveis modificações no modelo *PAM-L2* no momento em que este é utilizado para dar conta da percepção fônica de uma LE, especialmente nos contextos em que o idioma alvo não é amplamente utilizado fora do espaço da sala de aula. Tal situação corresponde largamente ao cenário da percepção do sistema sonoro do inglês, por exemplo, por falantes nativos do português brasileiro, que não estão inseridos em um ambiente onde a língua alvo é a principal ferramenta nas interações verbais.

Muito temos a considerar quando tentamos adaptar um modelo de percepção da fala em L2 para o cenário da percepção fônica em LE. No que concerne ao *PAM-L2*, especialmente, sugerimos as modificações a seguir:

experiências, mas também resultam desses processos e se fazem presentes desde os primórdios da existência humana. Outro tipo de plasticidade se refere à capacidade de recuperar funções cognitivas após danos cerebrais (MILDNER, 2008, p. 139).

- Ao contrário de Gibson (1986), Best (1995) e Best e Tyler (2007), que não preveem o uso de inferências e representações mentais para o evento perceptual, defendemos uma abordagem cognitiva da percepção e da aprendizagem perceptual, em que o aprendiz/percebedor opera tanto mentalmente, nas relações abstratas entre os sons da L1 e da LE, quanto fisicamente, nas rotinas articulatórias.
- Em detrimento da visão de Fowler (1986, 1986), clamamos pela adesão ao gesto articulatório na visão de Albano (2001), pois este se relaciona a outras unidades linguísticas, como morfológicas e sintáticas, sendo bastante conveniente para o tratamento da fala enquanto parte da gramática e, principalmente, pelo fato de tal unidade fônica percorrer o léxico dos usuários da língua.
- Sugerimos que o conceito de ambiente ecológico seja desvinculado das situações de percepção fônica em LE, especialmente nos contextos de sala de aula, uma vez que, neste cenário, a comunicação muitas vezes não é autêntica e o percebedor carece de uma exposição que seja, de fato, massiva à língua em uso efetivamente.
- Diferentemente de Gibson (1986), Best (1995) e Best e Tyler (2007), que concebem o realismo direto como forma de acesso aos objetos do mundo, prezamos pelo realismo indireto (JACKSON, 1977, 2010) como sendo o meio de acesso às unidades fônicas disponíveis ao percebedor. Optamos pelo realismo indireto perceptual porque ele dá conta do fato de que a L1 permeia as relações gestuais já internalizadas e aquelas que estão por adentrar a sua gramática fônica, além de reconhecer que existe sempre um aparato (seja mental ou físico) entre o percebedor e o objeto a ser percebido.

Reconhecemos, entretanto, que ainda se faz necessário incorporar ao modelo uma teoria que dê conta de mecanismos cognitivos como atenção e processamento do *input*, já que se faz indispensável que, para a percepção fônica em LE, o percebedor esteja atento aos contrastes finos da massa sonora. A atenção ao *input* fornece ao percebedor não somente a possibilidade de distinguir unidades fônicas, mas também de conectar tais informações ao seu léxico e, conseqüentemente, à sua gramática. Neste sentido, atentar ao *input* permite a seleção dos aspectos acústico-articulatórios primordiais para os contrastes funcionais na LE, uma vez que línguas distintas podem eleger diferentes aspectos acústico-articulatórios prioritários, essenciais no sistema alvo, para veicular oposições de significado (ALVES; ZIMMER, 2015; SCHWARTZHAUPT; ALVES; FONTES, 2015).

Ainda no que diz respeito a uma maior elaboração em termos de processamento do *input*, é preciso, também, considerar a multimodalidade da informação linguística. Dentro dessa variada gama de estímulos de caráter multimodal, outro compromisso do modelo é explicar o impacto das relações grafo-fônico-fonológicas (ZIMMER; SILVEIRA; ALVES, 2009) sobre a percepção da fala, uma vez que o conhecimento da língua escrita pode causar assimilações perceptuais equivocadas, resultando em sobreposições de categorias fônicas no nível interfonológico.

Destacamos, igualmente, a necessidade de se elaborarem tarefas de percepção que se proponham a captar com mais precisão a maneira como os percebedores categorizam os sons da LE. Dessa forma, as predições estabelecidas pelo *PAM-L2* poderiam ser testadas empiricamente e, como consequência, poderíamos ter fortes indicadores das possibilidades de assimilação perceptual. Tal constatação revelaria importantes fatos sobre percepção de distinções sonoras e também sobre a organização abstrata das unidades da fala, implicando conhecimento acerca das relações gestuais no âmbito tipológico.

Enfatizamos que as fronteiras entre LE-L2 estão longe de serem bem delimitadas. Sempre haverá casos em que o aprendiz de um novo idioma estará inserido no país em que a língua alvo é falada, mas poderá vir a não se relacionar em diversas situações comunicativas com os falantes nativos locais. Por outro lado, sempre haverá aprendizes de um novo idioma que, mesmo não tendo a oportunidade de vivenciar a língua no país em que tal sistema é falado, irão se dedicar ao estudo da nova língua e procurar se cercar do *input* estrangeiro a partir de inúmeras ferramentas (seriados, vídeos, músicas, *podcasts*, etc.). Outrossim, o contexto de aprendizagem de L2 pode vir a ser, de fato, ecológico, mas nem sempre o é – assim como alguns contextos de aprendizagem de LE podem vir a ser ecológicos também. Neste trabalho, porém, pensamos em como adaptar as premissas da proposta original de Best e Tyler (2007) para contextos de percepção fônica em LE.

Em nosso entendimento, os pontos positivos do *PAM-L2* residem, principalmente, na filiação do modelo à unidade gestual e nas previsões de assimilação perceptual, uma vez que as categorizações feitas pelos ouvintes são, em grande medida, nelas baseadas. Além disso, como o *PAM-L2* prevê a adoção do gesto articulatório, as estipulações e os resultados advindos de pesquisas que assumem o modelo podem contribuir para o próprio construto da Fonologia Articulatória.

É nossa tarefa, no entanto, esclarecer que algumas instâncias primordiais do modelo articulatório não estão sendo seguidas (seja pela falta de compromisso com a unidade gestual postulada pelo *PAM-L2*, seja por concepções divergentes em relação ao paradigma realista direto ou pela própria dificuldade ao se pensar em um ambiente ecológico de fato). Isso deixa margem para a possibilidade de revisões, ou, até mesmo, da elaboração de um novo modelo (o que seria um objetivo a ser atingido em longo prazo, e que pode ter, em seu cerne, a presente reflexão), decorrente de contínuos reparos e ajustes teóricos que deem conta de retratar mais fielmente o cenário da percepção da fala em LE. Como resultado, estudos futuros (que tentem aprofundar as modificações tanto para LE como para a L2) podem conduzir o modelo para a noção de percepção indireta e para o gesto acústico-articulatório, bem como à equivalência referente aos contextos de LE e L2 quanto à percepção fônica.

Referências bibliográficas

ALBANO, E. **O Gesto e suas Bordas**: Esboço de Fonologia Acústico-Articulatória do Português Brasileiro. Campinas: Mercado de Letras, 2001.

ALLAN, D. **Oxford Placement Test 1**. Oxford: Oxford University Press, 2004.

ALVES, U.; ZIMMER, M. Percepção e Produção dos Padrões de VOT do Inglês por Aprendizes Brasileiros: O Papel de Múltiplas Pistas Acústicas sob uma Perspectiva Dinâmica. **Alfa: Revista de Linguística**, V. 59, p. 157-180, 2015. **crossref** <http://dx.doi.org/10.1590/1981-5794-1502-7>

ANTONIOU, M.; TYLER, M.; BEST, C. Two Ways to Listen: Do L2-Dominant Bilinguals Perceive Stop Voicing According to Language Mode? **Journal of Phonetics**. V. 40, p. 582-594, 2012. **crossref** <http://dx.doi.org/10.1016/j.wocn.2012.05.005>

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M.; CROFT, W.; ELLIS, N.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a Complex Adaptive System: Position Paper. **Language Learning**, V. 59, p. 1-26, 2009. **crossref** <http://dx.doi.org/10.1111/j.1467-9922.2009.00533.x>

BEST, C. A Direct Realist View of Cross-language Speech Perception. In: STRANGE, W. (Ed.). **Speech Perception and Linguistic Experience**: Theoretical and Methodological Issues in Cross-Language Speech Research. Timonium: York Press, 1995. p. 167-200.

BEST, C.; TYLER, M. Nonnative and Second-Language Speech Perception: Commonalities and Complementarities. In: BOHN, O.; MUNRO, M. (Orgs.). **Language Experience in Second Language Speech Learning**: In honor of James Emil Flege. Filadélfia: John Benjamins Publishing Company, 2007, p. 13-34. **crossref** <http://dx.doi.org/10.1075/llt.17.07bes>

BETTONI-TECHIO, M. **Perceptual Training and Word Initial /s/-Clusters in Brazilian Portuguese/English Interphonology**. 2008. 214 f. Tese (Doutorado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2008.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Gestures as Phonological Units. **Haskins Laboratories Status Report on Speech Research**. V. 100, p. 69-101, 1992.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: An Overview. **Haskins Laboratories Status Report on Speech Research**. V. 112, p. 23-42, 1992. **crossref** <http://dx.doi.org/10.1159/000261913>

BROWN, D. Indirect Perceptual Realism and Demonstratives. **Philosophical Studies**. V. 145, p. 377-394, 2009. **crossref** <http://dx.doi.org/10.1007/s11098-008-9237-x>

BROWN, P.; ROEDIGER, H.; MCDANIEL, M. **Make It Stick: The Science of Successful Learning**. Cambridge: Harvard University Press, 2014. **crossref** <http://dx.doi.org/10.4159/9780674419377>

BUNNIN, N.; YU; J. **The Blackwell Dictionary of Western Philosophy**. Malden: Blackwell, 2004.

CHAMOT, A; BARNHARD, S.; EI-DINARY, P.; CARBONARO, G.; ROBBINS, J. **Methods for Teaching Learning Strategies in the Foreign Language Classroom and Assessment of Language Skills for Instruction: Final Report**. Washington: Georgetown University, 1993.

DANCY, J. **Epistemologia Contemporânea**. Cambridge: Harvard University Press, 1985.

DE BOT, K.; LOWIE, W.; VERSPOOR, M. A Dynamic Systems Theory Approach to Second Language Acquisition. **Bilingualism: Language and Cognition**, V. 10, n. 1, p. 7–21, 2007. **crossref** <http://dx.doi.org/10.1017/S1366728906002732>

DELATTRE, P.; LIBERMAN, A.; COOPER, F. Acoustic Loci and Transitional Cues for Consonants. **Journal of the Acoustical Society of America**, V. 27, n. 4, 1955, p. 769-773. **crossref** <http://dx.doi.org/10.1121/1.1908024>

FEIDEN, J.; ALVES, U.; FINGER, I. O Efeito da Anterioridade e da Altura na Identificação das Vogais Médias Altas e Médias Baixas do Português Brasileiro por Falantes de Espanhol. **Letras de Hoje**. V. 49, n. 1, p. 85-94, 2014. **crossref** <http://dx.doi.org/10.15448/1984-7726.2014.1.14715>

FEIDEN, J.; PEROZZO, R.; FINGER, I.; FONTES, A. Percepção de Vogais Médias Altas e Médias Baixas do Português Brasileiro por Falantes de Espanhol em Tarefa de Discriminação Categórica. In: ALVES, U. (Org.). **Aquisição Fonético-Fonológica de Língua Estrangeira: Investigações Rio-Grandenses e Argentinas em Discussão**. São Paulo: Pontes (no prelo).

FLEGE, J. Second Language Speech Learning: Theory, Findings, and Problems. In: STRANGE, W. (Ed.). **Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research**. Timonium: York Press, 1995. p. 233-272.

FOWLER, C. Coarticulation and Theories of Extrinsic Timing. **Haskins Laboratories Status Report on Speech Research**. V. 8, p. 113-133, 1980. **crossref** <http://dx.doi.org/10.3758/BF03211675>

FOWLER, C. An Event Approach to the Study of Speech Perception from a Direct-Realist Perspective. **Journal of Phonetics**, V. 14, p. 3-28, 1986.

FOWLER, C. Invariants, Specifiers, Cues: An Investigation of Locus Equations as Information for Place of Articulation. **Perception & Psychophysics**. V. 55, n. 6, 1994, 597-610.

FOWLER, C. Listeners Do Hear Sounds, not Tongues. **Journal of the Acoustical Society of America**, V. 99, p. 1730-1741, 1996. **crossref** <http://dx.doi.org/10.1121/1.415237>

GASS, S.; SELINKER, L. **Second Language Acquisition: An Introductory Course**. 3. ed. Nova Iorque: Routledge, 2008.

GAZZANIGA, M.; HEATHERTON, T.; HALPERN, D.; HEINE, S. **Psychological Science**. Nova Iorque: W.W. Norton & Company, 2012.

GIBSON, J. **The Ecological Approach to Visual Perception**. Nova Iorque: Psychology Press, 1986.

GOLDSTEIN, L.; FOWLER, C. Articulatory Phonology: A Phonology for Public Language Use. In: SCHILLER, N.; MEYER, A. (Eds.). **Phonetics and Phonology in Language Comprehension and Production**, Mouton de Gruyter, 2003. p. 159-207. **crossref** <http://dx.doi.org/10.1515/9783110895094.159>

GUTIERRES, A.; BATTISTI, E. A Variação da Nasal Velar por Aprendizes Brasileiros de Inglês. **Anais do X Encontro do CELSUL**, 2012.

JACKSON, F. **Perception**. Nova Iorque: Cambridge University Press, 1977.

JACKSON, F. Representative Realism. In: DANCY, J.; SOSA, E.; STEUP, M. (Eds.). **A Companion to Epistemology**. 2. ed. Malden: Blackwell, 2010. p. 702-705.

KLUGE, D. **Brazilian EFL Learners' Identification of Word-Final /m-n/**: Native/Non-native Realizations and Effect of Visual. 2009. 189 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2009.

KLUGE, D. **Perception and Production of Final Nasals by Brazilians Learners of English**. 2004. 175 f. Dissertação (Mestrado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2004.

KOERICH, R. **Perception and Production of Word-Final Vowel Epenthesis by Brazilian EFL Students**. 2002. 261 f. Tese (Doutorado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2002.

KOSSLYN, S.; KOENIG, O. **Wet Mind: The New Cognitive Neuroscience**. New York: Free Press, 1995.

KRAMER, A.; WIEGMANN, D.; KIRLIK, A. **Attention: from Theory to Practice**. Nova Iorque: Oxford University Press, 2007.

LECLERCQ, M.; ZIMMERMANN, P. **Applied Neuropsychology of Attention: Theory, Diagnosis and Rehabilitation**. Londres: Psychology Press, 2005.

LIBERMAN, A. M.; MATTINGLY, I. G. The Motor Theory of Speech Perception Revised. In: **Cognition**, V. 21, p.1-36, 1985. **crossref** [http://dx.doi.org/10.1016/0010-0277\(85\)90021-6](http://dx.doi.org/10.1016/0010-0277(85)90021-6)

LUND, N. **Attention and Pattern Recognition**. Filadélfia: Routledge, 2001.

MCLAUGHLIN, B. **Theories of Second Language Learning**. Londres: Edward Arnold Press, 1987.

MILDNER, V. **The Cognitive Neuroscience of Human Communication**. Nova Iorque: Taylor & Francis Group, 2008.

MOORE, D. H. **The Perception of English Word-Final /l/ by Brazilian Learners**. 2008. 140 f. Dissertação (Mestrado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2008.

MOUND, B. **Perception**. Durham: Acumen, 2003.

NISHIDA, G. **Sobre Teorias de Percepção da Fala**. 2012. 174 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal do Paraná, Curitiba, 2012.

OHALA, J. Speech Perception is Hearing Sounds, not Tongues. **Journal of the Acoustical Society of America**, V. 99, p.1718-1725, 1996. **crossref** <http://dx.doi.org/10.1121/1.414696>

PEROZZO, R. V. **Percepção de Oclusivas Não Vozeadas sem Soltura Audível em Codas Finais do Inglês (L2) por Brasileiros: O Papel do Contexto Fonético-Fonológico, da Instrução Explícita e do Nível de Proficiência**. 2013. 191 f. Dissertação de Mestrado (Mestrado em Letras)—Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

RAUBER, A. **Perception and Production of English Vowels by Brazilian EFL Speakers**. 2006. 218 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2006.

REIS, M. **The Perception and Production of English Interdental Fricatives by Brazilian EFL Learners**. 2006. 186 f. Dissertação de Mestrado (Mestrado em Letras)—Faculdade de Letras, Universidade Federal de Santa Catarina, Florianópolis, 2006.

REIS, M. **The Perception and Production of the English Voiceless Interdental Fricative by Speakers of European French and Brazilian Portuguese**. 2010. 271 f. Tese (Doutorado em Letras)—Faculdade de Letras. Universidade Federal de Santa Catarina, Florianópolis, 2010.

SCHWARTZHAUPT, B.; ALVES, U.; FONTES, A. The Role of L1 Knowledge on L2 Speech Perception: Investigating how Native Speakers and Brazilian Learners Categorize Different VOT Patterns in English. **Revista de Estudos da Linguagem**, V. 23, p. 311-334, 2015.

crossref <http://dx.doi.org/10.17851/2237-2083.23.2.311-334>

SILVA, A. A Variável Tempo nos Estudos em Aquisição. In: BRAWERMAN-ALBINI, A.; GOMES, M. (Orgs.). **O Jeitinho Brasileiro de Falar Inglês: Pesquisas sobre a Pronúncia do Inglês por Falantes Brasileiros**. São Paulo: Pontes, 2014, p. 33-47.

SILVEIRA, R. Uma Análise da Produção Acadêmica na Área da Interfonologia Português-Ínglês. In: RAUBER, A.; WATKINS, M.; SILVEIRA, R.; KOERICH, R. (Orgs.). **The Acquisition of Second Language Speech: Studies in Honor of Professor Barbara O. Baptista**. Florianópolis: Editora Insular, p. 3-19. 2010.

STYLES, E. **Attention, Perception and Memory: an Integrated Introduction**. Nova Iorque: Psychology Press, 2005.

ZIMMER, M.; ALVES, U. Uma Visão Dinâmica da Produção da Fala em L2: O Caso da Dessonorização Terminal. **Revista da ABRALIN**, V. 11, p. 221-272, 2012.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. **Pronunciation Instruction for Brazilians: Bringing Theory and Practice Together**. New Castle upon Tyne, UK: Cambridge Scholars Publishing, 2009.

Artigo recebido em: 30.10.2015

Artigo aprovado em: 16.02.2016