

UNIVERSIDADE FEDERAL DO PARANÁ

Jeniffer Imaregna Alcantara de Albuquerque

**ASPECTOS DA PERCEPÇÃO DA DESSONORIZAÇÃO TERMINAL DO INGLÊS
POR FALANTES NATIVOS DE PORTUGUÊS BRASILEIRO**

CURITIBA

2012

Jeniffer Imaregna Alcantara de Albuquerque

**ASPECTOS DA PERCEPÇÃO DA DESSONORIZAÇÃO TERMINAL DO
INGLÊS POR FALANTES NATIVOS DE PORTUGUÊS BRASILEIRO**

Dissertação em Estudos Linguísticos
apresentada ao Curso Pós-Graduação em
Letras da Universidade Federal do Paraná
para obtenção do título de Mestre em
Letras.

Orientadora: Profa. Dra. Adelaide H. P. Silva

CURITIBA

2012



PARECER

Defesa de dissertação da mestranda JENIFFER IMAREGNA ALCANTARA DE ALBUQUERQUE para obtenção do título de **Mestre em Letras**.

Os abaixo assinados ADELAIDE HERCÍLIA PESCATORI SILVA, UBIRATÃ ALVES e DENISE KLUGE arguiram, nesta data, a candidata, a qual apresentou a dissertação:

“ASPECTOS DA PERCEPÇÃO DA DESSONORIZAÇÃO TERMINAL DO INGLÊS POR FALANTES NATIVOS DE PORTUGUÊS BRASILEIRO”

Procedida a arguição segundo o protocolo que foi aprovado pelo Colegiado do Curso, a Banca é de parecer que a candidata está apta ao título de **Mestre em Letras**, tendo merecido os conceitos abaixo:

Banca	Assinatura	APROVADA Não APROVADA
ADELAIDE H. PESCATORI SILVA		aprovada
UBIRATÃ ALVES		APROVADA
DENISE KLUGE		APROVADA

Curitiba, 22 de agosto de 2012

Prof. Dr. Luis Gonçales Bueno de Camargo
Coordenador

Dedicatória

Irene no Céu

Irene preta
Irene boa
Irene sempre de bom humor.
Imagino Irene entrando no céu:
- Licença, meu branco!
E São Pedro bonachão:
- Entra, Irene. Você não precisa pedir licença.

(BANDEIRA, Manuel. Libertinagem. São Paulo: Nova Fronteira, 2008.)

Para minha vó Irene (*in memoriam*), que era tudo isso e muito mais! Pelo leitinho com canela antes de dormir e pelo incentivo pelas letras.

Agradecimentos

A minha mãe e meu pai, Solange e William, pelo apoio incondicional, pelas refeições quentinhas e por sempre me mostrar que posso fazer melhor. Ao meu irmão, que sempre me estendeu a mão quando eu precisei e que me fez ver que para tudo se dá um jeito. Sem vocês, nada disto teria acontecido.

À Prof^a. Dr^a. Adelaide H.P. Silva pela orientação desde a Graduação até o Mestrado. Agradeço pela ajuda, pela competência e também pela enorme paciência e tolerância com as perguntas constantes.

Aos professores doutores Denise Kluge e Ubiratã Alves pelas generosas contribuições na banca de qualificação e pela leitura final do texto. Um agradecimento especial para a prof^a. dr^a. Denise Kluge pelos diálogos por e-mail que com certeza fizeram parte das reflexões presentes no trabalho.

Aos integrantes do Lefon, Ana Paula, Flávio, Rita, Susana, Julia pelas reflexões teóricas compartilhadas. Em especial, gostaria de agradecer ao Gustavo Nishida, que em boa parte é responsável pelo rigor teórico que eu sempre busco ter, por algumas referências musicais que me acompanharam duração esses 2 anos e pela amizade, acima de tudo.

Aos professores doutores José Borges Neto, Clarissa Jordão, Rodrigo Gonçalves, pelas matérias ministradas no curso de pós-graduação. Elas contribuíram para a pesquisadora que sou hoje.

Ao Prof^o. Dr^o. Maximiliano Guimarães por me questionar sobre meus dados de produção ainda na graduação. As perguntas feitas nortearam este trabalho.

Ao Celin (Celin de Línguas e Interculturalidade) por ter permitido que a pesquisa fosse conduzida com os alunos da instituição e aos alunos que gentilmente se voluntariaram. À coordenação da área de inglês, Mariza, Alessandra, Erika e Álvaro, por terem me incentivado e me ensinado a priorizar as coisas na minha vida pessoal e acadêmica. Um agradecimento especial ao Álvaro, que assumiu algumas funções a mais para que eu pudesse terminar a dissertação. Por isso e pela amizade, obrigada.

À Janice por ter estendido a mão quando eu precisei. Pelo auxílio com a revisão e pelo profissionalismo sempre presente. Obrigada!

À Luana pela amizade cheia de veredas construída ao longo desses 7 anos. Obrigada pela compreensão, parceira e pela leitura do texto final. Em especial, obrigada por tentar fazer com que eu deixasse de ser tão *workaholic* e por continuar tentando.

À Francine pela paciência em entender alguém tão teimosa como eu. Por ter se oferecido gentilmente para ler o texto e pelas contribuições que auxiliaram na versão final. Pela amizade e carinho de sempre. Agradeço, também, à Evelyn, a quem eu passei a conhecer e admirar, pela personalidade forte que me faz perceber outros caminhos.

À Glaucia por ter sido uma das primeiras pessoas que me encaminhou na prática docente, pela paciência em me ensinar os caminhos da sala de aula. Obrigada pelo bom humor e pelas palavras amigas em todas as situações.

À Lívia pela parceria desde a infância até agora. Obrigada por ser um exemplo de perseverança e comprometimento. Ainda que longe, continua me ensinando muito.

A todos os meus amigos da Reitoria, que fizeram parte da minha formação pessoal e acadêmica.

À Fundação Araucária e à CAPES pela bolsa concedida.

Ao Patrick que aceitou a tortuosa tarefa de me acompanhar numa trajetória que espero que dure por muitos anos. Por cuidar de mim mesmo quando isso parecia difícil. Por termos crescido juntos e, com isso, por ter me ensinado a respirar mais leve. Por tudo isso e muito mais: aishiteru.

Epígrafe

“There are things known and there are things unknown, and in between
are the doors of perception.” Aldous Huxley

Sumário

1 INTRODUÇÃO	1
2 MODELOS E ABORDAGENS DE AQUISIÇÃO DE L2	4
3 DESSONORIZAÇÃO TERMINAL (DT): REVISÃO BIBLIOGRÁFICA DE EXPERIMENTOS DE PRODUÇÃO DE BASE GERATIVA	19
3.1 Estudos da década de 1980	20
3.2 A proposta de Oostendorp dentro da OT: (O que fazer com as evidências do dado fonético?)	25
4 A DT SOB UM PONTO DE VISTA DINÂMICO	31
4.1 Experimentos com DT: informantes brasileiros aprendizes de inglês como L2 à luz do Conexionismo e da Fonologia Gestual.	35
4.1.1 DT (?): Investigação da influência de parâmetros acústicos e contextos adjacentes na aquisição de L2 a partir de modelos dinâmicos.....	40
4.1.2 DT: Experimento de produção de oclusivas finais do inglês com falantes Brasileiros a partir da Fonologia Gestual.	44
5 EXPERIMENTOS DE PERCEPÇÃO DE FALA EM L2	60
5.1 Modelos de percepção de fala em L2	61
5.2 A DT em experimentos de percepção: Investigação da influência de parâmetros acústicos e contextos adjacentes	77
6 METODOLOGIA EXPERIMENTAL DE ESTUDOS DE PERCEPÇÃO	87
6.1 Tipos de tarefas: identificação e discriminação.....	88
6.2 Estudos sobre identificação e discriminação	90
7 METODOLOGIA DO PILOTO E EXPERIMENTO	101
7.1 Piloto.....	101
7.1.1 Os informantes	103
7.1.2 O <i>corpus</i> e a gravação de estímulos do teste de percepção	104
7.1.3 O teste de discriminação perceptual.....	110
7.1.4 Resultados do piloto	112
7.2 Experimento	116
7.2.1 Os informantes	117
7.2.2 Método.....	121
7.2.2.1 Gravação dos estímulos e coleta dos dados	121
7.2.2.2 Coleta dos dados dos informantes do GE E GC.	122
7.2.2.3 Corpus.....	123
7.2.2.4 Teste de Discriminação e Identificação	127
7.2.3 Análise Estatística	131
7.2.3.1 Resultados obtidos para o teste de discriminação	132
7.2.3.2 Resumo dos resultados do teste de discriminação: relação entre GE e GC e pontos de articulação.....	138
7.2.3.4 Resultados obtidos para o teste de identificação.	141

7.2.3.5	Resumo dos resultados do teste de identificação: relação entre GE e GC e pontos de articulação.....	147
7.2.3.6	Análise do Tempo de Reação	150
7.2.3.7	Análise da Semelhança entre os Níveis de Proficiência....	152
8	CONSIDERAÇÕES FINAIS	158
8.1	Discussão dos resultados e implicações teóricas	158
8.2	Limitações do estudo e sugestões de pesquisas futuras.	164
9	REFERÊNCIAS BIBLIOGRÁFICAS	166
10	ANEXOS	182
10.1	Questionário sobre o perfil dos informantes brasileiros.	182
10.2	Questionário sobre o perfil do grupo controle	184
10.3	Perfil dos informantes	185
10.3.1	Perfil dos informantes do GE Básico	185
10.3.2	Perfil dos informantes do GE Pré- Intermediário	186
10.3.3	Perfil dos informantes do GE Intermediário	187
10.4	Resultados do Teste de Discriminação po indivíduo.....	189
10.5	Resultados do Teste de Identificação por indivíduo.....	191

Lista de ilustrações

Figura 1 Pauta gestual da palavra “bad” (Browman & Goldstein 1992:158 apud Nishida, 2009: 35). VEL: abertura vélica; CL: corpo de língua; PL: protusão labial; GLO: abertura glotal.....	43
Figura 2 Pauta gestual da palavra “pad” (Browman & Goldstein, 1992:158 apud Nishida, 2009: 36). VEL: abertura vélica; CL: corpo de língua; PL: protusão labial; GLO: abertura glotal.....	44
Figura 3 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento bag na primeira repetição sentença “say bag promptly”, produzida pelo informante GU.....	49
Figura 4 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento bag na segunda repetição da sentença “say bag promptly”, produzida pelo informante GU.....	50
Figura 5 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento cab na primeira repetição sentença “say cab promptly”, produzida pela informante PM.....	50
Figura 6 No espectrograma, entre linhas verticais pontilhadas, a oclusão com a presença da barra de sonoridade do segmento cab na terceira repetição sentença “say cab promptly”, produzida pela informante PM.....	51
Figura 7 O espectrograma da quinta repetição da palavra rag , na sentença “say rag” produzida pelo informante JU.....	52
Figura 8 O espectrograma da primeira repetição da palavra rag, na sentença “say rag”, produzida pelo informante JU.....	53
Figura 9 No espectrograma, entre linhas verticais pontilhadas, a barra de sonoridade e a não produção do burst, na primeira repetição sentença “say cab”, produzida pelo informante PM (nativo).....	54
Figura 10 Gráfico do percentual de duração de vogal para o contexto 1.	56
Figura 11 Gráfico do percentual de duração de vogal para o contexto 2.	56
Figura 12 Gráfico do percentual de duração de vogal para o contexto 3.	57
Figura 13 Gráfico do percentual de vozeamento para o contexto 1.	58
Figura 14 Gráfico do percentual de vozeamento para o contexto 2.	58
Figura 15 Gráfico do percentual de vozeamento para o contexto 3.	59

Figura 16 O espectrograma da sentença “say bad promptly” produzida pela informante PM (nativa), sem a manipulação da palavra-alvo “bad”.	107
Figura 17 O espectrograma da sentença “say bad promptly” produzida pela informante PM (nativa), com a manipulação da palavra-alvo “bad”, sem burst, sem barra de sonoridade (indicação visual do parâmetro vozeamento).	107
Figura 18 O espectrograma da sentença “say bat promptly”, sem manipulação, produzida pela informante PM (nativa). Duração da vogal da sequência “bat”: 0,2345 ms.	108
Figura 19 O espectrograma da sentença “say bad promptly”, sem manipulação, produzida pela informante PM (nativa). A duração da vogal é da sequência “bad” é de 0, 3459 ms.	109
Figura 20 O espectrograma da sentença “say bad promptly”, com a duração da vogal manipulada, produzida pela informante PM (nativa). A duração da vogal é da sequência “bad” é de 0,2497 ms.	109
Figura 21 Tela do Praat visualizada pelos informantes ao receberem o estímulo auditivo.	111
Figura 22 Tela do Praat visualizada pelos informantes ao clicarem na resposta que julgassem correta.	111
Figura 23 Esquema informativo da quantidade de dados obtidos.	112
Figura 24 Esquema informativo dos dados dos experimentos de identificação e discriminação.	125
Figura 25 Espectrograma do estímulo bat. Entre as linhas pontilhadas a duração da vogal: 0.1386 ms.	126
Figura 26 Espectrograma do estímulo bad sem a manipulação. Entre as linhas pontilhadas a duração da vogal: 0.1895 ms.	126
Figura 27 Espectrograma do estímulo bad com a manipulação da duração da vogal. Entre as linhas pontilhadas a duração da vogal: 0.1338 ms.	127
Figura 28 Tela do Praat visualizada pelos informantes no teste de discriminação.	128
Figura 29 Tela do Praat visualizada pelos informantes no teste de discriminação ao clicarem na resposta que julgassem correta.	129
Figura 30 Tela do Praat visualizada pelos informantes ao receberem o estímulo auditivo no experimento de identificação.	130
Figura 31 Tela do Praat visualizada pelos informantes no experimento de identificação ao clicarem na resposta que julgassem correta.	131

Figura 32 Gráfico com a média de acertos das bilabiais do teste de discriminação. Notação: p,b1= Dado Completo ;p,b2= padrão Duração de Vogal; p,b3= Vozeamento.	140
Figura 33 Gráfico com a média de acertos das alveolares do teste de discriminação.	140
Figura 34 Gráfico com a média de acertos das velares do teste de discriminação.	141
Figura 35 Gráfico com a média de acertos das velares do teste de identificação.	149
Figura 36 Gráfico com a média de acertos das velares do teste de identificação.	150
Figura 37 Gráfico com a média de acertos das velares do teste de identificação.	150
Figura 38 Valores do teste de Spearman para o teste de discriminação.	151
Figura 39 Valores do teste de Spearman para o teste de identificação.	152
Figura 40 Box plot da média de acertos por nível para o teste de discriminação.	154
Figura 41 Box plot da média de acertos por nível para o teste de identificação.	155

Lista de Quadros

Quadro 1 Estímulos utilizados no experimento piloto de percepção com oclusivas finais do inglês.....	105
Quadro 2 Resumo dos dados dos informantes do GE. X: média; LE: Língua Estrangeira.....	120
Quadro 3 Perfil dos informantes do GC.....	121
Quadro 4 Resumo dos resultados para o teste de discriminação.....	138
Quadro 5 Resumo dos resultados para o teste de identificação.....	147

Lista de tabelas

Tabela 1 Perfil dos informantes falantes de PB como L1 que participaram do piloto.	104
Tabela 2 Resultados do experimento de discriminação para as oclusivas bilabiais. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais ou diferentes.....	113
Tabela 3 Resultados do experimento de discriminação para as oclusivas alveolares. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais e diferentes.....	114
Tabela 4 Resultados do experimento de discriminação para as oclusivas velares. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais e diferentes.....	114
Tabela 5 Resultados da média e percentual de acertos para o teste de discriminação para o GE Básico. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em percentagem para cada padrão; % DP = percentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.	133
Tabela 6 Resultados da média e percentual de acertos para o teste de discriminação para o GE Pré-Intermediário. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em percentagem para cada padrão; % DP = percentagem do desvio padrão; DC: Dado Completo; DV: Duração de Vogal; V: Vozeamento.	134
Tabela 7 Resultados da média e percentual de acertos para o teste de discriminação para o GE Intermediário. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em percentagem para cada padrão; % DP = percentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.	136
Tabela 8 Resultados da média e percentual de acertos para o teste de discriminação para o GC. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em percentagem para cada padrão; % DP = percentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.	137
Tabela 9 Resultados do teste de Wilcoxon para o teste de discriminação para o GC. Z: Wilcoxon; p: p valor.....	137
Tabela 10 Resultados da média e percentual de acertos para o teste de identificação para o GE Básico. Notação: N = número total	

de dados; X = média de acertos para cada padrão; % X = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.	142
Tabela 11 Resultados da média e percentual de acertos para o teste de identificação para o GE Pré-Intermediário. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.	143
Tabela 12 Resultados da média e percentual de acertos para o teste de identificação para o GE Intermediário. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração de Vogal; V: Vozeamento.	144
Tabela 13 Resultados da média e percentual de acertos para o teste de identificação para o GC. Notação: N = número total de dados; X = média de acertos para cada padrão; % X = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; Vozeamento.....	146
Tabela 14 Resultados do teste de Wilcoxon para o teste de discriminação para o GC. Z: Wilcoxon; p: p valor.....	146

Lista de abreviações

A = estímulo auditivo

AV = estímulo áudio-visual

CL = corpo de língua

DC = Dado Completo

DV = Dado com a manipulação da duração de Vogal

V = Dado com a manipulação do Vozeamento

DIM = diminutivo

DT = Dessonorização Terminal

GC = Grupo Controle

GE = Grupo Experimental

GLO = gesto largo de glote

GU = Gramática Universal

H1 = Hipótese 1 segundo o SLM

I-I = estímulos inglês-inglês

I-J = estímulo inglês-japonês

J-J = estímulos japonês-japonês

L1 – língua 1/língua materna

L2 = língua 2/língua estrangeira

LAD = language acquisition device

LEFON = Laboratório de Estudos Fônicos

OT= Teria da Otimalidade

PAM = Perceptual Assimilation Model

PB = Português Brasileiro

PL = ponta de língua

SLM = Speech Learning Model

SPE = Sound Pattern of English

V = estímulo visual

VOT = voice onset time

[+voc] = traço + vocálico

[ant] = traço anterior

Resumo

A discussão sobre a Dessonorização Terminal (DT) ganhou força com os trabalhos de Eckman (1981, 1987) e Major (1987), os quais a descreveram como sendo a perda do traço [sonoro] em algumas obstruintes em posição final. Contudo, o fenômeno ganha novos desdobramentos quando alguns estudos passam a relatar dados de neutralização incompleta (e.g. Pye, 1986 – para o russo – e Slowiaczek e Dinnsen, 1985 – para o polonês). Do mesmo modo que os trabalhos de produção, Shrager (2002), Warner, Jongman, Sereno e Kempes (2004), Broersma (2005), Kleber, John e Harrington (2010) demonstraram que os aprendizes de L2 são capazes de perceber o contraste surdo-sonoro em posição final, a partir do reconhecimento de diferentes pistas acústicas, a saber: duração de vogal; duração de vozeamento; e *burst*. No entanto, apesar de os estudos de produção e percepção corroborarem o fato de que a DT não pode ser tratada como um fenômeno binário, como havia sido descrito por Eckman (*op. cit.*) e Major (*op. cit.*), como incorporar o detalhe fonético fino à representação fonológica? Alguns trabalhos, Zimmer (2004), Zimmer e Alves, (2007, 2008), Oostendorp (2007) e Albuquerque (2010) ofereceram um novo tratamento para a DT, sem descartar os achados fonéticos. Oostendorp (*op.cit.*) propõe uma nova representação da DT a partir da Teoria da Otimalidade (OT) e, com isso, reconhece a gradiência presente nos dados da DT, mas não a incorpora à representação. Ou seja, embora Oostendorp trabalhe com dados gradientes, o primitivo de base permanece categórico. Nesse sentido, foi realizado um experimento de percepção com falantes de PB como L1, aprendizes de inglês como L2, a fim de que se analisasse a gradiência a partir de modelos que tenham como primitivos simbólicos os gestos articulatorios, e a observação do fenômeno da Dessonorização Terminal (DT) para esses dados. O experimento contou com três grupos (com 20 sujeitos em cada) de diferentes níveis de proficiência do inglês (básico, pré-intermediário e intermediário) e um grupo de 8 falantes nativos da Carolina do Norte que realizaram dois testes perceptuais no Praat: discriminação e identificação. Os resultados do experimento mostraram fortes indícios para que uma relação entre ponto de

articulação e parâmetro acústico (duração de vogal e vozeamento) seja estabelecida, tanto para informantes nativos como para aprendizes. Foram corroborados os achados de Warner *et al.* (2004) e Kleber *et al.* (2010), que relataram a possibilidade de um aprendiz perceber um contraste da L2 que não existe na sua L1. Além disso, o tempo de reação não foi significativo para que uma relação com o tipo de experimento fosse estabelecida. Contudo, observou-se uma correlação fraca com a variável repetição, uma vez que a cada repetição, o informante teve uma maior média de acertos. Os resultados de ambos os testes indicam que para que a DT receba um tratamento mais parcimonioso é necessário que esta seja analisada via modelos dinâmicos (Browman e Goldstein, 1986, 1989, 1990, 1992), uma vez que os informantes foram capazes de perceber as pistas necessárias para a distinção do contraste sonoro em posição final de modo gradiente. Além disso, os resultados parecem se encaixar nas categorias estabelecidas pelo *Perceptual Assimilation Model L2* (PAM; Best and Tyler, 2007). Os resultados em geral apontam para uma aproximação entre a percepção de nativos e aprendizes, uma vez que as dificuldades e acertos no reconhecimento de alguns parâmetros foram semelhantes.

Palavras-chave: percepção na L2; Dessonorização Terminal (DT); *design* experimental.

Abstract

The discussion about Terminal Devoicing (TD) emerged with the studies of Eckman (1981, 1987) and Major (1987) who described it as the loss of a [voice] feature in some final stops. However, the phenomenon is revisited when incomplete neutralisation data is found (e.g. Pye, 1986 – for the Russian – and Slowiaczek and Dinnsen, 1985 – for the Polish). In addition to production studies, Shrager (2002), Warner, Jongman, Sereno and Kemps (2004), Broersma (2005), Kleber John and Harrington (2010) noticed that L2 learners are able to distinguish final sonorous contrast by using different acoustic cues: length of the vowel that precedes the obstruent, length of voicing and burst. Although production and perception studies observed that TD cannot be seen as a binary phenomenon, as it had been previously described by Eckman (*op. cit.*) and Major (*op. cit.*), how is it possible to incorporate the phonetic detail to phonology? Zimmer (2004), Zimmer e Alves, (2007, 2008), Oostendorp (2007) and Albuquerque (2010) shed light on TD by offering a new analysis when incorporating the phonetic data. Oostendorp (*op. cit.*) proposes a new representation for TD in the light of Optimality Theory (OT). By giving this step forward, the author recognizes the gradient TD data, but he does not incorporate gradience inside phonology. That is, although Oostendorp works with gradient data, his unit analysis remains categoric. By taking this discussion into account, a perception experiment with students who had PB as L1 and were learning English as L2 was conducted. It aimed to analyze the gradience from models which work with gestures as their unit analysis; also, this work had the objective of observing the TD phenomenon for the analyzed data. Three groups of Brazilian students from different proficiency levels – basic, pre-intermediate and intermediate (20 people in each group) – and one group of 8 native speakers from North Carolina participated in the experiment. Each group took both the discrimination and the identification tests in Praat. The experiments' results showed a strong relationship between the place of articulation and the acoustic cues (the vowel that precedes the obstruent and the length of voicing), both for native speakers and L2 learners. The findings

were similar to the ones for Warner *et al.* (2004) and Kleber *et al.* (2010). These works have shown the possibility of a learner to perceive a L2 contrast which does not exist in his/her L1. In addition, the reaction time measured in the perception study was not considered relevant for differentiating the experiment type, but it showed weak relevance to the repetition sequence, since in each repetition the learner was able to get more correct answers. Both tests results indicate that a more adequate approach to treat TD would be to analyse it through dynamic models (Browman and Goldstein, 1986, 1989, 1990, 1992). This may be stated once learners were able to perceive the necessary clues to distinguish the gradient in the sonority contrast in final position. Moreover, the results seem to agree to the categories established by the *Perceptual Assimilation Model L2* (PAM; Best and Tyler, 2007). The general results lead us to notice that native speakers and learners' perception is similar. Both groups of speakers presented the same difficulties and amount of correct answers.

Key words: L2 perception; Terminal Devoicing (TD); experiments' methodology.

1 INTRODUÇÃO

Este trabalho teve seu início a partir de um experimento realizado durante um projeto de Iniciação Científica (IC) na Universidade Federal do Paraná (UFPR), conduzido em parceria com o grupo da UCPel (Universidade Católica de Pelotas), coordenado pela profa. dra. Márcia Zimmer no ano 2010.

Esse experimento analisava a produção de oclusivas¹ finais, surdas e sonoras, do inglês, realizadas por informantes paranaenses aprendizes de inglês como segunda língua. Na ocasião da apresentação desse trabalho de IC, a banca examinadora questionou sobre as relações entre os resultados do experimento de produção² conduzido e sua contraparte perceptual. Dentre os resultados, o trabalho apontava para uma gradiência na sonorização das produções dos informantes, consideração que se situa dentro de um arcabouço teórico que atribui natureza dinâmica à produção dos sons da fala (Browman & Goldstein, 1986 e os seguintes; Albano, 2001). Contudo, como a gradiência encontrada na produção se verificaria na percepção? Os informantes aprendizes são capazes de reconhecer tal gradiência ou, para a percepção, a informação encontrada nos dados de produção seria categórica?

Tais questionamentos instigaram a investigação sobre as pistas acústicas que conduzem os informantes a reconhecerem um determinado som como tal. Para tanto, este trabalho propõe, num primeiro momento, que se faça um levantamento sobre estudos de aquisição de segunda língua e como a noção de “erro” é tratada em modelos teóricos distintos (capítulo 2). Tal investigação é necessária uma vez que modelos distintos procuram dar um tratamento, de acordo com seus pressupostos teóricos, a fenômenos linguísticos pertencentes ao processo de aquisição de L1³ e L2⁴. Ao mesmo

¹ Os termos “oclusiva” e “plosiva” serão tratados como sinônimos neste trabalho.

² Este experimento será posteriormente relatado (seção 4) a fim de se discutir a relação entre produção e percepção.

³ Neste trabalho, L1 e língua materna são sinônimos.

⁴ Embora haja uma distinção teórica entre L2 (Língua 2) e LE (Língua estrangeira), não faz parte dos objetivos deste traçar diferenças entre os termos.

tempo, é importante para algumas abordagens explicar como os aprendizes lidam com o sistema da L2 e de que maneira se dá a adequação das novas formas adquiridas. Na tentativa de produzir sequências corretas na L2-alvo, o aprendiz comete erros, desvios de produção, que acusam a sua não proficiência na L2. Contudo, o modo como essa falha é vista pelas diversas abordagens, que lidam não só com o processamento linguístico, mas também com a relação ensino-aprendizagem, mostra-se determinante na investigação da L1 e L2.

De forma mais geral, serão mencionados também alguns fenômenos que são contemplados ou deixados de fora dos diferentes paradigmas teóricos. A partir da discussão realizada, no capítulo 3, toma-se o fenômeno da dessonorização terminal a fim de discutir os resultados de experimentos de produção a partir de modelos de base gerativa e, no capítulo 4, passa-se à resenha de experimentos de produção de base dinâmica e a implicação direta desses achados para a percepção, de modo a entender como os resultados podem contribuir para a discussão sobre a gradiência ou discretude na percepção. A partir dessa discussão, no capítulo 5, o trabalho se desenvolve com o objetivo de investigar as pistas acústicas envolvidas na percepção de oclusivas finais do inglês em diversas línguas. A partir da descrição desses estudos de percepção, no capítulo 6, discorre-se sobre questões relacionadas à metodologia experimental utilizada em experimentos de percepção em L2, uma vez que essa pode influenciar nos resultados finais de um experimento (Pisoni, 1975, *apud* BRASILEIRO, 2009). Diante dessa exposição, o fenômeno da dessonorização terminal será analisado a partir de um piloto e de um experimento de percepção com falantes brasileiros e os respectivos resultados perceptuais serão discutidos no capítulo 7. Os resultados do piloto serviram como base para que a metodologia do experimento fosse reformulada. Entre os achados do piloto se encontra a utilização de pistas secundárias por alguns informantes e a predileção por um dado parâmetro acústico na percepção do contraste surdo-sonoro em posição final. Com a metodologia reformulada, o experimento foi organizado a partir dos seguintes objetivos: a) a influência do

ponto de articulação na discriminação e identificação do contraste surdo-sonoro para oclusivas finais do inglês para falantes de PB como L1; b) se existe uma diferença significativa entre um maior percentual de acerto entre diferentes níveis de fluência (nível básico, pré-intermediário e intermediário); c) a influência do tipo de tarefa (identificação e discriminação) dada ao sujeito na identificação do contraste em posição final; d) a correlação entre o tempo de reação para responder aos estímulos e os dados tidos como “errados”; e) semelhanças e diferenças na produção da DT para o grupo experimental (GE) e o grupo controle (GC), através da comparação entre os resultados obtidos para ambos os grupos. O experimento reforçou o achado de que os alunos parecem utilizar mais de uma pista na discriminação e identificação do contraste surdo-sonoro e que, além disso, essa escolha estaria relacionada ao ponto de articulação, a saber: bilabial, alveolar e velar.

Por fim, no capítulo 8, apontam-se as considerações finais acerca dos resultados do experimento, as possíveis implicações para a análise de fenômenos de mesma natureza e os desdobramentos no ambiente da sala de aula.

2 MODELOS E ABORDAGENS DE AQUISIÇÃO DE L2

Vários estudos que lidam com fenômenos ligados à aquisição de L1 e L2 se debruçaram sobre as semelhanças e distinções de ambos os processos de aquisição (L1 e L2), a exemplo dos que serão citados abaixo. Brown (1977) já fazia menção ao fato de que, embora as produções não sejam iguais em um e outro processo (aquisição de L1 e L2), existe uma semelhança, por exemplo, das sequências gramaticais em algumas línguas, aspecto que facilitaria o processo de aquisição da língua-alvo pelo aprendiz. No entanto, ao longo do aprendizado da língua materna e de uma L2, o falante-aprendiz testa constantemente as estruturas a que é exposto e produz, muitas vezes, sequências que são consideradas desviantes, os chamados “erros”. Assim, nesta introdução, faremos um caminho por alguns modelos e abordagens teóricas que tratam da aquisição de L1 e L2, em particular, sobre o segundo processo, de modo a observar como essas falhas de produção são incorporadas ou não dentro deles. Além de Brown (*op. cit.*), Krashen (1982) também se ocupou de processos de aquisição de L2. Nesse estudo, a questão gira em torno da aquisição de uma língua estar conectada com o que o autor chama de “filtro afetivo”, ou seja, mesmo que um dado aprendiz compreenda um enunciado, se este tiver barreiras com relação ao aprendizado da L2-alvo, não conseguirá produzir sentido com a informação ouvida. De acordo com Stevick (1976), o aluno não conseguiria fazer uso da informação, pois o *input* recebido não atingiria a parte do cérebro responsável pela aquisição de L2. Ainda que este estudo não tenha como objetivo posterior o de questionar os conceitos de aquisição e aprendizagem⁵, é importante ressaltar que, para Krashen, existem diferenças entre ambos os conceitos. O primeiro está relacionado a um processo natural de aprendizagem e o segundo é resultado de instrução formal.

⁵ Embora este trabalho não tenha como objetivo se estender na distinção entre os conceitos de aquisição e aprendizagem, será mantida a distinção feita por Krashen, conforme mencionado nesta página.

Pode-se dizer, então, que estudos como o de Krashen nos guiam a uma noção de aquisição e aprendizagem de L2 que se encontra pautada em acertos e erros. A motivação de trabalhos como os de Krashen parece ser a descoberta de como fazer com que o aprendiz de L2 seja bem sucedido em seu aprendizado, de maneira a descrever as limitações do aprendiz com relação aos mecanismos da língua-alvo que ele desconhece.

Com relação ao desenvolvimento do processo de aquisição de L2 relacionado à pronúncia e aos trabalhos que se debruçam sobre a produção de fala em L2, pode-se dizer que a noção de “erro”⁶ foi sendo vista de maneira distinta a depender do paradigma teórico. A partir de uma visão behaviorista de aprendizagem, a noção de “erro” é vista como a “incapacidade” do falante-aprendiz de perceber os sons da língua-alvo (OLIVEIRA, 2011). Essa visão behaviorista do processo de aquisição acaba por encerrar a produção do falante em categorias, de modo que um indivíduo deve adquirir o sistema da L2 que está estudando, através da imitação e repetição das estruturas da língua-alvo. Nesta concepção, há também a noção de uma correspondência biunívoca entre o processo de percepção e o de produção de fala. Assim, quando se pensa em produção e percepção de fala, a implicação desse paradigma teórico é a de que, se o indivíduo imita errado é porque recebeu um estímulo incorreto e, portanto, não seria capaz de reproduzir o som de maneira adequada. O “erro” é visto como uma falha de assimilação dos sons da L2, mas é passível de ser eliminado a partir do reforço positivo, da repetição da forma-alvo correta. A consequência mais extrema com relação a uma produção incorreta é a de se assumir que o falante não é capaz de produzir um determinado som da L2, porque ele não o percebeu.

Nesse paradigma teórico, o desenvolvimento da linguagem é tido como um comportamento, e, no caso da aquisição de uma L2, vincula-se à ideia de que o aprendiz irá transferir tais hábitos da L1 para a L2 (LADO, 1964). Dito

⁶ A partir deste momento, ainda que uma distinção com relação a tipos de desvio de produção seja feita, o vocábulo “erro” aparecerá entre aspas a fim de marcar um posicionamento teórico, o qual será discutido ao longo do texto.

isso, pode-se assumir que esta abordagem é, frequentemente, alocada junto à Hipótese da Análise Contrastiva, que surge no final da década de 1950, embasada no estruturalismo norte-americano, rendeu um maior número de trabalhos na década de 1960. Tal hipótese postulava que os sons da L2 que fossem semelhantes aos sons da L1 seriam mais facilmente adquiridos e, os que apresentassem um contraste maior, distintos dos existentes na L1, ofereceriam uma dificuldade maior para o aprendiz no processo de assimilação. Entre as questões mais importantes dessa corrente teórica, destacam-se a noção de transferência e a noção de interferência da L1 na aquisição da L2, as quais são tratadas por autores como Weinreich (1953), Lado (1957) e Lee (1968)⁷.

No entanto, apesar de essa abordagem ter embasado os primeiros estudos sobre o processo de aquisição de L2, a abordagem behaviorista (cujos trabalhos tiveram início no final da década de 1950) não se propõe a explicar algumas questões que são levantadas ao longo do processo de ensino-aprendizagem da produção de fala. Embora se tome a imitação como parte do aprendizado, sabe-se que os falantes não foram expostos a todo o conjunto de sentenças que, no entanto, são capazes de produzir. Além disso, pesquisas como a de Bloom, Hood e Lightbown (1974) relatam que as crianças, no processo de aquisição da L1, parecem imitar as estruturas da língua até que esse conhecimento se torne sólido, então param de imitar e iniciam um processo de fazer generalizações sobre o sistema da língua que estão adquirindo. O que se questiona a partir desse momento, então, é como seria possível fazer tais generalizações uma vez que, a princípio, todo o conhecimento linguístico é recebido através da imitação.

Assim, essa visão behaviorista aplicada ao ensino de línguas estrangeiras começa a ser suplantada, por um lado, pelo surgimento de novos modelos dentro da psicologia, que passam a ser empregados na relação de ensino/aprendizagem, como o modelo interacionista ou o modelo cognitivista,

⁷ O primeiro autor a tecer informações sobre a Hipótese Contrastiva foi Charles Fries. Para maiores informações, verificar Fries (1945).

e, por outro lado, dentro da linguística, pela necessidade de se ter modelos explicativos que investiguem como os falantes assimilam e constroem regras sobre um dado sistema linguístico.

É em meio a esse cenário que as abordagens inatistas (década de 1960) ganham força e se propõem a investigar alguns aspectos sobre a aquisição de L1 e L2, os quais haviam sido desconsiderados por uma abordagem de base descritiva. Ao contrário do behaviorismo, que vê o indivíduo aprendiz como uma “tábula rasa”, as teorias inatistas, que têm como um de seus precursores Noam Chomsky, preconizam que o aprendiz faz parte do processo de aprendizagem no sentido de que ele não é apenas influenciado pelo ambiente que o cerca, em que recebeu parte de seu conhecimento sobre a língua, mas também é capaz de levantar hipóteses sobre a estrutura da língua que está adquirindo, de modo a estabelecer e criar relações entre as diferentes estruturas que lhe são apresentadas. Desse modo, as teorias inatistas procuram investigar como a gramática se desenvolve na mente do falante e o papel da mente humana nesse processo. Diante disso, os processos de aquisição de linguagem não se originam da imitação, como na abordagem behaviorista, mas são determinados por princípios mentais, organizados a partir de uma ordem específica e alimentados por conteúdo linguístico. As estruturas mentais vão se formando a partir do que Chomsky (1965) chama de *language acquisition device* (LAD), um mecanismo composto de um conjunto finito de dimensões através das quais as línguas se organizam. Ele é responsável por agrupar a informação recebida, a qual irá preencher categorias gramaticais que, de acordo com os pressupostos de base desse paradigma teórico, são inatas, para posteriormente se obterem as regras da língua que será, em um momento seguinte, produzida. Uma vez que o LAD⁸ é um mecanismo, quando ativado, ele faz com que o aprendiz seja capaz de processar e, ao mesmo tempo, descobrir as regras que operam o sistema

⁸ Mais tarde, o LAD é substituído pela noção da Gramática Universal (GU), a qual é tida como um conjunto de princípios, os quais são seguidos por todas as línguas existentes, como enunciam Chomsky (1981), Cook (1991) e White (1989).

linguístico, a língua que se está adquirindo. É importante ressaltar que quando Chomsky descreve o funcionamento do LAD, o autor tem em mente aquisição de L1. Mas no que concerne a aquisição de L2, é possível dizer que, conforme menciona Galasso (2002), a operação do mecanismo (LAD) dentro da mente, de um falante-aprendiz adulto, é distinta, funcionando de modo distinto do processo de aquisição de informação linguística da L1.

Mas o conhecimento inato não é o único que opera sobre as generalizações feitas pelo aprendiz: o conhecimento que é adquirido pelo meio no qual o falante vive também possui um papel importante. Contudo, diferentemente da abordagem behaviorista, em que o meio é responsável por todo o processo de aquisição, a questão para as abordagens inatistas é a de que, tendo o falante sido exposto a um ambiente em que a língua-alvo é falada, as condições do ambiente não possuem influência no plano da representação simbólica, na competência linguística do falante para operar sobre aquela determinada língua. Tal competência é inata aos indivíduos.

Embora Chomsky não postule nenhuma implicação do seu modelo teórico para a aquisição de L2, vários teóricos lançaram mão dos pressupostos da Gramática Universal (GU) para discutir aspectos relacionados à L2, chegando a considerar a GU como um dos melhores modelos para compreender a aquisição de L2 (GASS & SCHACHTER, 1989). Contudo, alguns autores, como Rocca (2003), argumentam contra o fato de que a GU dá conta de questões de L2, e isso se deve, em grande parte, à noção de “período crítico” de desenvolvimento linguístico, segundo o qual os aprendizes possuem um limite ótimo para aquisição de uma língua, uma vez que o mecanismo inato de aquisição de linguagem funciona de maneira mais efetiva quando é estimulado no período certo. A literatura trouxe à tona alguns casos (ITARD, 1962) que corroboram a teoria do período crítico, como o do menino Victor, que foi encontrado perdido e crescera em um ambiente escuro e isolado de qualquer contato humano. Foi realizado um trabalho de inserção do menino no meio social e se tentou prover-lhe instruções sobre utilização de língua. No entanto, apesar de o menino ter desenvolvido a memória e seu relacionamento

com outras pessoas, ele não conseguiu adquirir língua, de acordo com os pressupostos gerativistas expostos até este ponto – ele não foi capaz de criar generalizações sobre a língua utilizada no ambiente e produzir sentenças mais ou menos bem estruturadas de acordo com a gramática da língua em questão. No entanto, apesar de casos como o de Victor endossarem a existência do período crítico, tal noção não é vista de forma pacífica por autores que tomam teorias de base inatista para explicar fenômenos da aquisição de L2. Como discorre White (1989), os pesquisadores de L2 não possuem um posicionamento único com relação à possível aplicação dos princípios inatistas aos estudos de aquisição de L2, pois, tomando ao menos esse postulado sobre o período crítico, como se explicaria a aquisição de uma L2 na fase adulta? A princípio, ela só ocorreria se os sistemas de ambas as línguas fossem idênticos, mas a literatura da área⁹ aponta que, apesar de termos algumas semelhanças, existem também diferenças significativas nas gramáticas de duas línguas.

No mesmo período em que emergem as teorias inatistas, na década de 1970, outros modelos de aquisição de L2 procuram trabalhar com a noção de “erro” baseando-se em princípios inatistas. Tais modelos têm como pressuposto que o falante possui um papel ativo no processo de aquisição e os “erros” cometidos não recebem reforço positivo, eles fazem parte do aprendizado e auxiliam como uma espécie de termômetro sobre o ponto em que se encontra o desenvolvimento do aprendiz. O modelo que surge baseado nesse pressuposto inatista foi conhecido como Modelo de Análise de Erros, tendo como trabalhos afiliados a esse modelo teórico Corder (1967).

Dois tipos de “desvios” são relatados por Corder (*op cit.*): *mistakes* e *errors*. Embora não exista uma tradução para o português que distinga essas duas palavras, a partir da descrição feita pelo autor, pode-se traduzir *mistakes* por “equivocos” – ou seja, relacionados a desvios casuais – e traduzir *errors* por “erros” – desvios que ocorrem de modo sistemático. Os equivocos

⁹ Referências para essa afirmação podem ser encontradas no início da Introdução deste estudo.

acontecem devido a um lapso de memória, relacionado ao desempenho do falante. Nesse sentido, eles podem aparecer quando o indivíduo-aprendiz está cansado ou quando está sob algum tipo de pressão emocional, e, assim, esses equívocos são assistemáticos, sendo produzidos quando os falantes não aplicam, apropriadamente, as regras da língua. Contudo, segundo Corder (*op. cit.*), tanto os aprendizes de L2 quanto os falantes nativos podem cometer esse tipo de “erro”. Por outro lado, os erros se dão pelo desconhecimento de algumas regras por parte do aprendiz, e, portanto, são relacionados a erros de competência.

Outros autores, como Edge (1989), oferecem outra classificação para os erros cometidos pelos aprendizes, dividindo-os em: *slips*, *errors*, *attempts*. É importante ressaltar a importância da inserção dos *attempts* (tentativas), uma vez que isso confere ao aprendiz o papel ativo dentro do seu aprendizado, pois ele faz uso do repertório dos padrões da L2 que possui para se comunicar. Outros autores¹⁰ trabalham com classificações distintas para os “erros”, mas estas não serão resenhadas nesse momento.

No entanto, embora o “erro” seja visto como parte do aprendizado e não como algo a ser excluído, não existe também um consenso no que diz respeito ao papel da correção dos “erros” cometidos pelos alunos em fase de aquisição de uma L2 com relação ao que é internalizado na gramática desses falantes. Autores que adotam a existência da GU possuem, contudo, visões divergentes acerca de como a correção afeta o processo de aquisição de uma L2. Krashen (1982) e Schwartz (1993) discorrem sobre o efeito não eficaz da correção, uma vez que o aprendiz adulto não modifica, na maioria das ocorrências, a sua produção após receber um estímulo correto, pois o conhecimento sistemático abstrato da língua não é afetado ao receber a forma correta; apenas a forma superficial linguística, o *output*, se altera.

Associada a esse entendimento do “erro” como sendo parte do processo de aprendizagem do falante de uma L2, começa a ser veiculada a noção de

¹⁰ Richards (1985:173) menciona a ocorrência de erros de interlíngua, os quais se originam da transferência de itens da L1 para a L2.

Interlândia. Essa noção foi inserida na literatura de aquisição de L2 por Selinker (1972) a partir de observações referentes ao modo de sistematização do conteúdo linguístico feito pelos indivíduos-aprendizes de uma L2. A interlândia possui um sistema chamado “intermediário”, que se caracteriza pela transferência que os padrões estabelecidos na L1 têm nas produções em L2 (ELLIS, 1997). Contudo, com relação à transferência de comportamentos entre a L1 e a L2, Ellis (*op. cit.*) menciona que existe uma diferença na maneira como esse conteúdo é passado de uma língua para a outra, no sentido de que as estruturas internalizadas pelos falantes são distintas e, por vezes, esse processo de transferência pode se mostrar parcial ou completo.

Embora, como foi dito acima, a noção de “erro” da Interlândia esteja relacionada parcialmente às correntes inatistas, é possível dizer que dentro dos pressupostos de base da Interlândia encontra-se um direcionamento conexionista para o aprendizado. Os primeiros estudos conexionistas datam de 1943, com os trabalhos do neurofisiologista, filósofo e poeta Warren McCulloch, e do lógico Walter Pitts. Esses autores possuem um papel importante no início dos estudos conexionistas por terem desenvolvido o primeiro modelamento, através da matemática, para as sinapses realizadas entre os neurônios.

Os estudos conexionistas, com relação à aquisição de linguagem, possuem uma distinção basilar quando comparados aos inatistas no sentido de que, para os conexionistas, pode-se falar em comportamentos universais das línguas, mas isso não necessariamente estaria associado a um conhecimento, um processo que se encontra nos genes, ou seja, como apontam Elman, Bates, Johnson, Karmiloff-Smith, Parisi, Plunkett (1996), o conteúdo inato (como afirmam existir os inatistas) é distinto dos mecanismos inatos (os quais fazem parte de uma visão conexionista). Para Elman e outros autores, como Ellis (1999), não existe, portanto, como foi mencionado acima, um conhecimento inato, mas uma série de conexões sinápticas, a partir das quais é gerado o desenvolvimento linguístico de um dado aprendiz. O conexionismo se distancia, portanto, de abordagens inatistas, pois não entende a linguagem como faculdade inata, mas pode ser colocado ao lado do behaviorismo uma

vez que também preconiza a associação de informações como parte do processo de aquisição. Contudo, ao mesmo tempo em que compartilha de alguns pressupostos behavioristas é importante ressaltar que, para as abordagens behavioristas, um comportamento é aprendido em função da sua repetição e do reforço positivo que o aprendiz recebe. O aprendizado para um modelo conexionista, no entanto, se dá através das conexões estabelecidas e da força entre elas nas redes neurais, da generalização de um padrão que se obtém pela observação do funcionamento de um número reduzido de dados/fatos.

Os mecanismos neurais são os responsáveis por gerar o conhecimento não só linguístico, mas de outros processos que envolvem a interação humana. Ao contrário de pressuporem regras universais, como os inatistas, os conexionistas assumem que os aprendizes podem ter representações únicas a depender de uma série de fatores que estão associados à L1 do aprendiz, à natureza do *input* recebido, dentre outros fatores. Em suma, os modelos conexionistas levam em conta fatores variados que atuam no processo de aquisição de linguagem, diferentemente dos modelos inatistas. No entanto, a literatura de língua estrangeira é controversa no que concerne a quais fatores especificamente se deve considerar, por exemplo, como o aprendiz teve contato com a língua-alvo, como se deu o aprendizado, quais foram as condições em que o aprendiz adquiriu a língua, as diferentes metodologias de ensino pelas quais passou durante o aprendizado formal da língua, a idade em que aquisição da L2 teve início, etc.

A partir do caráter associativo entre as informações linguísticas e dados da experiência do indivíduo (como foram mencionados no parágrafo anterior), a mente humana procura criar ligações, formar elos que se tornam mais fortes a depender da recorrência com que as formas são apresentadas e recebidas pelo aprendiz. Com relação ao funcionamento do modelo, é possível estabelecer que este prevê que as redes neurais funcionam através de um sistema de força de conexões no interior das redes neurais ou, melhor, o modelo trabalha com um “ajustamento da força da conexão” (RUMELHART,

HINTON E WILLIAMS, 1986). Nesse sentido, por exemplo, existem certos aspectos gramaticais da L2 que terão conexões mais fortes com a L1 e outras conexões mais fracas. De acordo com Poersch (2004:451), “as redes aprendem, alterando a força das conexões como resposta à atividade neural” e “a repetição de experiências de aprendizagem ocasiona um incremento na força das conexões”.

Tomando alguns dados de aquisição de L1 a título de exemplificação do modelo conexionista, é possível citar o trabalho de Rumelhart *et al.* (1986) que estudaram o aprendizado do passado simples do inglês por crianças. Os autores notam, ao longo do estudo, que a aquisição do passado ocorre em três estágios. Existe um primeiro estágio no qual os aprendizes utilizam alguns verbos no passado, em sua maioria os regulares, e, ao usarem tais formas, as produções são corretas. No segundo estágio, as crianças-aprendizes passam a necessitar e utilizar mais verbos, agora, em sua maioria, irregulares, o que faz com que uma regra para o passado seja criada. A regra de inserir a terminação *-ed* para marcar o passado é utilizada pelos aprendizes em diversos contextos, inclusive nas produções de verbos irregulares, os quais, em um primeiro momento, eram produzidos corretamente. Finalmente, num terceiro estágio, as crianças já conhecem ambos os padrões, os verbos regulares e os irregulares e ambos coexistem no sistema da L1 do indivíduo. Nesse sentido, percebe-se que o processo de aquisição de um determinado aspecto gramatical ocorre através de etapas, as quais vão sendo ultrapassadas a depender dos padrões que vão sendo criados pelo aprendiz.

O que foi discutido acima é aplicado para a aquisição de L1, mas como a noção de linguagem e aprendizado de língua se adequaria para a L2? Apesar do processo associativo se manter para aquisição de L2, de acordo com Ellis (2007:8), as redes neurais do indivíduo já possuem a configuração dos padrões que operam na L1 “chegando a um ponto de fixação em que a L2 é percebida através dos mecanismos otimizados para a L1” (tradução minha)¹¹. Nesse

¹¹ “... it has reached a point of entrenchment where the L2 is perceived through mechanisms optimized for the L1” (Ellis, 2007:8).

sentido, pode-se dizer que as forças de conexão que operam nas redes neurais funcionam em um sistema de atração e repulsão entre as formas existentes na L1 e que podem ser ou não encontradas na L2. Algumas dos padrões para aquisição de determinados aspectos da L1 são transferidos para o aprendizado da L2. As redes neurais não são estáticas e, assim, elas se modificam a partir das novas conexões que são realizadas, fazendo com que o indivíduo continue assimilando novas informações mesmo já possuindo outro sistema (L1). No entanto, conforme Ferreira-Junior (2007), as redes possuem graus de sensibilidade distintos a depender dos estágios de aprendizado, de modo que em estágios iniciais elas são mais sensíveis às mudanças. O autor continua:

(...) à medida que o treinamento avança, as ativações das unidades são alargadas (em termos de variação e tolerância). Ou seja, vão do centro para os extremos (em escala). Isso diminui a velocidade da aprendizagem e, em última instância, gera o que chamamos em ASL de “fossilização” (ou seja, exemplos positivos não mais afetam a interlíngua ou o sistema). (FERREIRA-JUNIOR, 2007:229)

A “fossilização” ocorre com mais frequência para aprendizes adultos do que para aprendizes crianças de uma L2, e, como é mencionado acima, advém da concepção de que o cérebro dos adultos é menos plástico (SOKOLIK, 1990) do que o das crianças e, em decorrência disso, as conexões feitas pelas redes neurais são distintas. A partir disso, é possível dizer que a rede neural passa por um treinamento e, após esse processo, alterações na rede podem ser mais difíceis de serem feitas. Com relação a essa informação, o modelo procura explicar que, ao adquirir a L2 enquanto adulto, um indivíduo já possui sua rede de conexões da L1 treinada e, portanto, as conexões que são adquiridas posteriormente podem, em sua maioria, não ser incorporadas pela rede, resultando por vezes nos “erros” de produção na L2¹².

¹² Para uma leitura mais aprofundada em experimentos que testaram as restrições às regras estabelecidas por falantes-aprendizes de uma L2, verificar: Sokolik e Smith (1992), Ellis e Schmidt (1997), Sokolik e Smith (1992), Haskell, Macdonald e Seidenberg (2003).

Até este ponto, nas três propostas de aquisição de L2 analisadas (behaviorista, gerativista e conexionista), o “erro” foi sendo incorporado em um primeiro momento de modo a ser considerado como um desvio, para depois ser visto como parte do processo de aprendizagem, estando relacionado com a capacidade do aprendiz de estabelecer novas categorias para um sistema de L1 já formado, chegando, posteriormente, a ser um resultado parcial de um cálculo gerado a partir de conexões realizadas por redes neurais. É importante notar que a noção de “erro” recebe visões distintas a partir dos aspectos que são considerados parte do processo de aquisição de uma língua. As restrições das abordagens e modelos moldam o entendimento das produções com desvios.

Ainda que existam várias questões que as abordagens e modelos descritos aqui não consigam explicar, um passo muito importante foi dado ao se perceber que o aprendizado de uma L2 envolve a aquisição de aspectos distintos daqueles da L1 e que as línguas e o sistema cognitivo parecem funcionar de maneira integrada.

As pesquisas conexionistas recentes já demonstram uma preocupação com o fato de que os sistemas fonético-fonológico, sintático e semântico possuem uma integração e que essa relação deve ser dinâmica, uma vez que as mudanças ocorridas em um deles influenciam o outro (ZIMMER, 2010). Os sistemas dinâmicos, advindos da matemática, trabalham com a noção de que todos os fatores que contribuem para a formação de um sistema estão integrados e são necessários para o seu funcionamento. Nenhum elemento de um grupo passa por uma mudança sem que esta dependa do estado em que se encontram os demais elementos do mesmo grupo (BLANK & ZIMMER, 2011:30).

Como foi dito acima, os componentes de um sistema interagem de forma a estabelecer uma relação de dependência. As mudanças que incidem sobre um sistema podem ocorrer em qualquer momento ou estágio no qual tais elementos estejam. Se pensarmos na produção de fala, quando um indivíduo produz um enunciado como “bola”, ele não o faz pensando em cada som

isoladamente, para que o ouvinte após capturar todas as sequências individuais as agrupe, formando então um conjunto que seja reconhecido como tal palavra. A cadeia sonora para o exemplo exposto se dá no tempo, a articulação das vogais e consoantes é construída no *continuum* espaço-temporal. Até então, paradigmas teóricos como o gerativismo inseriam o tempo apenas na produção, ou seja, na realização física, como no exemplo descrito acima. Contudo, na abstração, na representação do dado, o tempo não era incorporado. Assim, para modelos de base gerativista, conforme discorre Silva (2009:2) “[...] decorre a posição de que cada segmento ocupa um *slot*¹³ na representação de um enunciado. Se é assim, não é possível haver sobreposição entre os segmentos num plano simbólico”. Uma vez que não há a possibilidade de sobreposição, em modelos gerativistas – apesar de que, como dissemos antes, o “erro” passa a receber uma explicação dentro da generalização feita pelo aprendiz – como são representadas as produções dos aprendizes de L2, as quais muitas vezes se encontram sobrepostas e ocasionam os chamados desvios de produção? A partir disso, é possível dizer que os sistemas dinâmicos vão além. Para estes o tempo é intrínseco, fazendo parte da representação linguística do indivíduo, ele não se encontra somente na realização física dos sons. Assim, ao incorporarem a unidade tempo, os sistemas dinâmicos passam a dar conta de maneira mais parcimoniosa, por exemplo, dos “desvios” de produção, em especial, para a L2, por não verem no “erro” uma não produção ou uma falha, mas sim, um degrau alcançado nas generalizações do falante-aprendiz. Dentro do campo da produção de fala, especificamente, modelos como a Fonologia Gestual de Browman e Goldstein (1986, 1989, 1990, 1992) trabalham sob uma perspectiva dinâmica. O modelo desses autores será visto com mais profundidade nos capítulos seguintes.

Outra característica dos sistemas dinâmicos se encontra no fato de que eles podem ser lineares ou não. Embora os modelos de base gerativa podem não ser lineares, esses preveem um processamento serial da linguagem. O modelo conexionista preconiza um processamento em paralelo e incremental,

¹³ Uma janela.

ou seja, não se processam todos os dados para depois implementar a fala; ao contrário, processa-se uma parte do enunciado que é implementado na sequência, e assim sucessivamente até que todo o enunciado tenha sido processado e implementado.

Embora os sistemas conexionistas e dinâmicos compartilhem de vários pressupostos de base, a diferença entre esses tipos de modelo basicamente consiste no fato de que os modelos conexionistas trabalham com a noção de que existem pesos ligando as diferentes camadas que arquitetam um sistema. Já os modelos dinâmicos, apesar de preverem, como os conexionistas, que há uma interação de diversas variáveis determinando a produção de fala, não preveem que um sistema seja organizado em camadas e que haja pesos ligando as variáveis de camadas distintas.

Voltando-se para questões de aquisição de L1 e L2, é possível que, como propõem Blank e Zimmer (*op. cit.*), tome-se o sistema linguístico dos aprendizes como dinâmico e que, portanto, todas as informações adquiridas pelo indivíduo desde a infância, sejam elas linguísticas ou não, vão sendo armazenadas e se encontram em um estado de interação constante. Da mesma forma que os sistemas fonético-fonológico, sintático e semântico foram muitas vezes descritos por outros paradigmas teóricos como isolados, mas agora passam a ser vistos numa relação de interdependência, de modo semelhante, o aprendiz utiliza o conhecimento da L1 para assimilar o novo conteúdo da L2-alvo, lançando mão da integração de recursos linguísticos e de experiência, aqui no sentido biológico. Como os sistemas são constituídos de inúmeras variáveis que se relacionam (De BOT, VERSPOOR, LOWIE *et al.*, 2005), a aquisição deve ser vista como um processo complexo que reúne tais variáveis somadas aos fatores que podem influenciar os processos individuais, como as condições iniciais de aprendizado da L2, os métodos utilizados no aprendizado, etc. Com isso, é interessante notar que os aprendizes possuem sistemas linguísticos distintos e, portanto, podem ter ritmos de aprendizado diferentes ou, ainda, precisam de mais tempo para adquirir uma determinada língua.

Cada um dos modelos citados contribuiu para as pesquisas tanto em aquisição de L1 como em L2. Embora em alguns paradigmas teóricos os pressupostos de base sejam radicalmente distintos, em outros, alguns pressupostos acabam convergindo, fazendo com que novos fenômenos linguísticos passem a receber atenção e sejam investigados. Nesse sentido, o termo “erro” foi sempre tratado entre aspas por não possuir uma única definição. Esse conceito surgiu, em um primeiro momento, para denunciar uma falha no processo de aquisição de linguagem, seja porque o aprendiz não tenha recebido reforço positivo suficiente, ou por não ter sido capaz de criar generalizações que dessem conta de construir estruturas adequadas para a língua-alvo. Mas essa palavra recebe outro olhar a partir do momento em que informações, como a experiência do aprendiz, são incorporadas pelos modelos. Com isso, pode-se dizer em última instância que, para os modelos dinâmicos e conexionistas, o erro não existe. O falante passa por um constante processo de maturação linguística, cognitiva e operacional (com relação aos seus outros sistemas em funcionamento, como o respiratório, por exemplo). Nesse sentido, algo que seja escutado como “erro” pode revelar informações sobre o que a literatura chama de contrastes encobertos¹⁴. Essa concepção de “erro” será utilizada nos capítulos a seguir para tratar de um fenômeno linguístico, a dessonorização terminal, foco deste trabalho.

¹⁴ Por contrastes encobertos entende-se uma diferença subfonêmica, que pode não ser percebida de outiva, mas que pode ser medida acusticamente. Pode-se dizer que, ao lidar com processos de aquisição, seja de língua materna ou estrangeira, os contrastes encobertos se tornam a regra e não a exceção (Munson, Edwards, Schellinger, Beckman e Meyer (2010). Para uma leitura mais aprofundada, ver Macken e Barton (1980), os quais desenvolvem um experimento com a aquisição do contraste sonoro em oclusivas por crianças.

3 DESSONORIZAÇÃO TERMINAL (DT): REVISÃO BIBLIOGRÁFICA DE EXPERIMENTOS DE PRODUÇÃO DE BASE GERATIVA

Neste capítulo, objetiva-se discorrer sobre os estudos que tiveram como foco a dessonorização terminal (DT). Em linhas gerais, antes de se passar à resenha desses trabalhos, é importante ressaltar o aparecimento da DT nos estudos de aquisição de L1 e L2.

A literatura fonética de aquisição de L1 trouxe importantes considerações sobre a realização de consoantes em posição inicial e final de palavra. Há o relato de que, em posição final, todas as consoantes oclusivas são realizadas como surdas, a exemplo de Ewert (1933), que descreveu os sons do francês antigo; Dinnsen e Eckman (1978), que pesquisaram um dialeto da região de Ferrara, na Itália; Rubach (1984), para o polonês; e Padgett (2002), para os dados do Russo.

Achados como os descritos acima motivaram outras pesquisas, uma vez que um questionamento permanecia: se há uma diferença fonêmica, os falantes nativos não deveriam produzir alguma distinção de sonoridade entre os segmentos? Além disso, experimentos com bebês ainda na fase do balbucio mostraram que eles procuram produzir as oclusivas finais surdas e sonoras (OLLER *et al.* 1976). Mas, conforme menciona Myers (2011), apesar da tentativa dos infantes de realizar a distinção sonora das oclusivas em posição final, eles produzem apenas as surdas¹⁵.

Esse processo de dessonorização no fim de sequências sonoras foi sendo visto como um erro, como uma produção desviante. Isso ficou ainda mais marcado na área de aquisição de L2, que apresentou dados de realizações de falantes-aprendizes tidas como inadequadas quando a consoante-alvo é vozeada e a sua produção é desvozeada. Tal fato ocorre,

¹⁵ Para um estudo aprofundado no processo de aquisição de oclusivas surdas e sonoras, verificar Velten (1943), Smith (1973 e 1979), Flege (1982).

segundo Myers (*op cit.*), independentemente de existir dessonorização na L2-alvo ou na L1 do falante. Esse fenômeno fonológico foi assim descrito a partir de algumas pistas fonéticas, como a “fonética da posição pré-pausa”, mencionada por Meyers (*op cit.*), segundo quem existe a possibilidade de que a pausa influencie o desvozeamento da consoante (como em SIEVERS, 1901; JESPERSEN, 1926; BLOOMFIELD, 1933; LINDBLOM, 1983), através de um processo de assimilação do silêncio da pausa pela consoante vozeada adjacente.

Como foi descrito acima, existe uma visão distinta entre a DT como fenômeno fonológico e as pistas acústicas que, se por um lado contribuem para a discussão sobre a DT, uma vez que inserem exemplos que são verificados em diversas línguas, por outro, levantam uma discussão sobre a representação e realização da DT.

Nesse sentido, será feito um levantamento de alguns estudos de produção e percepção, para que se discutam as contribuições de diferentes modelos para esse fenômeno. Com relação aos estudos de produção, serão mencionados os aspectos acústicos que estão relacionados à DT, para que se inicie uma discussão sobre quais pistas acústicas são relevantes na produção de oclusivas finais. Não serão mencionados neste momento, ainda, modelos de análise perceptual de L2 (tal discussão será feita no capítulo 5).

3.1 ESTUDOS DA DÉCADA DE 1980

Vários estudos discorreram ao longo dos anos sobre a dificuldade na distinção entre plosivas sonoras e surdas no nível da representação e sobre a existência de um processo de neutralização da oposição surdo-sonora em posição final por parte dos aprendizes de uma L2.

Nessa direção, temos os primeiros estudos realizados sobre a DT em aquisição de L2: tal fenômeno não é recente e vem sendo objeto de estudo desde a década de 1980, com o trabalho de Eckman (1981). Nesse estudo e em um posterior (1987), o autor caracteriza a DT como sendo um processo que envolve a perda do traço [sonoro] em algumas posições finais e que ocorre em

língua materna e estrangeira. Especificamente, no estudo de 1981, Eckman realizou um experimento com falantes de espanhol e chinês como L1 que estavam aprendendo inglês. Ao testar a produção de oclusivas finais do inglês com ambos os grupos, o autor notou que os hispanos dessonorizavam as oclusivas finais e os chineses inseriam um segmento semelhante a uma vogal. De acordo com Eckman, era previsto que os aprendizes, a partir da fonologia da L1, encontrassem as saídas descritas acima. O autor menciona que existe uma tendência a dessonorizar em L2 e isso se deve ao que ele descreve como noção de dificuldade, a qual está relacionada ao que as línguas marcam, ou selecionam como recorrente. Segundo ele, as línguas funcionam em um sistema de existência de contrastes em posições que apresentam uma menor dificuldade de ocorrência. Nesse sentido, um contraste como o de vozeamento, por exemplo, teria um alto grau de dificuldade de ser encontrado em posição de fim de palavra, por se tratar de um contexto em que diversos fenômenos ocorrem, a saber: apagamento de segmentos, neutralização, etc. No entanto, se uma dada língua exibe o contraste de vozeamento em posição de final de palavra, ela também possui tal distinção de sonoridade nas posições mediana e inicial, por serem posições menos “difíceis”. Para uma língua como o alemão, por exemplo, que não oferece contraste nas oclusivas em posição final, mas sim na posição mediana, isso faz com que o contraste exista naturalmente no início; tal assunção equivale à premissa de Pike (1948) relativa à tendência dos sistemas à simetria. O modelo com o qual Eckman (1987) trabalha propõe um sistema de restrições. Ao contrário dos pressupostos iniciais do modelo gerativo, no qual as manifestações linguísticas são expressas através da aplicação de regras, as generalizações, para Eckman, são feitas a partir das restrições que operam sobre cada língua. Há, portanto, como se mencionou no parágrafo acima, uma restrição que opera em todas as línguas e faz com que o contraste sonoro em final de palavra não seja impossível, mas dificilmente encontrado. No entanto, é importante mencionar que embora a restrição em

questão esteja presente em todas as línguas, seus efeitos só são sentidos em algumas, dependendo do *ranking*¹⁶.

O autor menciona que algumas restrições¹⁷ que operam na L2 podem ser as mesmas ou não das operantes na L1. Com isso, o aluno aprenderia a lidar com a presença do contraste em posição de início e meio de palavra, existente em sua L1, e generalizaria a sua distinção de sonoridade em posição de fim de palavra para a L2.

Outros autores, como Major (1987), também descreveram a DT como se tratando da perda do traço [sonoro], mas para este autor o processo da DT se dá na interlíngua. Como já havíamos mencionado, no capítulo anterior, o que estaria ocorrendo com a produção dos aprendizes ao tentarem produzir as oclusivas finais é uma tentativa de sistematização entre a L1 e a L2. Nesse caso, em uma língua em que não existem oclusivas em posição final, como é o caso do português brasileiro (PB), o aprendiz teria em sua interlíngua um segmento oclusivo que seria neutralizado, ou seja, não faria a distinção dos pares surdos e sonoros ou, até mesmo, o aprendiz não produziria a oclusiva final, uma vez que esta não pertence ao sistema da L1. Associados a um paradigma inatista, os trabalhos de Eckman e Major têm como premissa que a perda do traço [sonoro] deve ocorrer para a produção de oclusivas em todas as línguas. Esses trabalhos se debruçam sobre a representação da DT, tratando-a

¹⁶ O termo *ranking* referido nesse momento diz respeito a um conceito exposto nos primeiros textos que tratam da OT, Prince e Smolensky (1993), tal termo não existia nos trabalhos da década de 1980. Todas as línguas possuem o mesmo conjunto de restrições, que pertencem à gramática Universal. A diferença entre as línguas estaria, então, em como essas restrições são ranqueadas. Por exemplo, na existência de quaisquer restrições A e B, ou A domina B ou o oposto. O *ranking* das restrições para qualquer língua consiste no ordenamento das restrições de acordo com a relação de dominação exemplificada acima.

¹⁷ É importante ressaltar que o termo *constraints*, restrições, usado por Eckman tem um sentido diferente dentro de modelos como a OT. Ao invés de regras, como as existentes na fonologia gerativa até então, passa-se a trabalhar com as restrições. Estas determinam qual *output* é o mais harmônico com a gramática. Um output é tido como harmônico quando esse satisfaz as restrições e o ranqueamento delas.

como um processo de neutralização¹⁸ da sonoridade das oclusivas em final de palavra. Entretanto, dados de produção de fala, através da investigação de algumas pistas acústicas, apontam na direção oposta dessa previsão, uma vez que evidenciam que a neutralização não é total.

Um aspecto que deve ser considerado nos estudos resenhados é que os dados foram obtidos através de uma análise de ouviva, através da qual o pesquisador lança mão de seu aparato auditivo e de sua impressão acústica, sobre os dados de produção, podendo, por uma limitação do ouvido humano, não perceber informações que podem se mostrar relevantes na distinção de algumas sequências sonoras. Contudo, ainda que a análise tenha sido assim conduzida, os estudos de Eckman e Major abriram caminho para a necessidade de se conduzirem mais estudos sobre a DT, especialmente em outras línguas.

Com relação a dados de aquisição de L1, autores como Ernestus e Baayen (2006) verificaram que o fenômeno da DT se dá de maneira incompleta no holandês, pois algumas distinções se mantêm, como entre [t] e [d]. Para o alemão, Piroth e Janker (2004) discorrem sobre a existência de uma neutralização total para alguns parâmetros acústicos, como duração da vogal precedente e vozeamento na oclusão, mas não em outros, como duração da plosiva. Conforme mencionado por Oostendorp (2007:1):

1. Alemão:

i. bli[n][t] 'blind(PRED.)'- blind[ə] 'angry(ATT)'

ii. bunt 'colourful(PRED.)'- bunt[ə] 'colourful(ATT)'

O que podemos verificar acima é que em ambos os exemplos, temos as primeiras palavras terminando com a plosiva alveolar surda, no entanto, quando temos a presença de contexto vocálico adjacente em (i) temos a mudança da surda para sua contraparte sonora. Contudo, em (ii) temos o

¹⁸ Entende-se neutralização como sendo a perda de oposição entre unidades distintivas do sistema em função, por exemplo, de acento ou de posição de tal unidade na palavra. Essas unidades podem ser fonemas ou traços, a depender do paradigma teórico que se adote.

mesmo contexto adjacente e não temos a mudança da consoante surda em sonora, como é o caso do par mínimo “Bund” [t], que significa “povo” e “bundes”, “do povo”. Segundo Oostendorp, para línguas que apresentam casos como os do alemão em (1), a literatura acaba por explicar que o desvozeamento final é foneticamente incompleto: “[...] a neutralização não é completamente perdida, mas é passível de ser recuperada através do detalhe fonético fino” (OOSTENDORP, 2007:2, tradução minha)¹⁹, nesse caso a duração da vogal que precede a consoante final. De acordo com o autor, esse tipo de visão sobre o fenômeno faz com que se assumam que algumas teorias formalistas não são capazes de oferecer uma análise que dê conta da neutralização. Oostendorp continua a argumentação pontuando que trabalhos como os de Port e Leary (2005)²⁰ acabam por atribuir a não adequação de modelos formalistas para tratar fenômenos como a DT ao modelo proposto por Chomsky e Halle (1968), cujas representações, segundo Oostendorp, são demasiadamente simples.

Como Oostendorp (*op cit.*) aponta, a questão da neutralização incompleta e do tratamento que se dá a esse achado estão conectados a pressupostos teóricos distintos, ou seja, de se considerar que, por um lado, pode se tratar de uma visão, segundo a qual a fonética consegue ter acesso à estrutura profunda, à estrutura para uma dada palavra no sistema da língua e, nesse caso ela “não é cega” (OOSTENDORP, 2007:4). Por outro lado, é possível pensar que a fonética não possui tal independência e que “a fonologia já se encarrega do ‘realce’ que seria, de outra forma, considerado puramente fonético²¹” (Oostendorp, 2007:4). Trabalhos como o de Avery e Rice (1989) mostram algumas evidências de como o contraste sonoro e surdo nas

¹⁹ Trecho original: “[...] the neutralization is not completely lost, but recoverable in fine-grained phonetic detail”.

²⁰ Port e Leary (2005) apontam que o dado fonético ilumina a questão da DT, uma vez que alguns falantes produzem uma distinção entre dois sons de modo que esta seja parcialmente neutralizada.

²¹ Trecho original: “[...] the phonology already takes care of the ‘enhancement’ that would otherwise be considered purely phonetic.”

oclusivas pode ser verificado através da duração da vogal que precede a oclusiva, a fonologia “informa” qual deve ser a duração da vogal em função da natureza da consoante que a segue, e essa diferença de duração é implementada na sequência pela fonética.

A partir dessa constatação, será feito, abaixo, um apanhado dos argumentos principais da proposta de Oostendorp.

3.2 A PROPOSTA DE OOSTENDORP DENTRO DA OT: (O QUE FAZER COM AS EVIDÊNCIAS DO DADO FONÉTICO?)

Como mencionado acima, o texto de Oostendorp surge no interior das discussões entre o tratamento dado aos fenômenos como a DT por teorias discretas e outras de base dinâmica. Como já foi mencionado anteriormente, Oostendorp (2007) abre seu texto remetendo a uma pesquisa de Port e Leary (2005), a qual se propõe a discutir o tratamento dado por teorias formais fonológicas a certos fenômenos linguísticos²². Nesse texto, os autores, ao mencionarem o tratamento dado à dessonorização terminal, expõem que os relatos de neutralização incompleta apresentados pela literatura da área, como a neutralização final no Russo (PYE, 1986) no Polonês (SLOWIACZEK & DINNSEN, 1985) e no Alemão (WARNER, JONGMAN, SERENO, KEMPS, 2004) trazem dados de que as produções coletadas são sensíveis ao contexto, ou seja, a depender do contexto adjacente, a oclusiva em posição final pode ser sonora, surda ou até mesmo parcialmente sonora. A questão para Port e Leary (2005) parece ser, na verdade, a de que as categorias assumidas por modelos teóricos como de Chomsky e Halle (1968) não dão conta de explicar casos como o da neutralização incompleta, pois tais modelos não trabalham com as variações que são observadas entre as categorias, uma vez que essas não estão no *input*. Os autores chegam a comparar a representação abstrata apresentada por modelos fonológicos formais, como o de Chomsky e Halle (*op*

²² É mister esclarecer que as teorias formais criticadas por Port e Leary (2005) são as de base gerativa.

cit.), com a discretude do sistema ortográfico, alfabético da língua, o qual se encontra construído sobre uma abordagem categórica e limitadora dos sons da língua, uma vez que se propõe a representar apenas as propriedades da fala cujos contrastes se mostram mais relevantes para uma distinção entre sons de categorias distintas e acaba por deixar alguns fenômenos, como a DT, sem um tratamento adequado. Contudo, Oostendorp procura apontar, em seu estudo, que as teorias formais fonológicas passaram por mudanças e não se pode mais trabalhar com noções que foram estabelecidas num primeiro momento do modelo de fonologia gerativa.

Assim, o autor se encontra em um segundo momento da revisão da DT a partir da OT. Como foi visto na seção anterior, Eckman também se baseia em um sistema de restrições para descrever a DT, mas a questão, no momento, para Oostendorp, é distinta. Ao mesmo tempo em que o autor não pode ignorar as pistas oferecidas pelos dados de neutralização incompleta em línguas como o alemão, holandês, russo e polonês, ele tem que encaixar essas evidências dentro do modelo com o qual trabalha. Mas, para isso, como veremos abaixo, o autor terá de lançar mão de recursos que não são inicial e originalmente previstos no arcabouço da OT, como a *Turbidity Theory* (doravante Teoria da Turbidez) A contribuição da Teoria da Turbidez será tratada abaixo.

A OT trabalha com um sistema de restrições estabelecidas pela GU, operando nas línguas de modo distinto. Tais restrições podem ser as de marcação (*markedness*) e as de fidelidade (*faithfulness*): as primeiras acabam por realizar uma avaliação das representações do *output*, como por exemplo, se um determinado segmento exibe ou não traços que sejam incompatíveis entre si, como por exemplo, o traço [+anterior] e [+vocálico], admitindo-se, ainda como em SPE, que todo traço [+voc] implica, necessariamente, a especificação negativa para [ant]; as segundas verificam a semelhança entre *input* e *output*, *i.e.*, elas operam entre um *input* e um possível *output*. Elas exigem que os segmentos correspondentes sejam idênticos em algumas propriedades específicas.

Com relação às Restrições de Fidelidade, a OT apresenta duas vertentes teóricas. Por um lado, a *Correspondence Theory* explicita que as representações criadas pelo *input* e *output* são separadas assim como a relação entre esses dois sistemas (McCARTHY e PRINCE, 1995). Na *Containment Theory* (PRINCE e SMOLENSKY, 1993), o *input* se encontra dentro do *output*, ou seja, tudo que se manifesta no *output*, possui uma representação no *input*. De acordo com um exemplo dado por Oostendorp, em uma sequência na qual se tem algo como /capt/, na representação, no *input*, na realização pode se ter [cat]. Esse segmento produzido passa por um processo de parseamento, no qual existe a transposição de um segmento dentro de uma estrutura silábica. Nesse caso, não é possível que se fale em um segmento que tenha sido introduzido pela fonética e que não tenha uma representação no *input*. Contudo, o modelo, como o autor mesmo aponta, apresenta algumas questões que precisariam ser investigadas, como o caso da epêntese. De acordo com a operação de parseamento e preenchimento, um segmento inserido deve ser “vazio”. Nesse caso, se fosse permitida a inserção de segmentos epentéticos que não fossem “vazios”, então, não se teria como distinguir epênteses que acontecem de modo não previsível nas línguas, daquelas que estariam previstas no sistema de uma determinada língua-alvo.

Além dos processos sobre a inserção de segmentos, se torna importante discutir o que acontece com segmentos que sofrem apagamento. De acordo com o parseamento, de “transformação” do conteúdo abstrato (*input*) em realização física (*output*), como menciona Oostendorp (2007:6), “elementos apagados não ‘são parseados’ na estrutura fonológica”. Quando isso ocorre, o segmento deixa para trás a sua mora²³, a qual pode ser preenchida por outro segmento. Mas, que tipo de item ocupará esse lugar?

²³ De acordo com Ladefoged (1962) a mora é uma unidade equivalente a uma sílaba curta. Segundo o autor, ela é uma unidade de tempo. Trata-se de uma unidade da fonologia que determina o peso silábico em algumas línguas. Línguas como o japonês organizam seu ritmo em função da duração das moras que constituem a cadeia sonora. Oostendorp parece propor que as moras funcionam como “slots temporais” que precisam ser preenchidos. Assim, as moras constituem parte da informação fonológica da língua e cabe ao nível fonético preencher essas moras com sons. Se um som é apagado, o slot permanece e precisa ser preenchido.

Tal questão nos coloca diante da necessidade de discutir como as restrições são ordenadas no *ranking* de modo a prever a realização ou apagamento de segmentos representados no *input*. Tomando como base as restrições de fidelidade, como se dá o ranqueamento de um item que foi apagado ou que não foi pronunciado, qual seria, então, a representação no *input*? Essa discussão se mostra importante para se tentar dar conta da DT numa abordagem afiliada a esse paradigma teórico.

A partir disso, Oostendorp reconhece as limitações das duas linhas teóricas apresentadas para tratar de fenômenos que envolvem a inserção ou apagamento de um dado segmento. Nesse sentido, o autor (*op cit.*: 8) prevê um encaminhamento para esses e mesmo dados semelhantes através de outro modelo. Assim, autores como Goldrick (2000) lançam as primeiras ideias sobre a Teoria da Turbidez. Nessa teoria, dois tipos de relação operam: de projeção e de pronúncia. A primeira diz respeito a um processo estrutural do *input* que se dá entre o segmento e o traço, enquanto que a pronúncia se trata da contraparte fonética, da realização da estrutura. Oostendorp demonstra como se daria a explicação do parseamento de um dado que no *input* teria duas vogais e cuja vogal passa a ser longa no /ka+ezi/ → [ke:zi] ‘moon (DIM)²⁴, Oostendorp (*op cit.*: 9)”. Segundo o autor, uma análise mais tradicional diria que a vogal acaba por deixar sua mora, a qual é ocupada por outra vogal, fazendo com que o novo segmento se torne bimoraico e, portanto, longo. De acordo com os pressupostos da OT, as vogais não precisam necessariamente ter mora no *input*, mas, independentemente desse fato, o efeito causado por ela, a mora, permanece o mesmo. Na Teoria da Turbidez, “a mora é projetada pela vogal apagada, apesar do fato de essa vogal não ser pronunciada;

²³ DIM = diminutivo.

contudo, a mora é pronunciada na vogal que é de fato pronunciada²⁵” (Oostendorp, *op cit.*: 9, tradução minha).

Diante do que foi exposto acima, podemos pensar um pouco sobre como esse modelo descrito por Oostendorp enxerga a DT. De acordo com a Teoria da Turbidez, um determinado segmento projetado, como foi visto no exemplo do parágrafo anterior, pode ser pronunciado ou não, ou seja, existe uma possibilidade de representação para os segmentos que não são encontrados no *output*. De acordo com o autor, as representações para o *output* da DT seriam [t], [d] ou [d_o]. Um segmento que não possui vozeamento terá no *input* representações distintas quando for surdo ou quando se mostrar desvozeado.

O autor procura, então, através da Teoria da Turbidez, incorporar os achados de línguas como o alemão e o holandês, que previam a presença de algumas características fonéticas como a duração da vogal, a qual é tida como uma pista da sonoridade da consoante que foi dessonorizada. De acordo com Oostendorp (2007:12, tradução minha): “se a duração da vogal precedente é uma pista para o vozeamento, esse alongamento pode ser levemente iniciado, da mesma forma, por traços de vozeamento projetados²⁶”. Para o autor, então, um dado que seja vozeado no *input*, tem duas formas de ser produzido. Ele pode ter um traço que é projetado e, também, pronunciado ou, por outro lado, não.

No entanto, ao mesmo tempo em que o dado fonético é tido pelo autor como indiscutivelmente gradiente, ele procura manter a distinção clássica entre fonética e fonologia. Apesar da proposta exposta por Oostendorp (2007) dar continuidade aos trabalhos que investigam a DT a partir de modelos fonológicos, e de oferecer contribuições importantes para o tratamento desse e de outros fenômenos, a saída para dar conta da DT, como já foi inicialmente exposta, se deu a partir da inserção de outros mecanismos, como a projeção e

²⁵ No trecho original: “[...] is that the mora is projected by the deleted vowel, in spite of the fact that this deleted vowel is itself not pronounced; however the mora is pronounced on the vowel which is itself pronounced.”

²⁶ No trecho original: “[...] if the length of the preceding vowel is a cue to voicing, this lengthening may be triggered slightly by projected voicing features as well”.

a pronúncia, pressupostos que não se encontram nas descrições iniciais da OT.

A Teoria da Turbidez propõe maneiras de se tentar representar fatos que não são binários, mas tripartites, por exemplo, como é o caso da sonoridade, tal qual concebida por Oostendorp (2007). O estudo do autor reconhece que é preciso um maior detalhamento na representação fonológica e apresenta uma forma de se representar esse detalhamento. Assim, tal análise é capaz de explicar a DT de modo mais adequado quando comparada aos estudos da década de 1980. No entanto, é necessário que se olhe para fatos empíricos que reforçam a necessidade de incorporar o detalhe fonético fino à representação fonológica de modo que se analise a interação entre outros aspectos (como o vozeamento e o *burst*²⁷) que contribuem para a variação nos dados de produção da DT. Para isso, além da contribuição dada pela OT, passa-se a olhar para outras teorias que também se debruçaram sobre o fenômeno da DT, com o objetivo de fornecer uma análise dinâmica para ela.

²⁷ O *burst* é a "explosão" de ar resultante da soltura dos articuladores, durante a realização de uma oclusiva.

4 A DT SOB UM PONTO DE VISTA DINÂMICO

A partir do cenário exposto até este ponto, percebe-se que cada modelo/abordagem procura engendrar as evidências fonéticas a partir de regras fonológicas. No entanto, fatos como a neutralização incompleta parecem apontar para a necessidade de que, para contemplar fatos que envolvem mudanças temporais, ou seja, de natureza gradiente, é necessário prever primitivos de análise que incorporem a variável tempo. Tal fato vai de encontro a modelos inspirados na Fonologia Gerativa de Chomsky e Halle (1968), que consideram o tempo um produto da implementação – portanto extrínseco aos primitivos de análise. Um dos pesquisadores a tentar trabalhar com o tempo como fator intrínseco foi Gafos (2002). O autor procura associar a Fonologia Gestual à OT, propondo que os *rankings* de ordenamento dos padrões de organização dos gestos possibilitaria a incorporação da gradiência às gramáticas das línguas. Para se dar conta de dados que envolvem a ativação de fatores linguísticos e biológicos e que lidam com mudanças que se dão num determinado tempo e espaço é preciso que o tempo seja considerado intrínseco. Teorias que possuem como primitivos de análise unidades categóricas, como o traço (Chomsky e Halle, 1968) – o qual é geralmente utilizado em algumas análises via OT – não incorporam o tempo.

Um dos primeiros trabalhos a se valer de fatos empíricos para chamar a atenção para a necessidade de incorporar a variável tempo aos modelos de análise fonológica foi o de Keating (1985). Nesse, a autora faz uma reflexão crítica sobre o gerativismo, paradigma teórico com o qual trabalha, a fim de discutir alguns fatos que são tomados como universais, mas que são, na verdade, específicos de língua. A autora, ao discutir sobre a duração extrínseca da vogal, relata que, em línguas como o chinês, tcheco e árabe, as vogais que antecedem segmentos oclusivos sonoros não são, necessariamente, mais longas, como prevê a literatura (CHEN, 1970, *apud* NISHIDA, 2007). Segundo Nishida (2007:131), a autora (KEATING, 1985) “observa que o fato é específico de língua e não apenas um fato implementacional, fonético, sendo, com isso,

descrito na gramática da língua”. A autora aponta, ainda, para o papel que a fonética tem dentro do gerativismo. Segundo ela, a) a Fonética aparece como regra que converte os valores de binário para escalar, que faz parte da gramática da língua; b) a parte que é, convencionalmente, chamada de fonética não está, tecnicamente, na gramática. Trata-se de um componente universal e previsível que realiza a tradução de uma transcrição fonética em parâmetros físicos contínuos. Assim, Keating acaba por concluir que muitas das questões não resolvidas pelo modelo da fonologia gerativa partem do modo como se vê o dado fonético e, em última instância, o papel da Fonética na gramática de uma determinada língua.

Com o ponto levantado por Keating, passa-se, então, a observar o surgimento e funcionamento de modelos que possuem um primitivo de análise que seja dotado de tempo intrínseco.

Uma discussão interessante é levantada por Port e Leary (1995), ao mencionarem que os sistemas dinâmicos, advindos da matemática, são sistemas que se configuram pela presença de estados numéricos que acabam por se desenvolver através do *continuum* temporal, mas não de modo aleatório, de acordo com regras que são internas e externas ao sistema. Uma característica interessante desses sistemas é que estes são passíveis de mudanças, não sendo estáticos, portanto. Quando analisamos o funcionamento desses sistemas, observamos que um “estado jamais é separado de forma estanque do próximo, já que há uma ideia de continuidade entre os estados, que deve ser preservada na relação temporal” (ELMAN, 1998, *apud* BLANK e ZIMMER, 2011:29).

De maneira análoga à relação de dependência entre o dado do *output* e o *input*, pressuposto de base de teorias como a Turbidez, a relação entre os sistemas dinâmicos também estabelece uma relação de dependência, mas esta não é unilateral como na Teoria da Turbidez, na qual o *output* sempre está conectado com as regras estabelecidas no *input*, sem que a influência possa se dar no outro sentido (do *output* para o *input*). Sistemas como os dinâmicos funcionam a partir de associações em paralelo. Existe uma grande distinção de

base entre essas duas visões, de sistemas categóricos, como a Teoria da Turbidez, e os sistemas dinâmicos, como é o caso da Fonologia Gestual (a qual será trazida para comparação no final desta seção). Um sistema dinâmico não se inicia em um ponto estacionário, pois se entende que um dado sistema sempre está integrado a outros sistemas que já podem ter sido acionados. Não existe, portanto, uma relação hierárquica, como nos modelos categóricos, nos quais os segmentos vão sendo agrupados de acordo com um ordenamento que funciona para todas as línguas. É possível que se faça uma comparação entre os sistemas dinâmicos e os meteorológicos: a chuva não depende apenas de uma variável, como a evaporação da água dos rios, mas também da formação de frentes frias, por exemplo. Estas, por sua vez, têm sua ação condicionada a sistemas de pressão (que a impedem de atingir um determinado local ou permitem que ela o faça) ou a ventos, por exemplo, que a empurram para a costa ou para o continente. Em resumo: há uma diversidade de fatores que, conjugados, resultam em chuva numa determinada região. Da mesma maneira, o funcionamento dos sistemas dinâmicos pressupõe a interação de inúmeras variáveis.

As relações estabelecidas entre os sistemas dinâmicos é que determina o funcionamento de um dado sistema (KELSO, 1995). Como é mencionado por Kelso (*op cit.*), os sistemas não operam somente no tempo, mas também no espaço, o que quer dizer que dentro dos sistemas dinâmicos existem subconjuntos e alguns desses tendem a ocupar um lugar de preferência em detrimento de outros possíveis espaços. Isto é que confere a esses sistemas a possibilidade de funcionarem numa relação de atração ou repulsão. No entanto, essa relação opera de modo distinto para os modelos dinâmicos, como a Fonologia Gestual e o conexionismo. Modelos conexionistas são arquitetados com base em atratores e repulsores. De acordo com Blank e Zimmer (2011:30), “um pode ser entendido como um conjunto de soluções em que conjuntos de estados convergem, após um intervalo de tempo”. Da mesma maneira, existe a possibilidade de que os conjuntos não sejam atraídos por certos pontos e, neste caso, tem-se uma relação de repulsão. A Fonologia

Gestual (BROWMAN e GOLDSTEIN, 1996), por outro lado, utiliza uma equação que modela sistemas dinâmicos simples, na qual não estão presentes atratores ou repulsores, mas variáveis que se relacionam entre si e que terão maior ou menor influência no resultado dessa equação em função da sua grandeza, i.e., do valor que assumem na equação.

Tomando como base a perspectiva de sistemas dinâmicos, um olhar que assuma tal perspectiva deve permitir enxergar o processo sob uma perspectiva diferente. Como já mencionado, Oostendorp (2007) propôs em seu estudo um modelo que dava conta, em partes, de fatos gradientes (como é o caso da DT). Mas o tratamento que ele dá ainda se vale de um primitivo de tempo extrínseco, como o traço. Assim, é possível observar que apesar do autor passar a trabalhar com as restrições, o tratamento oferecido ainda não consegue captar a gradiência entre consoantes surdas e sonoras – considerando que as ensurdecidas se posicionariam entre esses dois pontos extremos do contínuo. A proposta de Browman e Goldstein pode ser promissora nesse sentido, porque eles incorporam a variável tempo ao primitivo de análise que elegem para seu modelo e que é o gesto articulatório. Através da incorporação do tempo ao modelo, pode-se dar conta dos fatos observados na relação de ensino/aprendizagem de uma L2, uma vez que as produções desviantes dos aprendizes envolvem, muitas vezes, tão somente uma “sintonia” dos gestos tal qual feita pelos falantes nativos, ou seja, trata-se de uma reorganização temporal dos movimentos articulatórios que os aprendizes já realizam.

A título de esclarecimento, é importante mencionar que a adoção de um primitivo de análise dotado de tempo intrínseco, num modelo de análise fonológica, surge na proposta de Browman & Goldstein (1986 e trabalhos posteriores) de uma Fonologia Gestual. De acordo com esse modelo, cada som se caracteriza pela presença de um ou mais gestos articulatórios, os quais podem fazer parte de um dos subsistemas articulatórios: oral, laríngeo ou vélico. Os gestos recebem especificações, em cada subsistema, para grau e local de construção. A partir de tais especificações, os gestos acabam por

representar variações, como as de rigidez ou amortecimento, as quais “geram variações de magnitude gestual e de ativação temporal do gesto” (SILVA, 2003).

É mister dizer que os autores precisavam assumir outra unidade organizadora que não fosse o traço, unidade assumida pelos modelos de base gerativa, uma vez que os traços possuem uma configuração binária, e isso não possibilitaria a análise de fenômenos que não se encaixam dentro de categorias bem definidas. Além disso, Browman e Goldstein, ao tomarem como unidade articulatória o gesto, marcam uma posição com relação à não distinção entre fonética e fonologia. Dada essa nova configuração, fenômenos relacionados à aquisição de L1 e L2 podem passar a receber um novo tratamento.

4.1 EXPERIMENTOS COM DT: INFORMANTES BRASILEIROS APRENDIZES DE INGLÊS COMO L2 À LUZ DO CONEXIONISMO E DA FONOLOGIA GESTUAL.

Assumir uma perspectiva dinâmica nas investigações sobre aquisição de L2 implica assumir que os aprendizes de L2 podem produzir os mesmos gestos que os falantes nativos: a diferença está no fato de que os aprendizes precisam coordenar os gestos da mesma maneira que os nativos. Desta forma, o ensino dos sons de uma L2 não deve consistir em ensinar o que os aprendizes não sabem, mas em ensiná-los a melhorar, a refinar, uma produção que já exibem. Essa discussão coloca a prática de correção de “erros” produzidos pelos alunos/falantes sob outra perspectiva, uma vez que, como já foi apresentado na introdução deste trabalho, não existe propriamente uma produção errada, mas um processo de adequação gestual, de acomodação/ sintonia dos gestos, que o aprendiz precisa adquirir.

Pode-se dizer que esse olhar dinâmico para com a DT, especialmente com relação a falantes de PB aprendizes de inglês como L2 teve início com o trabalho de Zimmer (2004). A autora conduziu um experimento sobre a

transferência grafo-fônico-fonológica²⁸ e concluiu que a DT se dava por conta da influência dos segmentos adjacentes à oclusiva final. Contudo, o estudo de Zimmer (*op cit.*) contou com uma análise de dados feita de ouvintes, ou seja, sem a utilização de ferramentas como a análise acústica²⁹. Desse modo, segundo a autora, os achados do seu experimento e, também dos conduzidos por Eckman (1981) e Major (1987), podem ter sido enviesados. A partir desse argumento, decidiu-se conduzir um novo experimento. Zimmer e Alves (2007) realizaram um estudo para verificar a presença da DT na interlíngua de brasileiros aprendizes de inglês como L2 e, para tanto, lançaram mão da análise acústica. Esse estudo contou com a participação de 08 estudantes do sexo feminino, de quatro níveis de proficiência em inglês (iniciante, intermediário, intermediário-superior e avançado – nomenclatura dada através do Oxford Placement Test³⁰). As informantes eram alunas das disciplinas de Língua Inglesa I e II, do curso de graduação em Letras (língua inglesa) da Pontifícia Universidade Católica do Rio Grande do Sul. Os sujeitos do experimento realizaram a leitura de palavras isoladas em uma ficha. Tal instrumento possuía 70 palavras, das quais algumas foram selecionadas para análise (*add, head, rob, rub, drug and bed*). Após a coleta de dados, estes foram analisados no Praat.

²⁸ O presente estudo utiliza o conceito de grafo-fônico-fonológico conforme exposto por Zimmer e Alves (2006). “conhecimento grafo-fônico-fonológico” Indica uma relação entre a forma grafada e os fones da língua em questão. O termo “fônico-fonológico” expressa, tal como o exposto em Albano (2001, 2002), a gradiência e a continuidade entre o fone físico e o fonema, visto aqui como sendo ativado de forma distribuída e em paralelo por unidades neuronais no cérebro. Entende-se por fonema não uma unidade distante daquela de superfície, mas a própria unidade de *output* produzida pelos falantes, de modo que, em uma fonologia de cunho conexionista, os conceitos tradicionais de “fone” e “fonema” correspondam a uma única realidade (conforme pode ser depreendido de Joanisse 1999, 2000, e Bonilha 2004). Acredita-se que a relação expressa no termo “fônico-fonológico” consegue representar bem tal relação, de modo que tal termo, baseando-se na concepção aqui descrita, não esteja fazendo relação a unidades de natureza simbólica.” (Zimmer e Alves, 2006:116)

²⁹ Essa consideração é feita partindo do fato de que pesquisas, como a de Zimmer (2004), apontam para a não percepção do ouvido humano, de certos aspectos fonéticos que podem influenciar na análise final dos dados.

³⁰ O *Oxford Placement Test* é constituído por questões de múltipla escolha de compreensão auditiva e de gramática do inglês. O teste foi validado a partir da aplicação, ao longo do período de cinco anos, a um universo de aprendizes de mais de 40 nacionalidades.

A partir da análise acústica, os autores encontraram uma gradiência na produção das oclusivas finais, ou seja, ao invés da dicotomia surda ou sonora, proposta por um sistema de traços, os autores puderam constatar que as produções dos aprendizes se davam num *continuum* que compreendia segmentos surdos, sonoros e parcialmente sonoros ou surdos. Tal análise foi feita à luz de modelos conexionistas e dinâmicos. Com isso, os resultados também apontaram para uma evidência contrária à exposta por Eckman e Major, de que a distinção surdo-sonora seria neutralizada. Nas considerações finais os autores mencionam a possível influência de segmentos adjacentes na sonoridade da oclusiva e suscitam a necessidade de que outros parâmetros acústicos, como a duração da vogal que precede a consoante, bem como o tempo de duração do período de oclusão e o de soltura, sejam analisados.

A revisão da DT através do conexionismo e da Fonologia Gestual, modelos de base dinâmica, possibilitou um tratamento mais parcimonioso à DT, uma vez que um dos pontos centrais para a representação desse e de outros fenômenos parece incidir sobre a natureza do primitivo de análise.

É nessa direção que os autores Zimmer e Alves realizam outro estudo (2008), com o objetivo de investigar os resultados obtidos no trabalho de 2007. A pesquisa conduzida em 2008 também é de base conexionista. Nesse sentido, é preciso lembrar que entre os pressupostos de base tem-se que o *input* é suficientemente rico para conduzir o aprendizado a partir de informações que estão relacionadas de forma a estabelecer redes de conexão. Segundo os autores,

- a) O ambiente linguístico é rico em regularidades que são distribuídas e guiam o aprendizado da língua;
- b) O processo de aquisição de uma língua necessita da exploração de restrições probabilísticas que estão inseridas em vários tipos de informação linguística e não-linguística;
- c) É difícil definir uma linha clara entre conhecimento linguístico e não-linguístico, uma vez que o sucesso do aprendizado depende

tanto da estrutura do *input* bem como do conhecimento prévio do aprendiz;

d) A informação distribucional pode oferecer evidências implícitas para a aquisição de uma língua.

(Rhode e Plaut, 2003; Seidenberg e McDonald, 1999, *apud* Zimmer e Alves, 2008: 43)

Além de o *input* ter um papel diferente dentro do modelo, os autores propõem que noções como as restrições de marcação propostas por Eckman (1981, 1987, já mencionadas na seção 3 deste trabalho) sejam substituídas. De acordo com os autores, as restrições de marcação possuem múltiplos sentidos e poderiam ser expressas por outro conceito menos ambíguo. Zimmer e Alves mencionam que Haspelmath (2006, *apud* ZIMMER e ALVES, 2008) descreve a existência de 12 sentidos para o termo e acaba agrupando os diferentes significados em quatro categorias: marcação como complexidade, como dificuldade, como anormalidade e como correlação multidimensional. Embora não sejam discutidas aqui as especificidades de cada uma das categorias, é importante notar que o termo marcação, a *markdness*, como já mencionado na seção 3, possui um sentido distinto também a depender do modelo teórico (a OT, por exemplo, utiliza-o de forma diferente, como já foi previamente discutido). Assim, Zimmer e Alves (2008), para explicar como as informações são selecionadas na representação, optam por utilizar a experiência linguística do falante e a noção de transferência linguística ao invés de entenderem que os sistemas possuem restrições que operam nas línguas e que tais restrições não levam em conta, necessariamente, o modo como os sujeitos em questão aprenderam uma determinada língua.

Com base no experimento desenvolvido em 2007, Zimmer e Alves (2008) possuíam as seguintes hipóteses iniciais: a) os estudantes brasileiros produzem um percentual significativamente menor de vozeamento durante o período de oclusão em comparação com a produção dos americanos; b) a vogal que antecede o segmento oclusivo final é um parâmetro significativo para a distinção do contraste surdo/sonoro.

O estudo de Zimmer e Alves (2008), que contou com 08 falantes brasileiras, com idade entre 19 e 26 anos, e 3 americanas, com idade entre 25 e 32 anos, corroborou a hipótese inicial que previa que os aprendizes brasileiros iriam apresentar uma menor porcentagem de vozeamento durante o período de oclusão, quando comparados com os falantes americanos. Foram rodados testes estatísticos que revelaram que, com exceção de dados como “rag”, todos os resultados mostraram que os dados dos brasileiros possuíam valores percentuais mais baixos. Segundo os autores, o fato da oclusiva velar sonora ter obtido um percentual distinto das demais está relacionado com os relatos da influência do ponto de articulação, que podem ser vistos em Ladefoged & Maddieson (1996). Os achados com relação à influência do ponto de articulação apontam para o fato de que a taxa de vozeamento diminui à medida que a língua é direcionada para a porção posterior do trato, ou seja, a região velar. Assim, os segmentos velares têm a tendência de possuir valores menores para a duração do vozeamento, quando comparados com segmentos bilabiais e alveolares.

A segunda hipótese, relativa à possível influência da duração da vogal antecedente à oclusiva final sobre a distinção dos pares surdo/sonoros, foi parcialmente corroborada. Os resultados foram significativos com relação a alguns pares (*cap/cab*, *back/bag* e *bop/bob*). De acordo com os autores, a diferença entre os pares pode ser explicada através da frequência de uso, pois as palavras com as quais os aprendizes têm menos contato tiveram resultados distintos. Ainda segundo os autores são necessários mais dados para confirmar se a duração da vogal pode ser considerada uma pista para a distinção entre oclusivas finais surdas e sonoras. Segundo Zimmer e Alves, os aprendizes parecem transferir o conhecimento fonético-fonológico da duração de vogal em PB para a L2, o que significa que como em PB a duração de vogal não é um processo que ajuda no contraste surdo-sonoro, os aprendizes ainda precisam aprender a utilizar tal pista para fazer a discriminação dos dados do inglês.

Embora os dados de duração da vogal que precede o segmento oclusivo final e do percentual de vozeamento sejam distintos dos americanos, é possível dizer que não houve casos de neutralização total, os informantes-aprendizes continuaram sonorizando ainda que com percentuais de sonorização menores do que os nativos e com duração de vogal mais longa em segmentos oclusivos surdos quando comparados com os americanos. Isso vai de encontro à proposta de que a dessonorização terminal seria um fenômeno que ocorreria na interlíngua. Uma vez que não se notou perda do traço sonoro nem neutralização do contraste surdo-sonoro, a hipótese de Eckman (1987) e Major (1987) de que a dessonorização terminal seria uma regra utilizada na transferência da L1 para a L2 não se apresenta como viável. Os autores se posicionam, então, a favor de um tratamento dinâmico para os dados de DT.

Tomando o estudo de Zimmer e Alves (2007 e 2008) como referência, mostraremos um experimento, de natureza preliminar, que procurou investigar a influência de outros parâmetros acústicos na DT.

4.1.1 DT (?): Investigação da influência de parâmetros acústicos e contextos adjacentes na aquisição de L2 a partir de modelos dinâmicos.

Como relatamos anteriormente, já foram realizados vários experimentos, em várias línguas, os quais analisaram o processo de neutralização total e parcial que se encontram associados à DT.

Nos experimentos conduzidos na década de 1980, foram realizados estudos com base em teorias gerativas que relataram a presença de uma neutralização da sonoridade em oclusivas finais para, pelos menos, falantes de chinês e espanhol, aprendizes de inglês como L2. Mais tarde, outros estudos, através da OT, tentaram incorporar os dados empíricos de neutralização incompleta descritos por alguns autores e, para tanto, lançaram mão de teorias auxiliares, como a Turbidez para darem conta da DT. Ainda assim, apesar de o fenômeno passar a ser representado de maneira distinta, sendo que o falante passaria a ter em sua fonologia [t], [d] e [d⁰], os dados de neutralização

incompleta trazem à tona o caráter gradiente da DT, fazendo com que os primitivos de base (como por exemplo, os traços) precisem ser repensados. A Fonologia Gestual, modelo que não dissocia fonética e fonologia, traz o tempo para a representação do falante e, assim, mostra que a aquisição dos segmentos passa por um processo de faseamento dos gestos³¹ que, no caso da DT, acaba por gerar uma produção que é gradiente. Entender a produção de segmentos como gradiente, ou seja, sendo realizada em um *continuum* temporal, indica que a produção sonora não se reduz a encaixar sons em categorias fixas, mas sim que existe um leque de possíveis produções de um determinado som. Dentro deste estudo, é possível apontar, por exemplo, a produção das oclusivas sonoras e surdas por falantes não-nativos. Como será visto posteriormente, um mesmo segmento [d] pode ser produzido de modo diferente pelos aprendizes de L2, a depender de como esses manipulam as características acústicas do segmento, a saber: percentual de vozeamento, presença ou não do *burst*, duração da vogal que precede a oclusiva. Ao tentar sonorizar o segmento, o falante-aprendiz pode não ativar ou ativar parcialmente o gesto de vibração de pregas, mas isso não significa, necessariamente, que ele não tenha produzido [d]. Outros parâmetros acústicos estão envolvidos na caracterização do segmento. O aprendiz aprende pouco a pouco a manipular a magnitude dos gestos necessários para a produção de uma determinada sequência sonora.

A possibilidade de se entender o sistema de articulação de sons a partir da coordenação dos gestos que se realizam no tempo, conforme mencionado na introdução deste texto, permite que se tenha o gesto articulatório como uma unidade contrastiva (Nishida, 2009). Segundo o pesquisador Nishida, os segmentos são contrastados pela “[...] a) presença ou ausência de um gesto; b) pelos diferentes valores das variáveis da equação dinâmica; e c) pelas diferentes organizações possíveis entre os gestos – o faseamento” (NISHIDA,

³¹ Conforme expõe Nishida (2007:138) “Como os gestos são unidades com duração interna, é possível representar diretamente a sobreposição temporal entre dois gestos. As diferentes sobreposições (chamado faseamento ou sintonia) dos gestos é que produzem diferentes sons”.

2009:34). Diferentemente da representação das outras abordagens/modelos teóricos já aqui trabalhados, que adotam o traço como unidade de análise, a Fonologia Gestual prevê que os gestos articulatórios são formalizados em pautas gestuais. Conforme pode ser visto nas figuras abaixo, retiradas de Nishida (2009) e baseadas em Browman e Goldstein (1992), a representação se dá em dois eixos: vertical e horizontal. O primeiro demonstra gestos envolvidos na articulação dos segmentos e a sua magnitude, já o segundo traz a representação do tempo de ativação e desativação dos gestos.

Browman e Goldstein (1992:158), os proponentes da Fonologia Gestual, demonstram como funcionam as pautas gestuais para um par mínimo do inglês: “*pad*” e “*bad*”. Tais exemplos serão importantes para a discussão sobre como a DT pode ser tratada a partir de um modelo dinâmico.

Conforme pode ser visto na Figura 1, para que se produza o segmento “*bad*”, é necessário o envolvimento de três gestos: corpo de língua (CL), ponta de língua (PL) e lábios. Cada um dos gestos se sobrepõe temporalmente, ou seja, um gesto inicia ainda durante o tempo de ativação de outro. O gesto de CL, que é responsável pela produção da vogal do segmento, inicia ao mesmo tempo em que o de lábio fechado (que representa a oclusão da oclusiva bilabial sonora [b]). Contudo, o gesto de PL (que caracteriza a oclusiva alveolar sonora [d]) começa quando o gesto de CL está sendo executado e ambos terminam ao mesmo instante temporal.

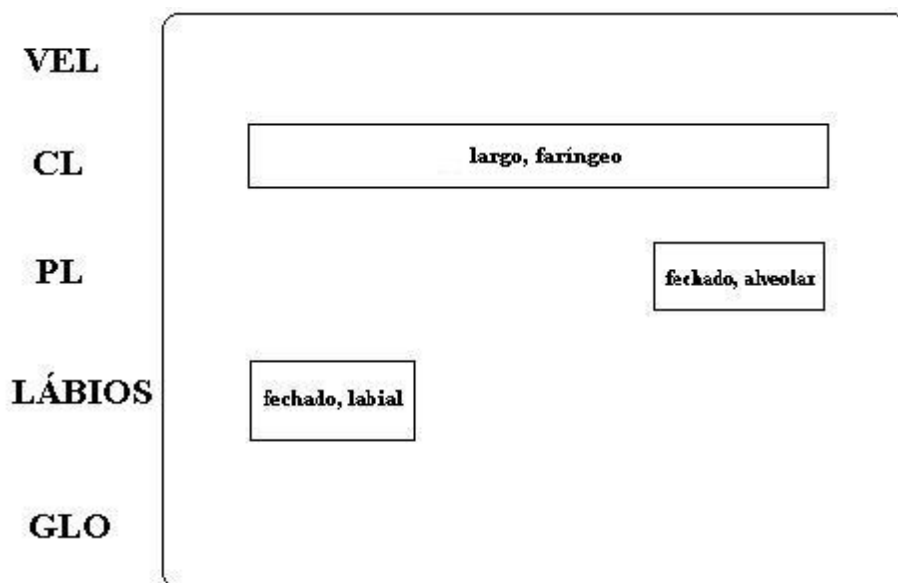


Figura 1 Pauta gestual da palavra "bad" (Browman & Goldstein 1992:158 apud Nishida, 2009: 35). VEL: abertura vélica; CL: corpo de língua; PL: protusão labial; GLO: abertura glotal.

Em oposição à Figura 1, quando se observa a representação gestual para o segmento "pad", na Figura 2, nota-se que não só os gestos ativados, bem como a magnitude³² desses se assemelham às configurações para segmento "bad". A pauta gestual de "pad" se distingue pela ativação do gesto largo de glote (GLO), que acarreta a produção da oclusiva bilabial surda [p]. É interessante notar, também, que, o gesto de GLO inicia após os gestos de CL e fechamento de lábios, o que, segundo, Nishida (2009:36) mostra "[...] que um segmento surdo não se caracteriza por seu ensurdecimento completo e tampouco acaba com o fim da oclusão bilabial".

³² A magnitude gestual é dada pela extensão vertical das "caixinhas" que representam os gestos, enquanto a ativação temporal é dada pela extensão horizontal das mesmas caixinhas.

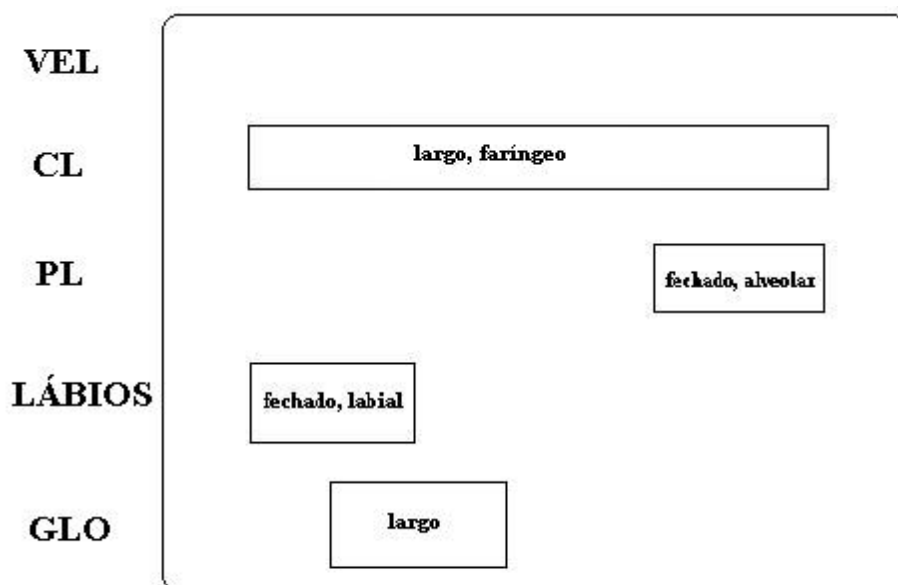


Figura 2 Pauta gestual da palavra “pad” (Browman & Goldstein, 1992:158 apud Nishida, 2009: 36). VEL: abertura vélica; CL: corpo de língua; PL: protusão labial; GLO: abertura glotal.

As Figuras 1 e 2 representam a possibilidade de fenômenos como a sobreposição de gestos. A sobreposição gestual gera, no limite, segmentos distintos. Mas, entre dois segmentos distintos (duras categorias, portanto), gera a gradiência observada. Mas é preciso lembrar que tal representação só se mostra possível uma vez que o tempo é intrínseco para a Fonologia Gestual. A demonstração das pautas gestuais, Figuras 1 e 2, serão elementos importantes na discussão dos dados do experimento, que, agora, passa a ser resenhado.

4.1.2 DT: Experimento de produção de oclusivas finais do inglês com falantes Brasileiros a partir da Fonologia Gestual.

Albuquerque (2010) toma informantes brasileiros, aprendizes de inglês como L2, para investigar a possível influência de parâmetros acústicos sobre a DT. Além do vozeamento da oclusiva, já observado por Zimmer e Alves (2007), Albuquerque verifica também a duração do vozeamento e a duração da vogal que precede a oclusiva. Esses parâmetros foram escolhidos para o estudo porque nosso ouvido não consegue fazer uma análise acurada de algumas nuances como, por exemplo, uma sonorização que não é mantida em toda a

extensão do período de oclusão. Com a análise acústica dos dados, percebeu-se também, como dissemos acima, que não bastava olhar para a presença da barra de sonoridade nos segmentos, mas que outros fatores, como a duração da vogal e a duração da oclusão podiam ser importantes para a análise, uma vez que a sonoridade do segmento oclusivo também dependia dos contextos adjacentes. Nesse sentido, se mostra de extrema importância que, ao falar da metodologia de coleta e análise de dados, seja levado em conta um *corpus* que contemple ambientes adjacentes distintos, para averiguar se há ou não uma distinção de produção a depender do contexto, e, também, para observar se os resultados encontrados são os mesmos para todas as oclusivas.

Com relação à metodologia, o *corpus* era composto por 72 sentenças, das quais 48 eram sentenças-alvo e 24 eram sentenças distratoras. As palavras-alvo foram escolhidas de modo que fossem obtidos dois exemplos de todas as oclusivas em posição final. A escolha das palavras pautou-se no fato de que, em português brasileiro, não há consoantes oclusivas nessa posição. Abaixo, pode-se conferir a lista das palavras utilizadas nas sentenças, as quais foram dispostas em fichas para serem lidas pelos informantes:

1. Seis itens terminados em obstruintes sonoras, escolhidos de forma a neutralizar a variável transferência grafo-fônico-fonológica: *bad, pad, bob, cab, rag, bag*;
2. Dois itens terminados em fricativas sonoras, escolhidos de forma a ensejar a transferência grafo-fônico-fonológica: *falls, bobs*;
3. Oito palavras correlatas terminadas em obstruintes e fricativas surdas das palavras dos itens 1 e 2, para comparação das medidas acústicas: *bat, pat, bop, cap, rack, back, false, pops*.

O *corpus* contemplou sentenças que possuíssem palavras com obstruintes de diferentes pontos de articulação. Além disso, Albuquerque (2010) contou com um controle dos contextos adjacentes apresentados aos informantes. Assim como como Zimmer e Alves (2007), o objetivo era verificar a influência dos contextos na produção das oclusivas: inexistente, vocálico e consonantal (plosivo). As sentenças-veículo utilizadas se encontram abaixo:

- 1.Say _____. (contexto 1)
- 2.Say _____ again. (contexto 2)
- 3.Say _____ promptly.(contexto 3)

Para este experimento, foi realizada uma coleta de 5 repetições de cada conjunto de sentenças, com todos os 6 informantes (5 paranaenses, entre 19 e 24 anos e 1 nativo, falante dos Estados Unidos, de Stevens Point - Wisconsin, de 27). Os informantes aprendizes eram estudantes do nível intermediário no Celin (Centro de Línguas e Interculturalidade). Foram gravadas 5 repetições de cada uma das sentenças, para que não houvesse nenhum viés idiossincrático por parte dos informantes. Os dados foram colhidos em cabine com tratamento acústico no LEFON³³, a uma taxa de amostragem de 44100 Hz. Em seguida, procedeu-se à sua análise através do *software* Praat (Boersma e Weenick, 2006).

Todas as sentenças foram recortadas no programa Praat. Após terem sido recortadas, foram extraídas a duração dos três parâmetros já aqui elencados. Com as medidas extraídas, foi calculada a duração relativa dos dados para observar qual era a duração real dos segmentos através de uma média. Não realizar a duração relativa dos dados implica obter valores que podem se mostrar robustos, mas possivelmente enviesados, pois as sentenças possuem tamanhos diferentes.

É importante enfatizar que a partir da utilização de programas de análise acústica do sinal de fala obtêm-se resultados com base nos parâmetros estabelecidos pelo pesquisador por julgá-los relevantes para uma investigação específica. Estes oferecem percentuais que são extraídos pelos pesquisadores, a partir do que se mostra relevante para uma dada pesquisa. Como menciona Price (1998:81), ferramentas tecnológicas nada fazem sem a instrução do

³³ O LEFON é o Laboratório de Estudos Fônicos, localizado na UFPR (Universidade Federal do Paraná). É um laboratório de fonética que possui uma cabine com tratamento acústico para fazer gravações de dados de fala.

pesquisador e, sem o olhar do pesquisador, “a tecnologia para dados da fala em nada contribui para o aprendizado da linguagem³⁴” (tradução minha).

Os programas de análise de fala têm sido largamente utilizados nos estudos dos últimos anos e têm contribuído para o estudo de novos e antigos fenômenos que envolvem a aquisição de L2. Muitos fenômenos que haviam recebido um tratamento de outiva, ou seja, cujos resultados dos experimentos se pautavam na impressão auditiva do próprio autor, hoje podem receber um tratamento mais acurado e, ao mesmo tempo, menos enviesado, uma vez que se tem relato da capacidade limitada de percepção do nosso aparelho auditivo para alguns detalhes finos de produção, como por exemplo, do fenômeno que é foco desse trabalho: a DT.

O objetivo de Albuquerque (2010), como relatado anteriormente, era investigar a realização da DT para aprendizes de inglês à luz de uma análise acústica, uma vez que poucos estudos sobre o fenômeno da DT dispunham naquele momento de análise acústica. A análise acústica mostrou-se necessária a partir das evidências de neutralização parcial em línguas como o alemão e o holandês, e de pesquisas como a de Zimmer e Alves (2007) que observaram a importância de se analisar os contextos adjacentes à plosiva-alvo. Diante desse quadro, elaboraram-se as seguintes hipóteses: 1) como a fonotaxe do português não permite a ocorrência de oclusivas em trava silábica, espera-se que os falantes introduzam uma vogal em sílabas do inglês terminadas com oclusivas; 2) os informantes ensurdeceriam as consoantes sonoras diante de outra plosiva surda e sonorizariam diante de um contexto vocálico, sonoro; 3) haveria uma transferência grafo-fônico-fonológica em dados como [falls] e [false] uma vez que a lateral final em inglês é pronunciada e no PB a lateral palatalizada sofreu um processo de vocalização, que leva à produção da aproximante lábio-velar [w] em posição final de sílaba ou palavra; 4) com relação ao ensino-aprendizagem, e de acordo com a abordagem

³⁴ Trecho original de Price (1998:81) “Without people who know how to use them, books do nothing for literacy, computers do nothing for computation, and speech technology does nothing for language learning”.

dinâmica, pode-se pensar que os alunos realizam produções bem próximas às de um falante nativo. Para verificar tais hipóteses, primeiro se discutirá a variabilidade presente em alguns dados – tanto da falante nativa, como dos aprendizes –, uma vez que as considerações com relação à gradiência na produção dos sons se deram a partir da observação da variação de alguns parâmetros acústicos nos espectrogramas.

A análise dos dados conduziu para alguns achados interessantes no que diz respeito à diferença na produção de alguns segmentos observada nos dados de um mesmo falante. Não é possível afirmar que um dado informante produzirá sempre um mesmo som para um mesmo conjunto de dados. Na Figura 3, tem-se a primeira repetição da sentença *say bag promptly*. Nas figuras 3 e 4, referentes a esse dado, é possível notar uma variabilidade nas produções. Enquanto na figura 3 não é possível visualizar, na parte em destaque do espectrograma, a barra de sonoridade, na Figura 4, na segunda repetição dessa sentença, nota-se a presença da sonoridade. É importante perceber, no entanto, que a caracterização da produção da oclusiva velar sonora [g] não se limita necessariamente à presença da sonoridade. Existem outras pistas, como a duração da vogal que antecede e o espaço de oclusão, que auxiliam na identificação do [g]. Isso reforça a necessidade de se fazer uso da análise acústica, pois, em uma análise de ouvinte, talvez, um pesquisador igualasse as duas produções, de modo a dizer que o falante sonoriza, ou não. Além disso, observou-se uma oscilação no que diz respeito às produções do informante GU, mas isso não indica, necessariamente, que ele não entenda a articulação de sons plosivos finais, mesmo que este não tenha essa produção em sua língua materna. Este dado é representativo da produção de todos os informantes-aprendizes. É importante ressaltar que o mesmo se verifica para as produções da falante nativa PM, conforme as Figuras 5 e 6.

Além dessas observações, é possível notar, também, que o contexto adjacente é uma plosiva surda, o que, de acordo com nossas hipóteses prévias e estudos anteriores, faz com que o informante produza algo diferente do esperado. Se observarmos novamente a Figura 4, notamos que o informante

GU sonoriza mesmo diante de um segmento surdo. Com isso é possível endossar, aos poucos, a hipótese de que os alunos estão produzindo os segmentos-alvo de uma maneira que não se distancia das produções de um nativo, como poderá ser visto em outro espectrograma.

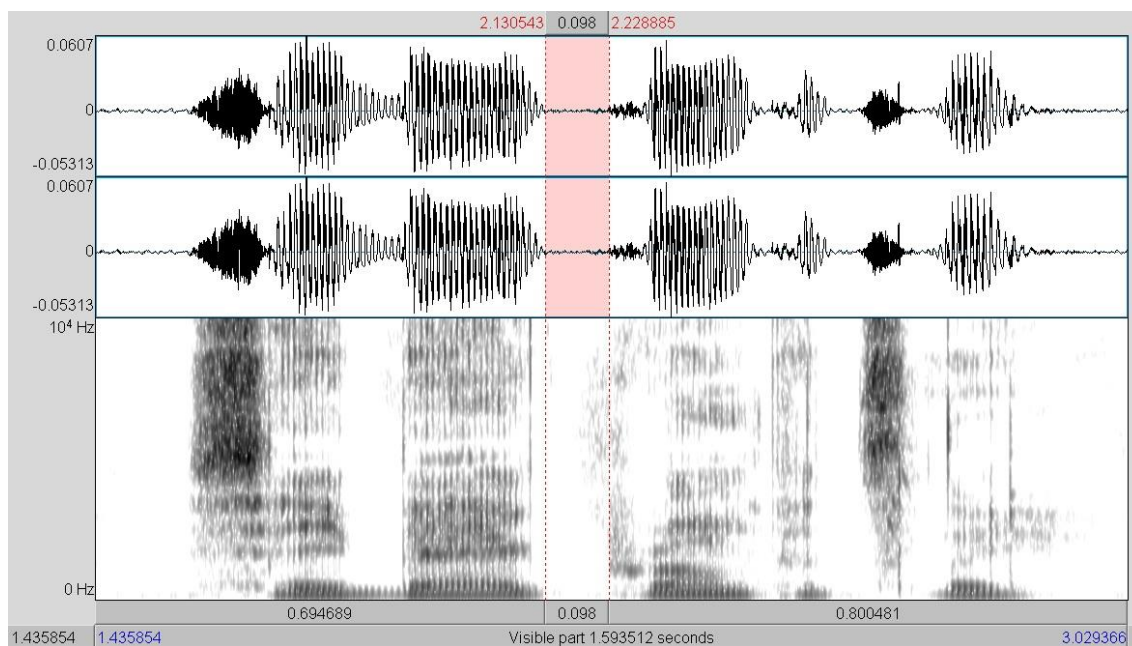


Figura 3 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento bag na primeira repetição sentença “say bag promptly”, produzida pelo informante GU.

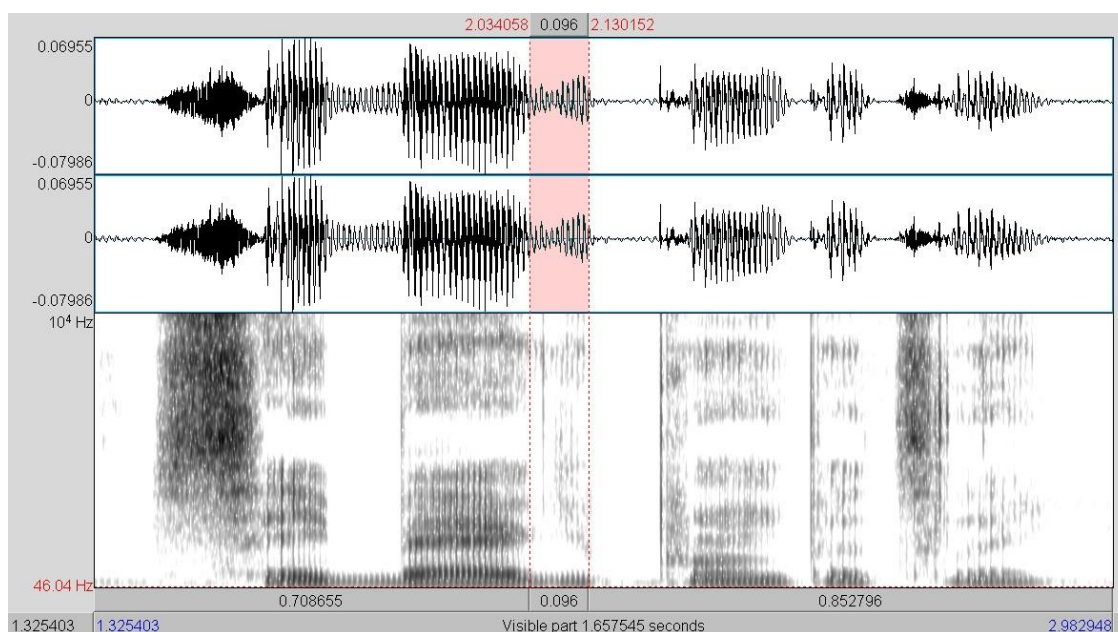


Figura 4 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento bag na segunda repetição da sentença “say bag promptly”, produzida pelo informante GU.

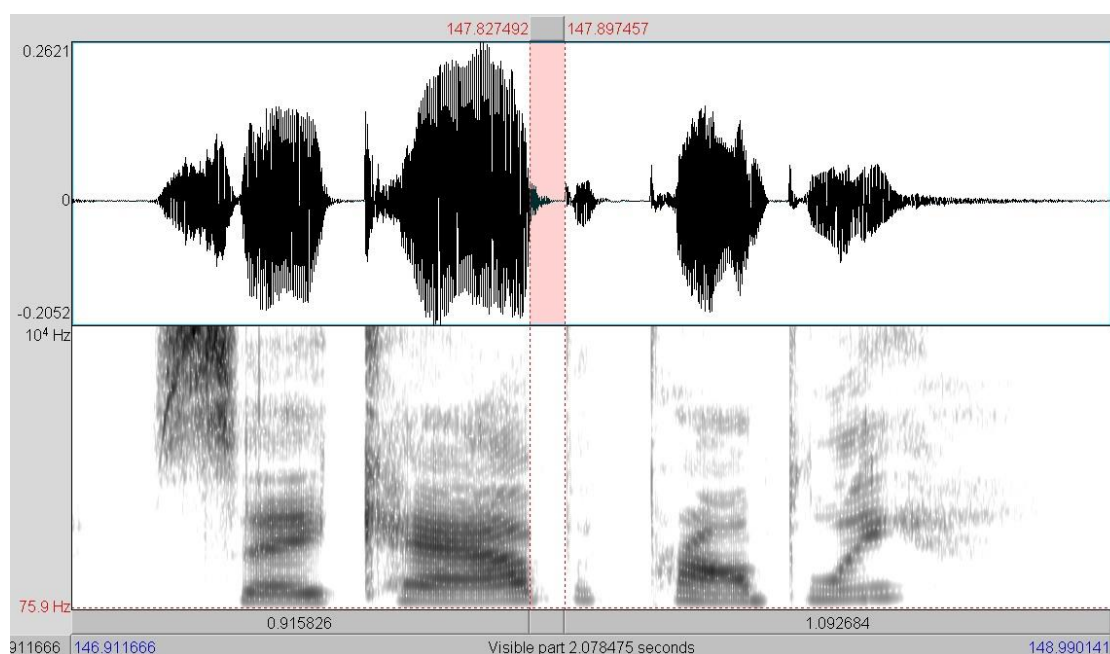


Figura 5 No espectrograma, entre linhas verticais pontilhadas, a oclusão após o segmento cab na primeira repetição sentença “say cab promptly”, produzida pela informante PM.

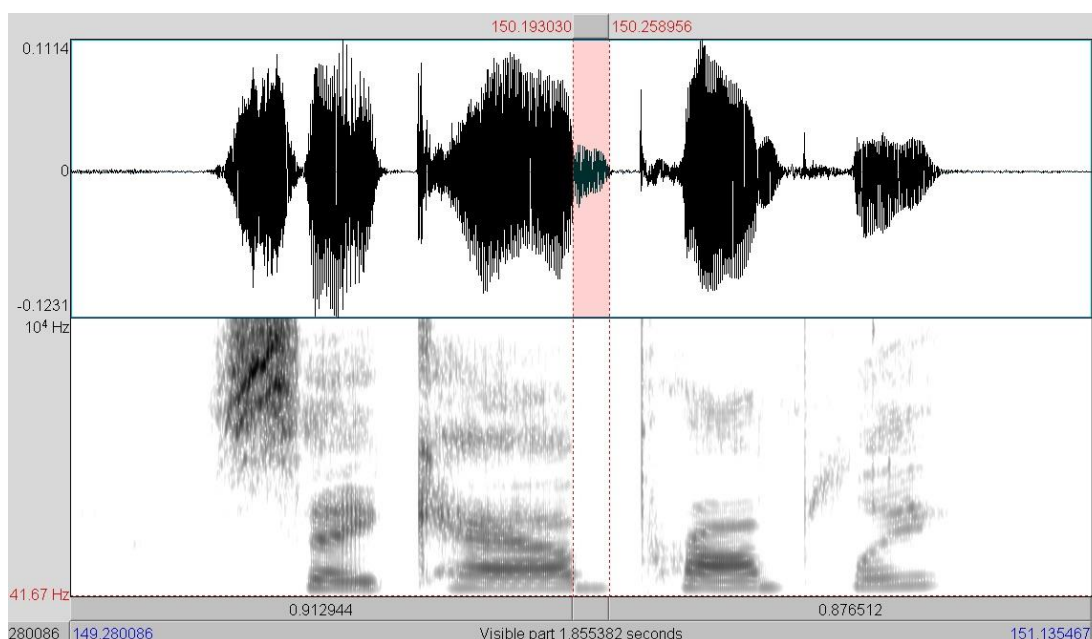


Figura 6 No espectrograma, entre linhas verticais pontilhadas, a oclusão com a presença da barra de sonoridade do segmento cab na terceira repetição sentença “say cab promptly”, produzida pela informante PM.

Outro aspecto interessante é que nem sempre foi encontrada a produção da *burst* nas oclusivas finais, sejam elas surdas ou sonoras. Em uma inspeção de outiva, um falante nativo poderia dizer que o aprendiz produziu ou [t] ou [d], mas como pode ser visto nos espectrogramas 7 e 8, há uma variação entre a visualização da barra de sonoridade, característica de sons sonoros e do *burst*, parâmetro acústico das oclusivas. Na figura 7, não se observa a presença da barra de sonoridade, nem do *burst*. Já na Figura 8, nota-se uma pequena concentração de energia no espectrograma, mas não é possível interpretar a barra de sonoridade como indicativa da sonoridade da consoante. Na imagem, observa-se a atividade das pregas seguinte à realização das vogais. Mais uma vez, caso a análise acústica não tivesse sido utilizada, por conta da sonoridade do contexto adjacente, uma vogal, poder-se-ia dizer que o informante teria sonorizado na Figura 7. Além disso, é preciso, também, lançar um olhar crítico para o percentual de sonorização nesta e em outras pesquisas.

Entre os dados analisados, existe uma oscilação na pista de sonoridade. Uma vez que os informantes-aprendizes estão adquirindo os parâmetros acústicos para produzir as oclusivas em posição final, a realização desses segmentos possui características específicas, como por exemplo, presença de uma duração de vozeamento incompleta, variação no tamanho da vogal que antecede a consoante, presença ou não do *burst*, etc. As figuras 7 e 8 funcionam como um exemplo da não presença de aspectos acústicos, o *burst* e a barra de sonoridade, nas produções de aprendizes.

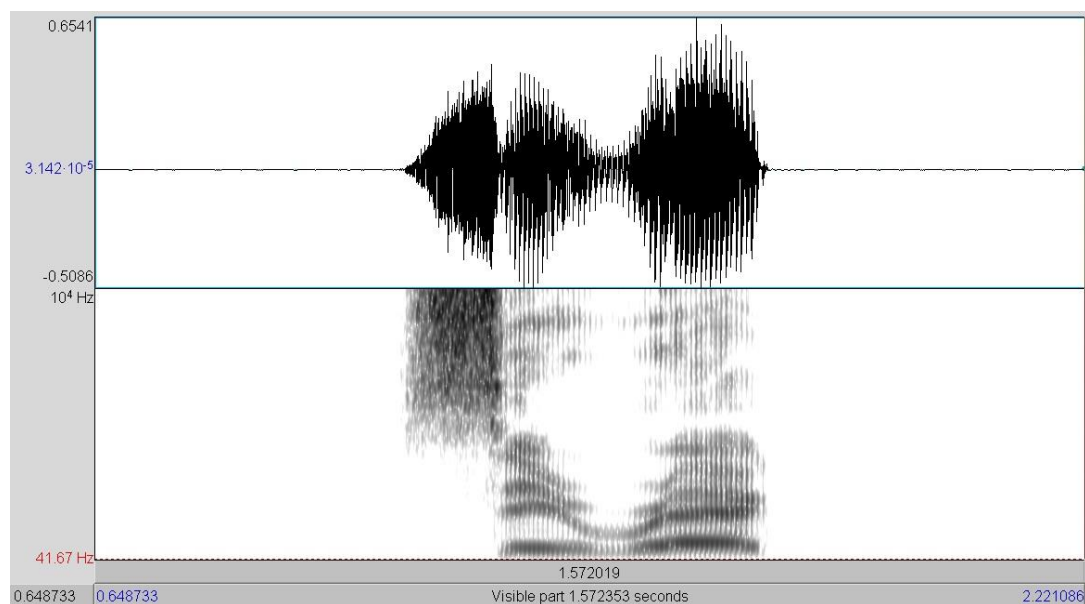


Figura 7 O espectrograma da quinta repetição da palavra rag , na sentença “say rag” produzida pelo informante JU.

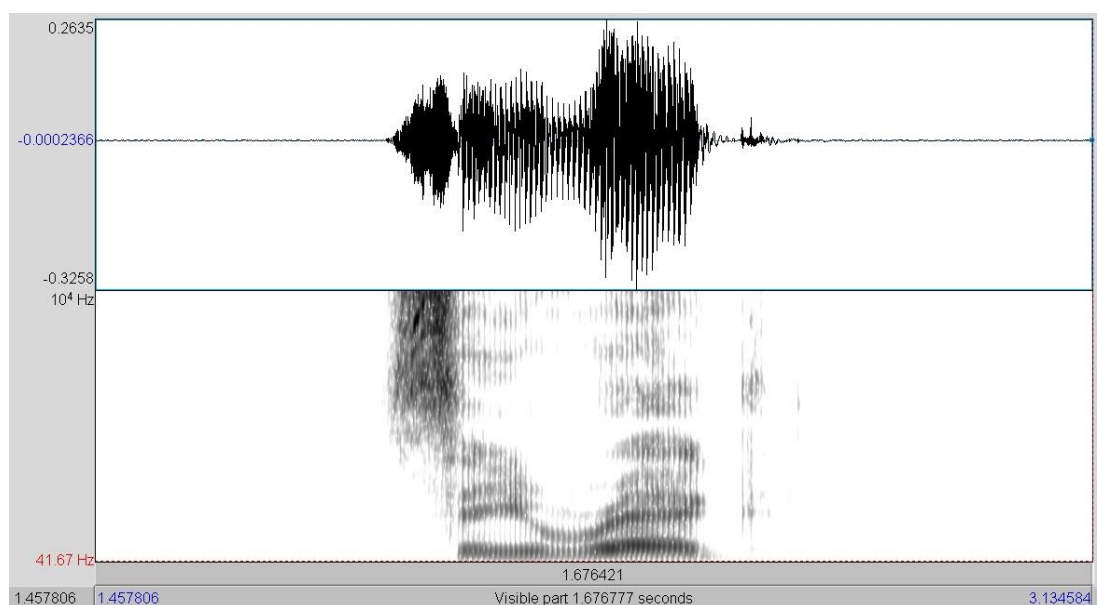


Figura 8 O espectrograma da primeira repetição da palavra rag, na sentença “say rag”, produzida pelo informante JU.

Conforme foi mencionado para os espectrogramas do informante brasileiro JU, não se pode observar a presença da barra de sonoridade, nem o *burst*. Da mesma maneira, é interessante notar que, na produção da informante nativa PM (Figura 9), não se observa, também, a presença do *burst*. A informante produziu um [g], pois se sabe que a sentença tinha um segmento que era *bag*, porém, como não há presença do *burst*, é possível caracterizar esse segmento como uma oclusiva? Sim, é. A *conditio sine qua non* para a caracterização de uma plosiva é a oclusão. Assim como estudos mais antigos, Selkirk (1982), Ladefoged (1993), e mais recentes, Alves (2004, 2008), Bettoni-Techio (2005) já descreveram, o *burst* pode, muitas vezes, ser realizado pela ação tão tênue dos articuladores, que o som produzido é muito pouco intenso e, por isso, não captado pelo microfone durante a gravação. De acordo com Kent e Read (1992), em final de palavra, especialmente, é comum que o *burst* não seja realizado. Nesse sentido, pode-se aproximar a produção de PM do informante paranaense JU.

Comparativamente à Figura 8, a duração da barra de sonoridade é mais longa na Figura 9. A barra de sonoridade, observada a partir da concentração

de energia no espectrograma, não é propriamente da consoante, mas ainda a vibração das pregas, seguinte ao movimento de dorso de língua responsável pela produção da vogal. Ou seja, assim como na Figura 8, também aqui a pista da sonoridade não está propriamente na consoante, que parece mesmo não ter sido realizada – ao menos não há indícios disto no sinal acústico -, mas na vogal. A questão subjacente a este espectrograma e ao da Figura 8 pode ser disposta assim: que pistas visuais nos permitem, no sinal acústico, identificar a sonoridade e, inclusive, a realização da consoante? Nestes dois casos, não há nenhuma pista da consoante. O resquício da consoante, no dado acústico, está na vogal que a antecede.

A pista para a visualização da oclusiva está na transição da vogal que a precede. É, inclusive, através desse aspecto, que é possível recuperar a informação sobre o ponto de articulação da consoante.

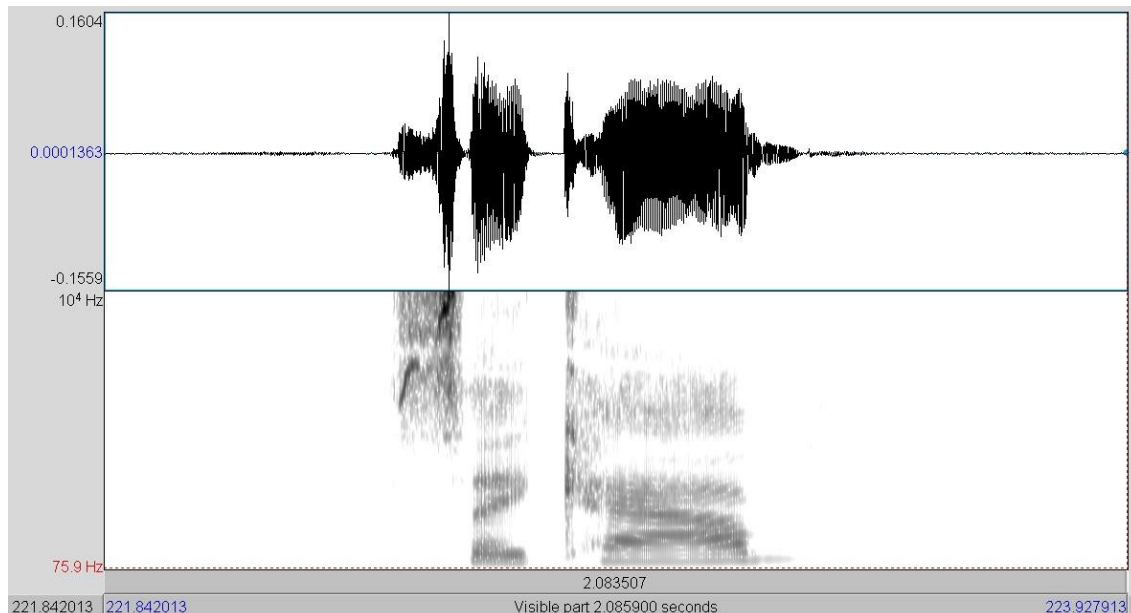


Figura 9 No espectrograma, entre linhas verticais pontilhadas, a barra de sonoridade e a não produção do burst, na primeira repetição sentença “say cab”, produzida pelo informante PM (nativo).

No que diz respeito à transferência grafo-fônico-fonológica, ela ocorreu com todos os informantes aprendizes. As palavras [falls] e [false] foram produzidas da mesma maneira no que diz respeito à produção da fricativa final como surda, em ambas as produções. Contudo, foi interessante observar que a maioria dos informantes tentou deslocar a distinção final para a vogal e as vogais foram produzidas com aberturas diferentes. Para a produção de [falls] os alunos realizaram uma vogal aberta e para [false] fechada. Mesmo que seja possível notar, aqui, a presença da influência dos sons da L1 na L2, observou-se, também, que o aprendiz marca as diferenças utilizando-se de outros parâmetros acústicos.

Passa-se, então, a algumas informações sobre os parâmetros acústicos, os quais podem ser observados nos gráficos a seguir.

Os primeiros resultados relativos à duração da vogal diante de oclusivas sonoras e surdas apontam para um aspecto interessante: a duração da vogal é menor diante de oclusivas produzidas na porção anterior do trato. Nas Figuras 10, 11 e 12, observa-se que a vogal que precede a oclusiva velar surda [k] apresenta valor médio de duração maior que quando precede as demais oclusivas. Temos aí um fator que precisa ser investigado: a influência do ponto de articulação da consoante sobre a duração da vogal. Pye (1986) realizou um estudo de produção sobre a neutralização no russo. Nesse, o autor trabalhou com as mesmas medidas que Albuquerque (2010) investiga posteriormente. Entre os achados do trabalho, verificou-se que a duração da vogal é menos afetada diante das consoantes alveolares [t,d].

Também é possível notar uma diferença de percentual de duração para os três ambientes nos quais a palavra-chave ocorre. A duração é maior para o primeiro ambiente, no qual não há contexto adjacente à direita da palavra-chave (nas sentenças do tipo “say___”). Isso pode indicar que, quando não há a presença de um contexto adjacente, os informantes tendem a realizar uma vogal que apresente uma maior duração para ressaltar a produção. É interessante também que, apesar de possuímos contextos adjacentes distintos, (Figuras 9 e 10), nos quais se tem uma vogal e uma oclusiva (nas

sentenças do tipo “*say_again*” e “*say_promptly*”), o percentual de duração de vogal não é discrepante.

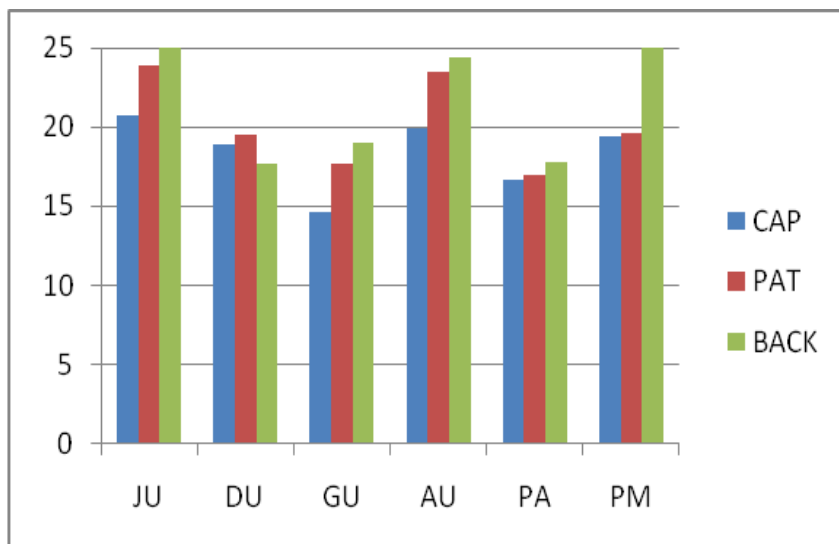


Figura 10 Gráfico do percentual de duração de vogal para o contexto 1.

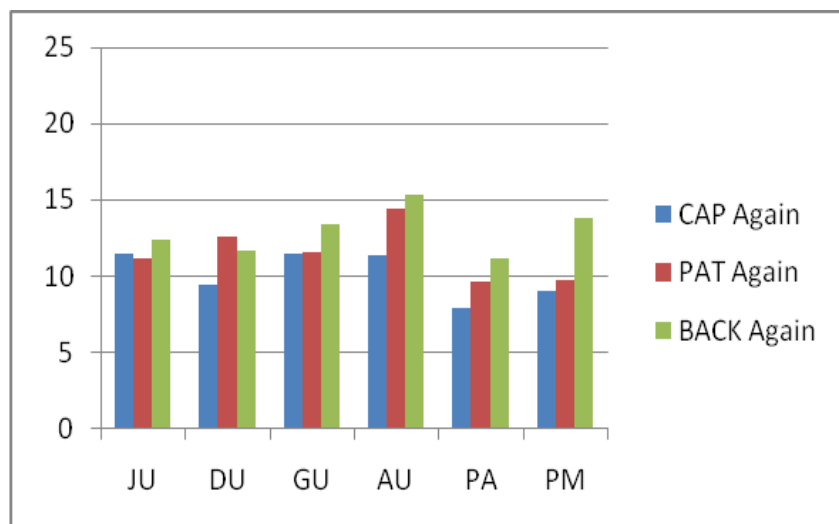


Figura 11 Gráfico do percentual de duração de vogal para o contexto 2.

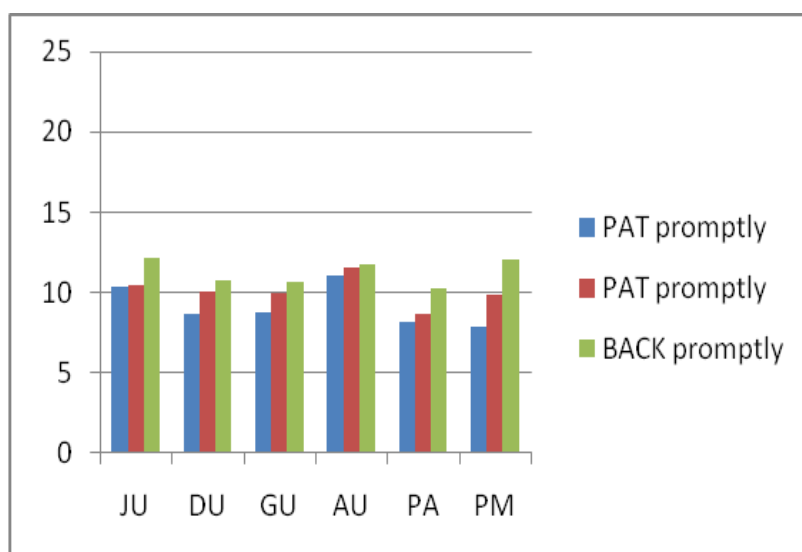


Figura 12 Gráfico do percentual de duração de vogal para o contexto 3.

Para o parâmetro vozeamento, mostrado nas Figuras 13, 14 e 15, observa-se que o seu percentual na palavra pode, também, ser influenciado pelo ponto de articulação da consoante. Embora a influência não pareça tão robusta para consoantes sonoras alveolares e velares, pois para alguns informantes, o percentual de vozeamento é o mesmo ou a diferença é muito pequena, para a bilabial sonora [b] observa-se um maior índice de vozeamento.

Em alguns pontos do gráfico, como por exemplo, em sentenças “*say bad again*”, os informantes paranaenses, mas de modo mais abrangente a nativa, realizam um *flap*, movimento que resulta na produção de um rótico e não de uma oclusiva, o que visualmente é observado por um decréscimo na taxa de vozeamento. Visualmente, quando se tem um *flap*, não há uma interrupção no sinal acústico. Esse tipo de fenômeno ocorre – denominado “*flapping*” na literatura relativa à fonologia do inglês (cf., e.g., Fox e Terbeek, 1977) – pois o rótico e as oclusivas [t] e [d] são articulados no mesmo ponto. A ocorrência do *flapping* suscita uma questão relacionada à maneira como tal fato se reflete nos resultados finais de uma análise.

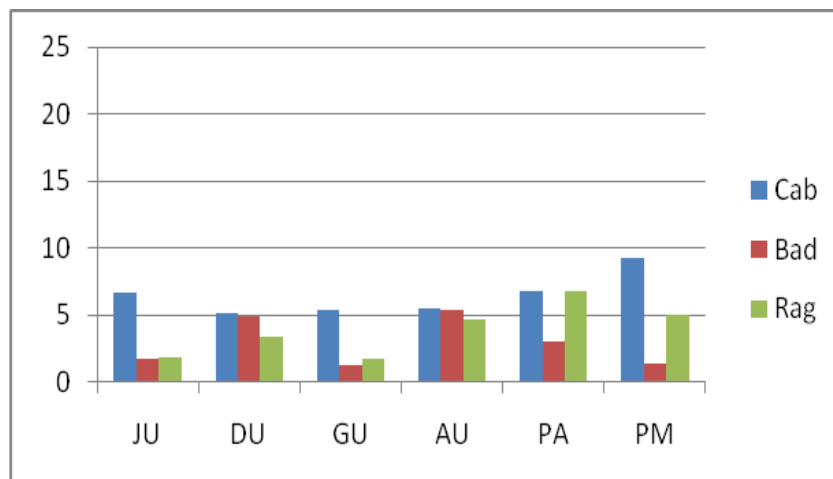


Figura 13 Gráfico do percentual de vozeamento para o contexto 1.

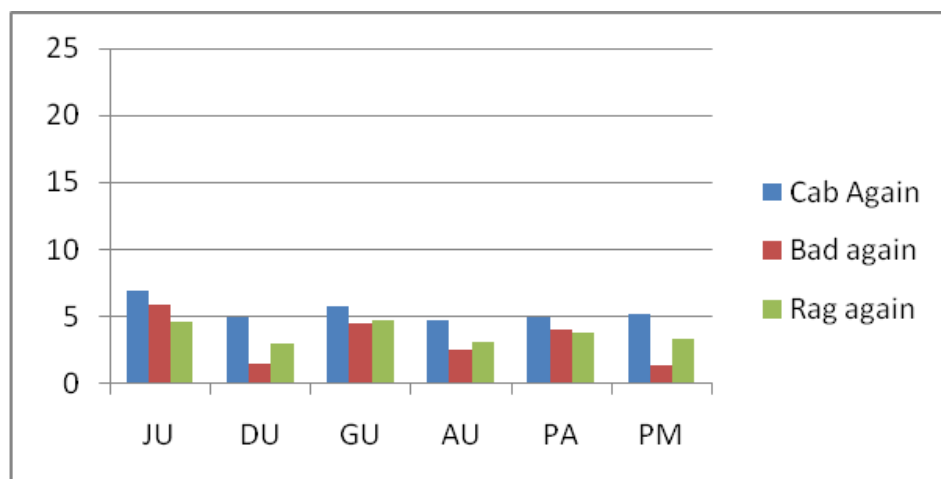


Figura 14 Gráfico do percentual de vozeamento para o contexto 2.

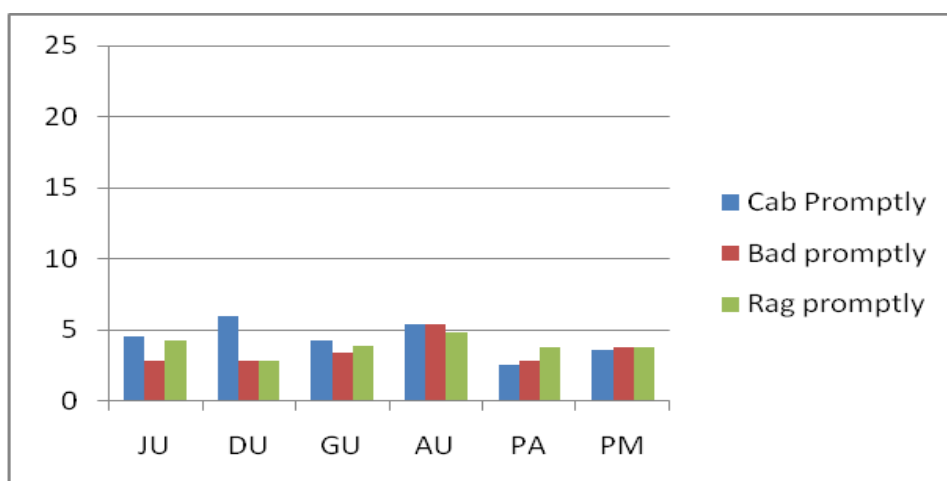


Figura 15 Gráfico do percentual de vozeamento para o contexto 3.

Além do que foi mencionado no parágrafo anterior, os contextos adjacentes influenciam no vozeamento: segundo relatos da literatura (ZIMMER e ALVES, 2007), quando o contexto adjacente é uma vogal, o segmento posterior tende a sonorizar³⁵. Além de se observar este fato, verificou-se que em outros contextos, mesmo diante de oclusiva, como é o caso de “*promptly*”, os informantes continuam sonorizando.

Com isso, pode-se notar que o experimento conduzido vai em direção aos achados de Zimmer e Alves (2007 e 2008), e concorda com o modelo da Fonologia Gestual de Browman e Goldstein. Contudo, como já foi levantado em um momento anterior, embora mesmo os modelos gerativistas mais recentes, como o modelo apresentado por Oostendorp, assumam a gradiência presente no *output*, como tal gradiência é percebida ainda permanece em discussão. Para tanto, antes de se proceder ao relato do experimento perceptual, foco deste trabalho, será feito um apanhado dos estudos de percepção que tomaram a DT como fenômeno de análise.

³⁵ Trata-se do mesmo efeito coarticulatório que faz com que diante de silêncio um som sonoro ensurdeça.

5 EXPERIMENTOS DE PERCEPÇÃO DE FALA EM L2

No início dos estudos fonológicos, muitos trabalhos se debruçaram sobre o desempenho dos aprendizes de L2. Como menciona Escudero (2005), Polivanov (1931, *apud* Escudero, 2005) relatou, em diversos estudos, a manifestação de fonemas da L1 na L2. De acordo com Escudero, esses achados poderiam indicar que algumas das dificuldades encontradas na produção da L2 seriam devidas à influência da percepção da L1 sobre a L2. A autora também faz referência aos trabalhos de Trubetzkoy (1939/1969, *apud* ESCUDERO, 2005), os quais indicaram a possibilidade de que “erros” de produção na L2 tinham influência da percepção, uma vez que, para Trubetzkoy, o sistema fonológico da L1 desempenha o papel de “filtro” e, assim, os sons da L2 são percebidos através desse filtro. O que podemos entender, de acordo com esse cenário teórico exposto por Escudero, é que as inadequações da produção de sons na L2 podem implicar a não percepção acurada desses sons. Em decorrência desses achados, observa-se que, como nos estudos de Llisterri (1995) e Leather (1999, *apud* ESCUDERO, 2005), a percepção se daria antes da produção, pois, assim como foi dito acima, uma vez que não se percebe, com acuidade, o som de um segmento da L2, a produção acaba por ficar comprometida. Um mau desempenho do aprendiz poderia estar associado, também, portanto, como Port e Leary (2005) apontam, ao nosso conhecimento estar pautado em um conjunto de símbolos alfabéticos. Nesse sentido, quando ouvimos algo como [rag] ou [ra], mas sabemos que [ra] não se configura como uma palavra possível numa dada língua-alvo, no nosso caso o inglês, o falante preenche tal produção com uma sequência sonora compatível, muitas vezes, com o padrão ortográfico da L2. É preciso notar, também, que produção e percepção, como menciona Port (1996), podem estar conectadas à ideia de expectativa vs. realização. De acordo com o autor, é o que a literatura costuma chamar de “*responses bias*”. Port relata que, à medida que os informantes são apresentados a um determinado conjunto de dados e, se estes parecem ter uma complexidade alta, os informantes ativarão outro processo

para fazer a distinção ou equiparação. Uma das questões extremamente interessantes que surgem a partir disso é que a decisão dos sujeitos por uma das produções pode ser afetada por diversos fatores, inclusive a expectativa destes antes mesmo de escutarem o conjunto de dados.

Nessa direção, os trabalhos que procuram dar um tratamento à DT do ponto de vista da percepção, acabam por endossar essa discussão maior sobre perceber categorias ou a gradiência entre os sons produzidos e, também, levantar questionamentos sobre a necessidade de se manter uma divisão entre os níveis fonético e fonológico.

5.1 Modelos de percepção de fala em L2

Nesta seção, serão apresentadas duas propostas de análise de percepção de fala na aquisição de L2: o *Speech Learning Model* (SLM; Flege, 1995) e o *Perceptual Assimilation Model* (PAM-L2; Best e Tyler, 2007). Ambos os modelos procuram analisar como se dão as diferenças na habilidade dos aprendizes de perceberem os contrastes presentes no inventário consonantal e vocálico da L2 com relação à base da L1 que possuem. O modelo PAM-L2 se baseia nas premissas do PAM (Best, 1995) para propor um tratamento de dados relativos à aquisição de L2. Nesse sentido, será feita uma apresentação do PAM para que depois se discuta o PAM-L2.

Antes de descrever seu modelo de produção e percepção da fala, Flege (1995) resenha alguns trabalhos que se debruçaram sobre a descrição e análise da existência do sotaque em diversas línguas. Entre essas pesquisas, se encontravam dados de que os falantes-aprendizes de L2 pareciam possuir uma dificuldade em seus sistemas motores, o que fazia com que as suas produções fossem diferentes daquelas dos falantes nativos (PENFIELD, 1965; LENNEBERG, 1967, *apud* FLEGE, 1995). Outros autores analisaram a presença de sotaque como um problema (pelo menos até certa extensão) gerado por uma “falha” na percepção dos falantes não nativos (FLEGE, 1992 a, b). Além disso, outros autores atribuem o sotaque ao fato dos falantes não-nativos não obterem um *input* adequado, uma vez que o *input* é obtido a partir

do professor da turma (que por vezes não é um falante nativo) e dos colegas. Questões motivacionais e psicológicas (alguns falantes não querem aprender a língua) podem, também, se configurar como fatores para a presença de produções distintas dos falantes nativos (FLEGE, 1988b).

A partir desses relatos, Flege (1995) descreve um experimento realizado por ele em 1981. Nesse, o autor notou que: a) se o som da L2 é identificado com um som da L1, o som da L2 será substituído pelo da L1 se a L1 e a L2 forem foneticamente distintas; b) contrastes que existem na L2, mas não na L1 não serão reconhecidos pelos aprendizes; c) contrastes que existem na L1 podem ser produzidos na L2 mesmo que esses não existam na L2³⁶.

Investigando as motivações para a existência do sotaque na L2, Flege (1995) descarta a possibilidade de que a sua existência tenha se dado a partir de uma “falha motora”, uma vez que vários experimentos comprovam que pessoas são capazes de produzir sons que não sabiam antes, obtendo produções adequadas. Segundo o autor, a relação entre “erros” de produção e “erros” de percepção ainda precisa ser investigada com mais cuidado, uma vez que os resultados para L2 são diferentes dos para L1. Alguns autores como Locke (1980) mencionam que os erros de produção possuem uma base articulatória e não mentalista. Como o sotaque parece não ter como base deslizos do aparato articulatório, isso significa que os padrões específicos de língua são passíveis, até certo ponto, de serem modificados.

Assim, Flege (1995) desenvolve um modelo que se baseia em 4 postulados e 7 hipóteses que têm como premissa que a produção de sons da L2 não será adequada se não houver alvos perceptuais acurados, que guiem o aprendizado sensório-motor do indivíduo. Por outro lado, apesar da importância do dado perceptual, não se pode esquecer que o modelo não entende que todos os erros de produção em L2 são perceptualmente motivados. O modelo prevê:

³⁶ “[...] 1) an L2 sound that is "identified" with a sound in the L1 will be replaced by the L1 sound, even if the L1 and L2 sounds differ phonetically; 2) contrasts between sounds in the L2 that do not exist in the L1 will not be honored; and 3) contrasts in the L1 that are not found in the L2 may nevertheless be produced in the L2” (Flege, 1995:235)

Hipótese 1 (H1): os aprendizes relacionam perceptualmente alofones posicionais concretos da L2 com alofones similares perceptualmente na L1.

Hipótese 2 (H2): se o aprendiz realizar uma diferença fonética entre um som da L2 e da L1, uma nova categoria fonética pode ser estabelecida para o som da L2.

Hipótese 3 (H3): a probabilidade de que o aprendiz torne-se consciente de tais diferenças aumenta na mesma proporção em que aumenta a diferença fonética entre o som da L2 e o som mais próximo na L1.

Hipótese 4 (H4): conforme aumenta a idade em que se dá a aprendizagem, os aprendizes são menos capazes de detectar essas diferenças fonéticas.

Hipótese 5 (H5): a formação de categoria pode ser impedida pelo mecanismo de classificação equivalente.

Hipótese 6 (H6), também é possível que a nova categoria estabelecida pelo aprendiz de L2 seja diferente da categoria equivalente do falante nativo monolíngue. Isso ocorre porque a categoria do falante da L2 pode ser desviada de uma categoria próxima da L1 para manter contraste no espaço fonológico comum da L1 e L2.

Hipótese 7 (H7): estabelece a conexão entre a representação categórica e a produção, afirmando que a produção de um som no curso da aquisição corresponde, eventualmente, às propriedades de sua representação de categoria fonética³⁷. (Flege, 1995: 239 adaptação minha)

A partir dos pressupostos do modelo, entende-se que mesmo que um dado sistema possua categorias estabelecidas para a L2, diferentes das já

³⁷ No original: "HI Sounds in the L1 and L2 are related perceptually to one another at a position sensitive allophonic level, rather than at a more abstract phonemic level."

H2 A new phonetic category can be established for an L2 sound that differs phonetically from the closest sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.

H3 The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.

H4 The likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are non-contrastive in the L1, being discerned decreases as AOL increases.

H5 Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production.

H6 The phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's.

H7 The production of a sound eventually corresponds to the properties represented in its phonetic category representation.

existentes para a L1, o segmento produzido nem sempre será realizado como seria esperado, de maneira semelhante à de um falante nativo. O SLM não prevê que os falantes-aprendizes irão produzir os sons da L2 como falantes nativos. O modelo se assemelha, conforme menciona Flege (1995), à visão de bilinguismo proposta por Grosjean (1989, *apud* Flege, 1995), na qual o autor sugere que não há como bloquear a mistura entre a L1 e a L2, uma vez que ambos os sistemas se encontram interligados. Segundo Flege (*op cit.*), isso implica em compreender que falantes que estão adquirindo uma L2 não possuem uma “chave” que liga e desliga o sistema de uma língua quando a outra está sendo utilizada. Tal observação vai de encontro a fatos como o *code-switching*, o qual foi utilizado para explicar fatos de bilinguismo, i.e., por que numa determinada situação um falante bilíngue utiliza uma língua e, noutra situação, usa a outra.

Outro fato que decorre do modelo reside na identificação de que existe uma suscetibilidade dos aprendizes realizarem produções mais acuradas quando adquirem a língua enquanto crianças do que quando adultos. Tal achado pôde ser confirmado a partir de trabalhos como o de Flege (1991b), que estudou o modo como os falantes de espanhol como L1 produziam as oclusivas do inglês. Os informantes que aprenderam inglês enquanto adultos apresentaram valores de VOT muito distantes das produções dos nativos; já os que adquiriram o inglês ainda crianças possuíam valores de VOT mais próximos às realizações dos nativos.

Dados de experimentos perceptuais (WARNER *et al.*, 2004; BROERSMA, 2005), mais tarde, irão de encontro à hipótese H1 do SLM, que prevê que alofones sensíveis à posição que se encontram na palavra (inicial, mediana ou final) estariam perceptualmente relacionados um ao outro. Nesse sentido, seria esperado, segundo Flege, que um dado falante cuja L1 não possui oclusivas em posição final não fosse capaz de relacioná-las com as oclusivas em posição inicial ou mediana existentes em sua L1. Isso se explicaria uma vez que as estruturas fonéticas da L1 não poderiam causar interferência na criação de novas categorias.

Flege tece, também, algumas considerações sobre como seu modelo trata a questão de “substituições” de segmentos da L2. De acordo com abordagens que tratam, entre outras questões, do processo de aquisição de uma língua, como é o caso da análise contrastiva, uma vez que o aprendiz precisa aprender um inventário de sons da L2 que não existem na L1 isso se apresenta como um problema de aprendizado (e.g., LADO, 1957; MOULTON, 1962). Nesses casos, o aprendiz, segundo Flege (1995) e outros autores (e.g., LEHISTE, 1988), procura substituir o som que desconhece da L2 pelo correspondente mais próximo da L1. No caso do inventário de sons, é comum que um segmento da L2 seja relacionado perceptualmente com um da L1. No entanto, os sons da L2 são associados na representação do falante-aprendiz de modo a serem categorizados como novos ou diferentes dos existentes na L1 (MUELLER e NIEDZIELSKI, 1963, *apud* Flege, 1995; BRIERE, 1966, *apud* Flege, 1995; HENNING, 1966, *apud* Flege, 1995; FLEGE, 1981, *apud* Flege, 1995; WODE, 1978, *apud* Flege, 1995). O SLM prevê que os sons da L2 podem vir a ser identificados de acordo com a sua posição de alofone na L1. O aprendiz passa por um processo de ganhar experiência como falante da L2-alvo e, com isso, pode, gradualmente, aprender a distinguir determinados sons da L2 e os sons da L1 que sejam próximos. Uma vez que o aprendiz é capaz de realizar tal distinção, o novo som da L2 passa a ser categorizado na representação do sujeito (Hipótese 2 do SLM). De acordo com o modelo, quanto mais cedo o falante-aprendiz começa a aprender a L2, menor é a distância fonética para que se inicie o processo de formação de categorias. Por outro lado, mesmo que o indivíduo comece o aprendizado ainda criança, sons da L2 que forem semelhantes aos da L1 serão difíceis de serem contrastados (Hipótese 4).

Além de resenhar experimentos cujos resultados parecem corroborar algumas hipóteses do modelo SLM, Flege, Munro e MacKay (1995) conduziram um experimento que tinha como objetivo verificar a relação entre a idade de aprendizado e a percepção do sotaque de 240 falantes nativos de italiano, aprendizes de inglês como L2. Tais sujeitos iniciaram seu aprendizado de

inglês entre 2 e 23 anos de idade e haviam morado por 32 anos no Canadá na época em que os testes foram realizados. Como já havia sido relatado para experimentos com falantes nativos de outras línguas, os resultados mostraram que o nível de produções mais acuradas começa a cair à medida que os indivíduos vão ficando mais velhos. Os autores apontam que além do fator idade outros aspectos como sexo, frequência de uso do inglês e tempo de residência no lugar podem ter contribuído para a diferença nas produções.

Em linhas gerais, é possível dizer que o SLM é um modelo de base fonológica, que procura analisar em que medida fatores como a idade em que o indivíduo adquiriu a língua e de que forma os sistemas de duas línguas diferentes se relacionam. Além disso, o SLM procura observar como os sons da L2-alvo são acomodados no inventário do falante de uma L2. Um dos pontos importantes do modelo reside na possibilidade de que o indivíduo é capaz de aprender os detalhes fonéticos relativos a um determinado som da L2, a ponto de que a produção do aprendiz é passível de ser modificada a partir da exposição que ele tem do segmento em questão. Isso não significa, contudo, que o falante-aprendiz irá realizar o som igual a um falante nativo, mas os valores dos parâmetros que caracterizam esse segmento serão alterados, não havendo uma discrepância entre a produção do nativo e não-nativo. Além disso, o modelo indica que os sons da L2 que existem na L1 têm uma maior possibilidade de não serem produzidos acuradamente, enquanto que os que não são encontrados no inventário da L1, por não compartilharem da mesma categoria fonológica, possuem mais chances de serem assimilados adequadamente.

Embora o trabalho de Flege (1995) ofereça uma interpretação interessante, o autor menciona que é necessária uma análise mais detida, pois embora a percepção esteja relacionada à produção, não é possível assegurar que aprendizes cujas realizações na L2 possuem valores diferentes das produções de falantes nativos não sejam capazes de perceber os segmentos da L2.

É importante esclarecer que o SLM lida com falantes que estão no seu último nível de aprendizado de uma L2, e que, portanto, já possuem uma grande exposição aos sons da L2. Por outro lado, o PAM (BEST, 1995) – o qual será discutido na sequência do texto – lida com falantes monolíngues, linguisticamente “ingênuos” com relação aos dados aos quais são expostos. Ao contrário do SLM, o PAM não se configura como um modelo explicativo sobre a percepção em L2. Apesar disso, como já foi mencionado no início da seção 5, o PAM-L2 tem como base os pressupostos teóricos do PAM.

Parte-se agora para a descrição do PAM, que embora tenha sido publicado no mesmo período que o SLM, possui princípios de base distintos.

O PAM prevê que o fato de um aprendiz conseguir distinguir os fones de outra língua, que não a sua, pode ter a influência da semelhança fonética e articulatória dos fones em ambas as línguas. Nesse sentido, o modelo dispõe que, ao entrar em contato com os sons de uma língua estrangeira, devido à exposição do sujeito-aprendiz a sua L1, este é capaz de perceber o fone não ativo em relação ao fone articulatoriamente semelhante na sua L1. De acordo com Best (1995), os contrastes perceptuais têm padrões, os quais estão relacionados com a capacidade de discriminação de sons de uma língua estrangeira. O PAM procura relacionar ambos os níveis, o fonológico e o fonético (aspecto distinto do SLM, uma vez que se trabalha com a separação dos sistemas fonológico e fonético). A questão para o modelo é tentar explicar como o sistema da L1 acomoda os fones aos quais é exposto a partir de fatos que são articulatoriamente motivados.

Um dos pontos de principal divergência entre os modelos se encontra, como foi mencionado anteriormente, nos princípios de base de ambos. O PAM está afiliado ao realismo direto, proposto por James Gibson (1966, 1979 *apud* Best, 1995) e Eleanor Gibson (1989, 1991, *apud* Best, 1995). Segundo essa teoria, o sujeito apreende o objeto perceptual de forma direta, sem nenhum tipo de mecanismo mental, nenhuma regra de transformação abstrata. De acordo com Best (1995), o estímulo que os falantes recebem é ambíguo e pobre.

Ainda de acordo com a autora, a fonologia tradicional (conforme foi discutido na seção 4, a propósito dos argumentos apresentados por Nishida, 2007) vê a implementação fonética das representações fonológicas como universal e não linguística. Diversos aspectos presentes nas línguas, como a coarticulação, são vistos como processos mecânicos, físicos. Nesse sentido, a fonologia faz parte do que é chamado de específico de língua e a fonética é, como já foi mencionado, universal e não linguístico. Contudo, estudos, como o de Port (1996) e Port e Leary (2005), discutem que a implementação fonética é específica de língua, não necessariamente universal, como preveem modelos/abordagens de base fonológica.

Algumas considerações, vindas de outros modelos, como por exemplo, da teoria motora (LIBERMAN e MATTINGLY, 1985) e psico-acústica (ASLIN, PISONI e JUSCZYK; DIEHL e KLUENDER, 1989; JUSCZYK, 1993), apontam para o fato de que o realismo direto não leva em conta o aprendizado. Segundo Best (1995), o aprendizado perceptual possui um papel muito importante no realismo direto. O modelo se preocupa com o processo de aquisição de pistas acústicas da língua materna e em como os indivíduos utilizam da experiência acumulada, da calibragem perceptual que vão realizando à medida que são expostos a dados da sua língua materna e de outras línguas para reconhecer características invariantes dos segmentos. Uma vez que o sujeito é capaz de melhorar sua percepção, ele consegue selecionar as pistas acústicas responsáveis pela caracterização de um determinado segmento e usá-las para identificar um dado som. O PAM trabalha diretamente com questões relacionadas à descoberta e sintonia (*atunement*). As informações de ordem mais alta (invariantes) são distintivas. O treinamento aumenta a eficiência da aquisição de informações do ambiente, deixando de lado informações de baixa ordem, irrelevantes para a distinção (economia de percepção) dos gestos importantes para que os sons sejam realizações existentes na língua. Os invariantes de baixa ordem são compostos de gestos genéricos do trato e os de alta ordem analisam os traços distintivos de uma dada língua.

Enquanto o SLM é um modelo de base fonológica e se utiliza dos traços distintivos para categorizar os sons da L2, o PAM tem como primitivos de base os gestos articulatórios (os quais foram descritos no início dessa seção). Assim, ao contrário da teoria motora que preconiza que os sujeitos percebem os dados de fala a partir da sua própria produção (o que implica, em última instância, em se pensar que um aprendiz não seria capaz de perceber os sons de outra língua adequadamente uma vez que possui dificuldades em produzi-los), o realismo direto (base do modelo PAM) prevê que os aprendizes percebem as informações acústicas a partir dos gestos produzidos pelos falantes, ou seja, não é necessário que eles saibam produzir tais segmentos³⁸.

Passe-se agora à discussão dos tipos de assimilações propostas pelo PAM. De acordo com o modelo, os segmentos não nativos podem ser percebidos de três modos:

- (a) como segmentos *incategorizáveis*, nos quais o som da L2 é percebido como um gesto de fala, mas a sua organização não é assimilada a nenhuma categoria específica da L1;
- (b) como *não assimilável*, quando o segmento da L2 não é percebido como um som de fala, a sua organização gestual ocupa uma área fora do espaço fonológico nativo, não sendo assimilado a qualquer categoria da L1;
- (c) como segmentos *categorizáveis*, nos quais as invariantes gestuais de ambas as línguas podem ser percebidas como similares. Neste caso, o som da L2 é assimilado numa categoria da L1. (Reis, M. 2011: 91)

De acordo com o exposto por Reis (2011), a partir do tratamento dado à acomodação dos sons da L2 por Best (1994a, 1995b), é possível mencionar que os pressupostos de base procuram analisar a relação entre sons da L1 e L2. Tal ligação entre sons nativos e não nativos pode se dar de modo que haja uma necessidade dos sons da L2, para serem assimilados, estabelecerem relações de similaridade com os da L1. Caso contrário, se os da L2 não

³⁸ Para mais informações sobre essa consideração, ver Fowler *et al.* (1990).

compartilham de gestos semelhantes ou existentes na L1, um dado som pode ser considerado incategorizável ou ser tido como um som que não pertence à fala. Nesse sentido, existe uma semelhança entre o SLM e o PAM: ambos desenvolveram modelos pautados na relação L1-L2.

Dentre os três padrões de assimilação existentes, o PAM apresenta 6 possibilidades de assimilação. É importante mencionar que não será feita uma discussão detalhada dos 6 pontos, sendo que apenas alguns serão analisados com o objetivo de relacioná-los, posteriormente, ao experimento de percepção a ser descrito mais tarde (capítulo 7). Abaixo, encontram-se as possibilidades propostas pelo modelo PAM:

- (*TC Type – Two-Category Assimilation*): Cada segmento não-nativo, de um par que oferece contraste, é assimilado a uma categoria diferente da L1, e a discriminação é excelente. Como exemplo disso tem-se o par [T] e [tʰ], do inglês (L2), que seria assimilado no PB (L1) como /s/ e /t/ respectivamente.
- (*CG Type – Category Goodness Difference*): Ambos os sons não-nativos, da L2, são assimilados pela mesma categoria da L1, mas eles diferem de modo discrepante com relação ao som nativo “ideal” (e.g., um som é aceitável, o outro é desviante). Espera-se que a discriminação seja entre moderada e boa.
- (*SC Type – Single-Category Assimilation*): Ambos os sons não-nativos (L2) são assimilados à mesma categoria nativa (L1), mas são igualmente discrepantes com relação ao som nativo “ideal”, ou seja, ou ambos são igualmente aceitáveis ou desviantes. A discriminação, nesse caso, é esperada que seja pobre.
- (*UU Type – Both Uncategorizable*): Ambos os sons não-nativos (L2) se encontram no espaço fonológico, mas não possuem correspondência com qualquer categoria da L1. A distinção é de difícil a muito boa.
- (*UC Type – Uncategorizable versus Categorized*): Ocorre quando um som é assimilado a uma categoria nativa (L1) e o outro não. A discriminação é prevista como boa.
- (*NA Type – Nonassimiable*): Ambas as categorias não-nativas recaem fora do domínio dos sons ouvidos como sons que não são

próprios da fala humana. Espera-se que a discriminação seja de boa a ótima. (Best, C., 1995:192)

Embora ambos os modelos possuam metodologias experimentais sólidas, é preciso lembrar que por trás das hipóteses do SLM ou dos padrões de assimilação do PAM, os dois modelos seguem visões distintas no modo como os dados de L1 e L2 interagem. O SLM é um modelo que mantém a divisão clássica entre fonética-fonologia; já o PAM, por ter como base o realismo direto e a Fonologia-Articulatória de Browman e Goldstein, está pautado em gestos como primitivos de análise e em uma integração do conhecimento fonológico e fonético.

Tanto o SLM quanto o PAM não se limitam a fazer previsões pautadas somente na influência dos contrastes fonológicos na L1. Ambos os modelos advogam a importância de semelhanças e diferenças fonéticas de elementos contrastivos e não entre fones da L1 e não nativos. Contudo, o PAM não está relacionado, apenas, às interferências fonológicas. Ele caracteriza a influência da L1 como sendo parte do aprendizado perceptual de padrões fonético-articulatórios em ambos os níveis: no nível abstrato e no detalhe fonético. Best e Tyler (2007) fazem uma análise dos postulados dispostos no SLM a partir do embasamento teórico do PAM, propondo um PAM-L2. Abaixo, listaremos os postulados e os comentários feitos pelos autores (BEST e TYLER, op cit).

P1: Os mecanismos e processos utilizados no aprendizado do sistema da L1, incluindo a formação de categorias, permanecem intactos ao longo da vida, e podem ser aplicados ao aprendizado da L2³⁹.

³⁹ No original: "P1: The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.

P2: Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.

P3: Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.

P4: Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.

Com relação ao primeiro postulado, Best e Tyler (2007) mencionam que o SLM e o PAM são semelhantes no sentido de acreditar que as habilidades que envolvem o aprendizado perceptual são as mesmas para crianças, que estejam aprendendo uma L1 ou L2, e adultos, aprendendo uma L2. No entanto, os modelos divergem no entendimento sobre a abstração da informação. Para o PAM, os informantes apreendem os invariantes linguísticos sobre os gestos articulatórios utilizados a partir do sinal de fala (Best e Tyler, 2007:20); o SLM entende que os aprendizes criam categorias a partir de pistas acústico-fonéticas.

P2: Aspectos dos sons da fala, específicos de língua, chamados de categorias fonéticas, são especificados nas representações da memória de longo prazo.

O PAM segue na direção oposta do segundo postulado do SLM. Enquanto o SLM pressupõe a existência de uma representação mental para o que é invariante na língua, no PAM, por outro lado, “[...] o informante percebe, de modo direto, os gestos articulatórios do falante e, através do aprendizado perceptual, passa a detectar os invariantes articulatórios de ordem mais alta no sinal de fala” (BEST e TYLER, 2007:20). Ainda de acordo com os autores, Gibson e Gibson (1955) mencionam que os falantes mais experientes em uma dada língua não passam por um processo de criação de categorias em suas memórias de longo prazo, mas desenvolvem suas habilidades perceptuais de modo a se tornarem mais “afinados” para que possam reconhecer invariantes em contextos diversos. Assim, para a teoria base do PAM, os falantes não precisam ter representações mentais das categorias fonéticas para que sejam capazes de perceber fones da L2.

P3: Categorias fonéticas estabelecidas na infância para os sons da L1 se desenvolvem ao longo da vida, para que reflitam as propriedades de todos os fones da L1 e L2 identificados como a realização de cada categoria.

Este postulado é compatível tanto com o SLM quanto com o PAM, pois em ambos os informantes continuam a refinar a sua percepção ao longo de suas vidas. Por exemplo, de modo semelhante ao SLM, no PAM, quando os informantes se deparam com um som para o qual não conseguem estabelecer a mesma “categoria” de um som já existente na L1, eles têm que encontrar outra maneira de estabelecer os invariantes para que possam reconhecer a configuração articulatória do novo fone. Contudo, Best e Tyler (2007) apontam para uma diferença no tratamento dado a duas questões fundamentais: “Como os ouvintes identificam fones não nativos como sendo equivalentes aos fones da L1” e “o nível ou níveis em que isso ocorre.” (BEST e TYLER, 2007: 21-22). O tratamento para essas questões está diretamente relacionado com o modo como modelos semelhantes ao SLM encaram o processo de percepção dos fones da L2: passiva ou ativamente. Segundo os autores, modelos como o SLM entendem que os informantes recebem, de forma passiva, estímulos proximais. Esses estímulos seriam correspondentes à parte acústica (os traços). Depois de receber o estímulo acústico, o falante computaria a distribuição dos fones no *input*. Por outro lado, modelos como o PAM, afiliados ao realismo direto, compreendem que os ouvintes buscam a informação distal no sinal acústico. A informação distal estaria relacionada aos gestos articulatórios e o modo como eles se organizam.

P4: Os falantes bilíngues lutam para manter o contraste entre as categorias fonéticas da L1 e L2, as quais compartilham do mesmo espaço fonológico.

Neste postulado, o PAM converge com relação ao SLM no sentido de que as categorias fonológicas de L1 e L2 podem existir no mesmo espaço. Contudo, o PAM se opõe ao SLM, pois prevê que os níveis fonético e fonológico interagem no aprendizado da L2 e que “isso depende crucialmente da relação entre os espaços fonológicos da L1 e L2” (BEST e TYLER,

2007:22). De acordo com os autores, se uma categoria fonológica da L2 foi assimilada, conforme o entendimento do PAM, ou é equivalente, – assim como menciona Flege (1995) nas hipóteses do SLM, a um fonema da L1, mas, foneticamente estes fones, da L1 e L2 forem distintos, i.e., sejam passíveis de serem discriminados – o PAM prevê que existe a possibilidade de existir uma única categoria fonológica para duas realizações fonéticas distintas.

A partir das distinções feitas por Best e Tyler (2007) entre o SLM e o PAM, os autores procuram estabelecer um PAM-L2, o qual também teria como base o realismo direto, mas teceria considerações sobre a percepção em L2. Os autores (BEST e TYLER, 2007) mencionam que o PAM-L2 não trabalha com relações de equivalência apenas no nível fonético (como o SLM), mas também no fonológico. Segundo eles, é necessário que exista uma equivalência num nível léxico-funcional. Isso significa que as categorias fonológicas estariam numa relação com as categorias ao redor, sendo que essa relação pode ser de contraste ou semelhança. Apesar de uma categoria fonológica se encontrar em uma relação de semelhança, isso não significa que relações de semelhança sejam percebidas no nível fonético. Tendo isso como base, Best e Tyler (*op cit*) elaboram 4 casos sobre uma possível interpretação de contrastes da L2 que os aprendizes de L2 percebem como sendo de fala.

1. *Apenas uma categoria fonológica da L2 é percebida como equivalente (percentualmente assimilada) a uma categoria fonológica da L1*⁴⁰.

⁴⁰ No original: 1) Only one L2 phonological category is perceived as equivalent (perceptually assimilated) to a given L1 phonological category.

2) Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but one is perceived as being more deviant than the other.

3) Both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but as equally good or poor instances of that category.

4) No L1-L2 phonological assimilation.

Neste tipo de situação, a previsão do PAM-L2 é de que o aprendiz percebeu ambas as categorias, fonética e fonológica, como equivalentes. Com relação ao aprendizado do fone da L2, o modelo prevê que mesmo que o aluno procure aprender tal distinção, é difícil que o fone mude de categoria. Segundo os autores, outra interpretação possível é a de que o fone seria identificado como desviante na categoria fonética. Apesar dos sons serem parecidos, eles são foneticamente distintos.

2. Ambas as categorias fonológicas da L2 são percebidas como equivalentes à mesma categoria fonológica da L1, mas uma é percebida como desviante em relação à outra.

Se comparado ao PAM, esse caso seria identificado como *Category Goodness*. É esperado, então, que os aprendizes distingam os fones de modo satisfatório. O fone considerado desviante será assimilado a uma nova categoria fonética e fonológica e o fone que tenha sido mais bem percebido, seria assimilado à mesma categoria fonética e fonológica da L1. Os autores não tecem, contudo, nenhum modelamento para como o fone considerado desviante se desenvolveria, seria aprendido. Uma hipótese, segundo eles, seria de que o fone desviante seria, primeiramente, aprendido como uma variante fonética do fone existente na L2, mas estando associado à categoria fonológica da L1. Com o tempo, o aluno poderia refinar os contrastes léxico-funcionais, como exposto pelos autores, passando a desenvolver uma nova categoria fonológica para o fone desviante.

3. Ambas as categorias da L2 são percebidas como equivalentes a uma mesma categoria existente na L1, sendo instâncias boas ou ruins dessa categoria.

Neste caso, segundo Best e Tyler (2007), os informantes terão problemas para identificar contrastes que sejam necessários na L2, uma vez que os dois fones da L2 serão assimilados fonética e fonologicamente a uma

única categoria fonológica da L1. Assim, numa situação em que o aluno tenha que discriminar dois sons que oferecem contraste na L2, este pode vir a considerar as palavras como homófonas. Tomando como base o SLM, poderia-se dizer que os fones da L2 se tornariam uma única categoria fonética da L1.

4. Não há nenhuma assimilação fonológica entre L1 e L2.

Uma vez que o falante não percebe o contraste entre os fones da L2 nem como pertencentes a uma única categoria da L1, mas como se os fones possuíssem uma configuração de características que se assemelhassem a mais de uma categoria fonológica da L1 (o que no PAM seria tido como Não-Categorizável), uma ou as duas categorias fonológicas novas será “relativamente fácil de ser percebida” (BEST e TYLER, 2007:28). A partir disso, os autores mencionam que este caso poderia ser comparado ao que é postulado no SLM como “nova” categoria. No entanto, os autores elucidam que existe uma diferença conceitual entre o caso 4 e o que é previsto pelo SLM como a criação de outra categoria. Para o PAM-L2, a não categorização de um fone não está relacionada somente com as semelhanças ou diferenças do fone pertencente à L2 e a categoria fonética da L1, existe uma preocupação de se levar em conta as relações existentes no sistema da interlíngua do sujeito. Não se trata de uma relação estritamente entre a assimilação de um determinado fone, mas do que já foi estabelecido no sistema fonológico do sujeito para fones semelhantes ao que se está analisando no momento.

A partir do que foi exposto sobre o PAM-L2, é possível pensar que tal modelamento leva em conta a natureza dinâmica da aquisição de percepção em L2, pois não só fatores como a exposição do falante à L2 são levados em consideração. O PAM-L2 entende que tanto o ambiente que o falante se encontra como a experiência que o indivíduo tem a língua se encontra em constante mudança.

Em suma, ambos os modelos, o de Flege (1995), e o de Best e Tyler (2007) lançaram questionamentos importantes sobre a relação entre conhecimento de L1 e L2, como as novas entradas são acomodadas no

sistema da L1 e como acontece a identificação de dados de L2. Ao final, no capítulo 8, se discutirá a possibilidade de um dos dois modelos embasar os achados dos resultados do experimento de percepção descrito no capítulo 7.

5.2 A DT EM EXPERIMENTOS DE PERCEPÇÃO: INVESTIGAÇÃO DA INFLUÊNCIA DE PARÂMETROS ACÚSTICOS E CONTEXTOS ADJACENTES

Voltando para os dados de produção apresentados pela literatura, para línguas como o Polonês, Holandês, Alemão e outros casos que foram aqui citados, autores como Kleber, John e Harrington (2010) procuram checar a existência da neutralização incompleta. Nesse trabalho, os autores se propõem, dentre outros objetivos específicos, a observar como se dá a percepção de oclusivas alveolares e velares em posição final no alemão, no sentido de tentar rastrear quais seriam as pistas acústicas utilizadas para fazer tal distinção. Antes de passar para os resultados desse experimento, é importante ressaltar alguns questionamentos feitos pelos autores sobre as pistas acústicas utilizadas em outros experimentos e que podem influenciar a percepção da neutralização incompleta.

Kleber *et al.* (*op cit.*) mencionam alguns fatos relatados pela literatura que já foram aqui mencionados, como por exemplo, a duração da vogal que precede um segmento vozeado ser maior do que para um segmento surdo⁴¹ (HAWKINS e NGUYEN, 2004). Pistas como a duração da vogal são utilizadas pelos falantes na tarefa de reconhecimento de segmentos (DAVIS, MARSLEN-WILSON e GASKELL, 2002; HAWKINS e NGUYEN, 2003; MANUEL, 1995). Os trabalhos que se debruçaram sobre o alemão relataram o mesmo (O'DELL e PORT, 1983; PORT e O'DELL, 1985).

⁴¹ Conforme já havíamos mencionado anteriormente, esse fato é controverso na literatura fonética e fonológica. Keating (1985) argumenta que esse fato não pode ser considerado universal e sim, específico de língua.

De acordo com os autores,

Dentre todas os correlatos acústicos medidos nesses estudos (e.g. vozeamento dentro e durante a oclusão), a duração da vogal se destacou como a pista mais importante para preservar o contraste de vozeamento em uma certa medida, até no Polonês (Slowiaczek e Dinnsen, 1985), no qual a duração da vogal não é a pista primária para diferenciar oclusivas vozeadas de não vozeadas em posição mediana da palavra (Keating, 1979). (Kleber, John e Harrington, 2010:186⁴², tradução minha)

A partir do que foi mencionado pelo estudo acima, apesar da duração da vogal ser considerada uma pista robusta para a identificação de sonoridade nas oclusivas finais, esta pode não se configurar como pista primária quando a tarefa é discriminar as mesmas consoantes em uma posição distinta da palavra. Existem outros parâmetros, que a depender da língua materna, serão mais robustos na distinção, como por exemplo, o *burst* ou a duração da oclusão como se verá ao final deste capítulo.

Com relação à observação acima, Kleber *et al.* (2010) notam que, para o alemão, embora a combinação entre uma maior duração da vogal e a oclusiva ser sonora exista, tal fato possui uma distribuição heterogênea. É possível destacar, por outro lado, que, diante das oclusivas surdas, na maioria dos casos, a vogal possui uma duração menor, quando comparada com as vogais que antecedem as oclusivas sonoras⁴³. Também é possível dizer, a partir do relato dos autores, que as vogais longas (em oposição às vogais curtas), em sua maioria, precedem oclusivas labiais e velares. Esse argumento vai de encontro aos achados do experimento conduzido por Albuquerque (2010), no qual a duração das vogais parecia ter alguma relação com o ponto de articulação. Embora ambos os estudos coincidam nesse aspecto, mais trabalhos precisam ser realizados para que se possa verificar a influência do ponto de articulação da consoante sobre a duração da vogal.

⁴³ Os autores mencionam que vogais curtas podem ser encontradas diante de oclusivas sonoras principalmente em casos de empréstimos (Kleber, F., John, T. e Harrington, J. 2010:186).

Com o objetivo de testar se a duração da vogal e o vozeamento são pistas que influenciam a percepção discriminatória entre oclusivas finais surdas e sonoras, a previsão dos autores (KLEBER *et al.* 2010) é a de que os indivíduos conseguem distinguir mais facilmente o contraste final entre /t/,/d/ quando as oclusivas forem precedidas de uma vogal longa, “porque o contraste sonoro só se mostra frequente no léxico para vogais longas seguidas de oclusivas alveolares, como por exemplo */laidən/* = sofrer vs. */lartən/* = liderar (KLEBER *et al.* 2010:187, tradução minha).⁴⁴ É preciso notar, no entanto, que os pesquisadores querem investigar se o mesmo tipo de distinção fonológica no alemão, entre vogais curtas e longas, e oclusivas surdas e sonoras em posição intervocálica, se mantém e pode ser percebida em posição final.

Os achados de Kleber *et al.*(*op cit.*) apontam na direção de que os informantes foram capazes de realizar a discriminação entre consoantes vozeadas e surdas em um contexto de neutralização e que, além disso, a discriminação com relação aos estímulos das velares mudou de forma gradiente, ao invés de modo categórico. Ao mesmo tempo em que esses resultados corroboram a percepção da neutralização incompleta, os dados ainda não se mostram robustos o suficiente para que se possa dizer que não há falhas na tarefa de discriminação entre as oclusivas.

Esse estudo lança, ainda, outras considerações importantes para o tratamento da DT. Como mencionam os autores, a frequência em que uma dada sequência aparece no léxico afeta a identificação dos sons. No estudo, as vogais curtas, mais do que as longas, se mostraram como uma pista importante na identificação da consoante que viria em seguida e demonstra que os falantes lançaram mão de seu conhecimento fonotático da língua.

Resultados semelhantes já haviam sido encontrados para o holandês (WARNER *et al.*, 2004). Nesse estudo, os autores investigaram, entre outros aspectos, a influência da duração da vogal e da duração da oclusão (ambos os

⁴⁴ Trecho original: “because the voicing contrast is only frequent in the lexicon for tense vowel plus alveolar stop sequences (e.g. */laidən/* = ‘to suffer’ vs. */lartən/* = ‘to lead’).”

parâmetros tiveram seus valores manipulados) na percepção do contraste sonoro da oclusiva final. A condução do trabalho se deu em etapas e os dados obtidos passaram por uma análise estatística. Ao todo, os autores realizaram cinco experimentos, sendo que os dados e alguns informantes do primeiro experimento (um estudo de produção sobre a DT no holandês) foram usados nos demais.

No primeiro experimento, os autores relatam que a influência da duração da vogal é menor do que para outras línguas. O estudo identificou que, apesar da duração da vogal ser menor para o holandês, 3.5 ms, e para línguas como o alemão e o polonês, de 10-15 ms (PORT e O'DELL, 1985; SLOWIACZEK e DINNSEN, 1985 *apud* WARNER *et al.*, 2004), ela causa um efeito na percepção da sonoridade das consoantes alveolares. No segundo experimento, agora de percepção dos dados do primeiro, a pesquisa apontou outro detalhe interessante. Informantes que produziram distinções mais acuradas para as alveolares, ao contrário do que previam Llisterri (1995) e Leather (1999), não demonstraram uma percepção mais sensível do contraste sonoro. Tal fato vai na direção oposta, portanto, de que, para perceber adequadamente os sons de uma língua, é necessário produzi-los de modo correto.

A terceira etapa do trabalho contou com manipulação da duração da vogal, para que se verificasse a influência desse parâmetro na identificação das oclusivas alveolares /t/ e /d/. Segundo os autores, dadas as condições ideais do experimento, de se isolar uma das pistas que ajudam na percepção da sonoridade final das oclusivas, criou-se um contexto distinto daquele da fala tida como natural e, sendo assim, os informantes foram capazes de usar a pista de duração de vogal na distinção. Em um contexto não controlado, talvez, os falantes não a usariam.

Além desses resultados, os autores, ainda, contribuem para a questão da natureza gradiente ou categórica da percepção da DT. Embora os informantes sejam capazes de distinguir o contraste sonoro entre o /t/ e o /d/ subjacentes em posição final, eles não estão realizando, na verdade, uma categorização linguística. Nas palavras dos autores, os “ouvintes podem ouvir

uma diferença entre o /t/ e o /d/ finais, mas não de maneira a estabelecer uma distinção categórica”⁴⁵ (WARNER *et al.*, 2004:267, tradução minha).

Como dissemos anteriormente, a duração da vogal foi um dos aspectos manipulados. No que diz respeito à duração da oclusão, a qual, no primeiro experimento, era apenas uma das pistas entre outras na discriminação da sonoridade das oclusivas, quando isolada, ela foi válida na distinção do vozeamento subjacente da oclusiva. Os informantes foram capazes de discriminar /t/ de /d/ de forma a identificar que a duração de oclusão era maior para a oclusiva alveolar surda do que para sua contraparte sonora.

Tendo em vista as contribuições do estudo, pode-se concluir que embora os informantes utilizem a duração da vogal e a da oclusão como pistas para a distinção dos pares /t/ e /d/, os aprendizes são capazes de saber qual pista está sendo isolada pelo experimento e, assim, são capazes de transferir sua aplicação para contextos diferentes (*e.g.*, se os informantes tivessem que usar as pistas para distinguir os mesmos pares em posição intervocálica e não final).

Isso nos leva, em última instância, a pensar na aquisição de L2. Nesse processo, os aprendizes precisam discriminar segmentos a partir de certas pistas que, podem ou não existir em sua L1 ou que, mesmo existindo, podem não ter a mesma relevância na distinção de sons pertencentes à L2 como tem na L1. Além disso, é possível pensar que mesmo os falantes-aprendizes que possuem mais dificuldades em adequar em suas produções parâmetros acústicos importantes para o sistema da L2 são capazes de utilizar 3.5 ms da duração da vogal (como nos dados de duração de vogal do holandês) como pista distintiva no contraste de sonoridade em contexto de final.

Nessa direção, outros estudos, como Broersma (2005), investigaram a percepção de contrastes semelhantes na L1 e L2, mas que se encontram em posições distintas das palavras, por exemplo, a distinção sonora das oclusivas

⁴⁵ “[...] listeners can hear a difference between final /t/ and /d/, but not as a categorical distinction”.

no holandês em contexto intervocálico e a mesma diferença, mas em posição final, no inglês. Embora o estudo de Broersma (*op cit.*) tenha investigado a percepção de fricativas e oclusivas em posição inicial e final, serão resenhados aqui os resultados para as oclusivas finais e para as de início de palavra. Estas serão trazidas como exemplos quando se julgar necessário.

A pesquisa, que contou com a participação de sujeitos falantes de holandês (que também possuíam conhecimento de inglês) e de inglês britânico, analisou a percepção do contraste sonoro em oclusivas em fim de palavra, utilizando logatomas⁴⁶. É importante mencionar que o experimento utilizou, em posição final, sequências com *burst* e sem *burst*, para verificar a influência desse parâmetro acústico na percepção. Entre os resultados do estudo, a autora menciona que, embora o holandês não estabeleça contraste de vozeamento em posição final, os informantes foram capazes de distinguir tal fato, tanto em posição inicial quanto em final, do mesmo modo que falantes de inglês. Isso demonstra que os falantes de holandês foram capazes de realizar uma discriminação existente em sua L1 (contraste sonoro em oclusivas) em uma posição da palavra que não oferece contraste na língua materna. Tal fato demonstra, assim como no estudo de Warner *et al.*(2004), que os aprendizes são capazes de utilizar os contrastes distintivos de sua L1 em segmentos que se encontrem em posição distinta na L2.

Com relação à duração de vogal, quando esta foi manipulada e era uma pista isolada, verificou-se que os falantes de holandês fizeram uso desse parâmetro na distinção dos pares /p/-/b/ e /t/-/d/. Contudo, tal pista foi menos utilizada por eles do que pelos informantes cujo inglês é a L1. Segundo a autora, isso pode ter ocorrido uma vez que a duração de vogal não é uma pista relevante na distinção de oclusivas surdas e sonoras em posição intervocálica no holandês. Outro fato que se soma ao anterior se encontra em uma maior exposição que esses falantes têm a outros falantes de holandês como L1 e de inglês como L2, do que a informantes falantes de inglês como língua materna.

⁴⁶ São palavras que não existem em uma dada língua e, portanto, não possuem significado, mas que respeitam a fonotaxe desta.

A percepção dos informantes do experimento seria influenciada pela duração das vogais que antecedem segmentos oclusivos surdos e sonoros que, no holandês (ELSENDOOM, 1985 *apud* BROERSMA, 2005), é menor do que no inglês. Assim, os holandeses podem ter aprendido a não tomar tal pista como tão relevante, principalmente, em um contexto que não é distintivo em sua língua: em posição final.

A adaptação perceptual pela qual aprendizes de L2 têm que passar é descrita por Sancier e Fowler (1997), as quais realizaram um experimento que suscita duas importantes questões: 1) a percepção de falantes nativos quanto à presença do que eles chamaram de “sotaque”, quando expostos às produções de uma informante que era falante nativa de português e de inglês como L2; e 2) a importância de verificar os valores de VOT⁴⁷, como fator relevante para a diferença na percepção da produção da informante. Um fato interessante sobre essa informante é que ela, periodicamente, realizava viagens para seu país de origem e, posteriormente, retornava aos Estados Unidos. Seu “sotaque”, como chamaram os autores, era percebido em ambas as línguas, indicando que havia influência da língua materna em sua L2 e vice-versa. Experimentos como o de Sancier e Fowler (*op cit.*) iluminam as discussões da área.

A partir do que foi mencionado acima, é preciso que se olhe, também, para a importância de outros parâmetros acústicos na distinção do contraste final de oclusivas, como o *burst*. Broersma (2005), como foi relatado anteriormente, expôs os informantes a dados com e sem o *burst*. O experimento indicou que a remoção do *burst* influenciou a percepção da distinção entre /b/-/p/, mas não para /t/-/d/. Como tal fato ocorreu tanto com os falantes de holandês, como de inglês, há a possibilidade de que exista alguma diferença dessa pista de acordo com o ponto de articulação: bilabial e alveolar.

⁴⁷ VOT é sigla para *voice onset time*, que é o tempo decorrido entre a soltura da oclusão de uma consoante plosiva até o início do vozeamento do som seguinte. Diferentes durações de VOT são associadas à sonoridade na literatura fonética. A diferença mais marcante é a seguinte: plosivas surdas têm valores positivos para o VOT e plosivas sonoras, valores negativos (cf. Kent & Read, 1992). Mas há também diferenças na duração do VOT que se refletem em consoantes aspiradas e não aspiradas, por exemplo, de modo que o VOT destas é menor do que o daquelas.

A influência do *burst* também é relatada para outras línguas. Flege (1989) menciona que o contraste de sonoridade existe no chinês, mas apenas em posição inicial. A distinção de vozeamento em início de palavra não se dá pela presença de sonoridade na oclusão, mas pelo *burst*. Quando submetidos a um experimento de discriminação do contraste sonoro de oclusivas em posição final, de sequências em inglês, os informantes chineses conseguiram reconhecer /t/ e /d/, ainda que o vozeamento tenha sido retirado, mas não foram capazes de distinguir quando o *burst* foi removido. Novamente, este e outros estudos, como Flege e Wang, 1989 (*apud* BROERSMA, 2005), endossam o fato de que os aprendizes utilizam pistas que são distintivas em suas L1 no reconhecimento de segmentos em uma L2.

Alguns estudos, como Ernestus e Baayen (2006) e Warner *et al.* (2004), mencionam a importância do *burst* como parâmetro acústico em estudos de produção, na presença de neutralização incompleta. A descrição dos autores aponta para o fato de que o *burst* é maior para oclusivas finais surdas do que para suas contrapartes sonoras. Warner *et al.* (*op cit.*) mencionam que o *burst* é 9 ms mais longo para as oclusivas finais surdas subjacentes, mas isso só ocorre quando elas estão diante de vogais fonemicamente longas. Em especial, Warner *et al.* (*op cit.*) relataram que o *burst* se mostrou um parâmetro importante na percepção das oclusivas sonoras, uma vez que essas possuem um *burst* menor.

Ainda, Port e Crawford (1989, *apud* SHRAGER, 2002) que realizaram experimentos com a DT no alemão, apontam que somente a duração da soltura do *burst* se mostrou como um parâmetro importante na distinção do vozeamento em oclusivas finais (entre as demais variáveis estavam: duração da vogal, duração da oclusão, duração do *burst*). Em um experimento de percepção conduzido posteriormente, os autores verificaram que 69,2% dos informantes identificaram com acuidade os pares vozeados e desvozeados. Shrager (2002) corrobora os achados de Port e Crawford (*op cit.*). Ao pesquisar sobre a influência de parâmetros acústicos distintos na percepção da neutralização terminal no russo, a autora aponta que o *burst* foi o único

parâmetro cuja presença foi significativa para a diferença do par /t-/d/. A partir da análise estatística, 74% dos dados foram identificados corretamente. Segundo a autora,

Quando as pregas vocais se abrem, e os pulmões estão contraídos, o ar passa pela boca. Como resultado, a pressão do ar se acumula atrás da oclusão na boca. Isso afeta a intensidade e a duração do *burst* na sua soltura. Uma vez que uma **duração mais longa do burst é resultado** do afastamento das *pregas vocais*, isso pode ser percebido como uma pista para **as oclusivas surdas**, tanto no processo de produção quanto no de percepção⁴⁸. (Shrager, Miriam, 2002, grifos da autora, tradução minha)

Cada um dos estudos resenhados contribuiu com a verificação da existência da neutralização incompleta e com as pistas envolvidas na produção e, em especial, na percepção do contraste sonoro de oclusivas em posição final.

As pesquisas, no geral, identificaram que embora produção e percepção estejam associadas, a relação entre os resultados dos experimentos com um e outro nível não são 1:1, ou seja, se um aprendiz, aparentemente, não produz um determinado segmento, isso não quer dizer que ele não o perceba. Não só isso, mas como além de perceber padrões distintivos em sua L1, um falante é capaz de adaptar, em certa medida, os padrões a segmentos semelhantes ou sequências sonoras que também existem na L2, mas em posições diferentes da palavra (e.g., posição inicial, mediana ou final), conforme apontou Broersma (2005).

Observa-se, também, que os ouvintes são capazes de aprender a utilizar as pistas que lhe são oferecidas na distinção de aspectos como o vozeamento nas oclusivas finais. Ainda que os parâmetros acústicos, e.g. duração da vogal, duração da oclusão e *burst*, quando isolados, tenham auxiliado na

⁴⁸ Trecho original: “When the vocal folds are open, and the lungs are contracting, the air passes to the mouth. As a result, a lot of air pressure is built up immediately behind the closure in the mouth. This affects the intensity and the duration of the burst at the release. Since **longer burst** duration is a duration **result** of an *openness* of the *vocal folds*, it may be perceived as a cue for **voiceless obstruents**, both in production and in perception processes.”

discriminação do contraste sonoro de oclusivas finais, a hipótese, de acordo com Warner *et al.* (2004), é a de que os sujeitos não utilizam todas as pistas com a mesma relevância em um contexto tido como fala natural. Isso corrobora, também, os achados de Flege (1989), nos quais a L1 influencia na seleção das pistas, dos parâmetros acústicos, que serão levados em conta no momento em que seja necessário distinguir dois dados. Pensando nos estudos conexionistas, pode-se dizer que os indivíduos selecionam as pistas de acordo com o peso distintivo que elas possuem na língua materna.

Os estudos mostraram também que alguns parâmetros são mais afetados por fenômenos como a DT do que outros. Contudo, como os resultados variam de língua para língua, não parece ser possível chegar a um acordo sobre qual fator se destaca como pista perceptual para a identificação do contraste sonoro final das oclusivas.

Além disso, fatores como a metodologia experimental influenciam nos resultados do experimento, a depender de quais e quantas variáveis são manipuladas, como elas são apresentadas – isoladas ou em grupo –, quantos são e qual o perfil linguístico dos sujeitos selecionados para o experimento, etc. Conforme menciona Port (1996), se o informante for apresentado a um conjunto de dados e ele é forçado a realizar uma discriminação de duas sequências sonoras, o resultado desse experimento pode ser enviesado, pois os falantes encerram um *continuum* de possibilidades em apenas duas categorias que lhes são apresentadas.

A partir dessa última questão, passa-se, agora, ao próximo capítulo, no qual será relatado um experimento de percepção que se insere no cenário dos estudos perceptuais da DT.

6 METODOLOGIA EXPERIMENTAL DE ESTUDOS DE PERCEPÇÃO

Os métodos experimentais na área de percepção são separados por alguns autores (YANTIS, 2001; LOPES e ABIB, 2002; GIBSON, J, 2002; RANDY, DIEHL, LOTTO e HOLT, 2004, VIANNA, 2006) em três grupos: comportamental, computacional e neurofisiológico. No primeiro grupo, se encontram os experimentos que têm como base a participação ativa dos informantes. Os sujeitos recebem estímulos e precisam tomar decisões conscientes sobre eles.

Os modelos computacionais se baseiam, em grande medida, no acesso a como os sujeitos de uma pesquisa reconhecem dados de fala, de modo que auxiliam na descoberta de como os estímulos são processados no cérebro. Uma das contribuições desses modelos para a compreensão da percepção relaciona-se à investigação acerca da maneira como o sinal de fala é processado para que as pistas acústicas sejam extraídas dele e, em última instância, para que os segmentos que constituem uma palavra sejam reconhecidos (McCLELLAND e ELMAN, 1986). Os experimentos que envolvem métodos neurofisiológicos (MINAGAWA-KAWAI, MORI, NAOI, KOJIMA, 2006), ao contrário dos comportamentais, utilizam um tipo de informação que não é consciente. Os informantes são apresentados a diversos tipos de estímulos e os resultados cerebrais são quantificados. A partir desse modelo experimental, as respostas podem revelar detalhes que antes não teriam sido captados pelo modelo comportamental. Um exemplo disso, pensando no foco deste trabalho, seria um sujeito ter como tarefa distinguir oclusivas sonoras de surdas, em posição final e, através de um teste discriminatório, não demonstrasse sensibilidade na distinção, mas, por outro lado, as suas respostas cerebrais

pudessem revelar que o informante é sensível a esse contraste. Um exemplo de método é a magnetoencefalografia⁴⁹.

Conforme descrito acima, como exemplo destes testes tem-se a identificação e a discriminação de itens. Tais testes oferecem ao pesquisador uma descrição de como os informantes percebem e categorizam os sons da fala. Embora existam diversos tipos de testes utilizados como parte da metodologia de experimentos de produção e percepção (BRASILEIRO, 2009), este capítulo concentra-se em dois tipos de tarefas experimentais: identificação e discriminação. Os estudos resenhados aqui seguirão, parcialmente, uma ordem cronológica, mas ao final serão discutidas questões sobre o modo de apresentação de estímulos que remonta a artigos da década de 1970.

6.1 TIPOS DE TAREFAS: IDENTIFICAÇÃO E DISCRIMINAÇÃO

Os testes de identificação dentro dos experimentos de percepção têm como objetivo fazer com que o informante decida, dentro de um conjunto finito de opções, qual dado corresponde ao estímulo recebido, seja ele auditivo ou visual (BRASILEIRO, 2009:37). Em geral, testes cuja tarefa consiste em identificar um dado avaliam a capacidade de um sujeito “etiquetar corretamente” um segmento. Como em outros tipos de metodologia experimental, precisa-se ter cautela na escolha dos dados que serão dispostos para o sujeito escolher e na quantidade de itens que farão parte da seleção. O pesquisador pode controlar quantas serão as opções apresentadas ao informante e isso pode tornar a tarefa mais ou menos exigente. Conforme menciona Brasileiro (*op cit.*), alguns pesquisadores consideram que um dos pontos positivos desse tipo de atividade reside no fato de que o informante não sobrecarrega sua memória. Um tipo de instrumento muito utilizado nesse tipo de teste são as fichas de leitura, com a apresentação das opções escritas ou

⁴⁹ A magnetoencefalografia “envolve a medição dos campos magnéticos associados à atividade eléctrica cerebral. Trata-se de uma técnica não invasiva que permite seguir, à semelhança da eletroencefalografia (EEG), a evolução dos processos eletrofisiológicos na escala do milissegundo” (TRINDADE, M. J. G. A , 2004).

disposição dos dados em uma tela de computador, que igualmente traz as informações escritas. Embora essa metodologia seja usada com frequência pelos pesquisadores (e.g. ESCUDERO, 2005), ela pode enviesar as respostas do sujeito, uma vez que ele pode se pautar no registro ortográfico ao invés do estímulo auditivo. Alguns experimentos tiveram seus resultados enviesados a partir desse quesito (e.g. ZIEGLER *et al.*, 2004). Por outro lado, além dessa ferramenta os pesquisadores lançam mão de figuras ou objetos que são apresentados aos sujeitos e esses precisam apontar/selecionar aquele que se mostra mais adequado a partir do estímulo recebido. Alguns pesquisadores optam por utilizar esse segundo tipo de ferramenta com crianças que ainda não foram alfabetizadas ou que estão passando pelo processo de alfabetização (BRASILEIRO, 2009). No entanto, é importante levar em conta que por trás da escolha de uma ferramenta ou outra se encontra não só a facilidade de identificação por parte do informante, mas também a precisão do estímulo. O pesquisador precisa ser criterioso na escolha da figura como ferramenta experimental, pois esta, se oferecer mais de uma interpretação, pode conduzir o sujeito a uma produção ou percepção inadequadas.

Dentre os experimentos de identificação, encontram-se os de julgamento, que verificam se uma sentença ou um segmento é adequado como estímulo ou não (BEDDOR e GOTTFRIED, 1995). Tal teste pode auxiliar, por exemplo, na avaliação de quais pistas acústicas são boas para o informante, ou seja, qual delas levam-no a identificar o estímulo recebido.

Muitas vezes, alguns experimentos optam por utilizar mais de um teste para obter resultados mais acurados ou para testar metodologias experimentais. Assim, outra tarefa possível é a de discriminação. Diferente da tarefa de identificação, um dos tipos de métodos de discriminação envolve escutar dois estímulos auditivos e perceber se existe alguma diferença acústica entre eles. Um tipo de estudo comum com este tipo de tarefa é a discriminação de dados em igual ou diferente. Esse tipo de teste é conhecido como AX. De acordo com Brasileiro (2009), uma das vantagens desse tipo de teste é o fato de que dois estímulos são apresentados aos informantes, fazendo com que, da

mesma maneira que os experimentos de identificação, a memória não seja sobrecarregada, uma vez que o sujeito não precisa lidar com uma grande quantidade de dados e variáveis ao mesmo tempo para escolher a resposta mais adequada.

Assim como na tarefa de identificação, a discriminação também oferece alguns pontos negativos. Ao ter que distinguir os dados em igual ou diferente, os informantes tendem a escolher um tipo de resposta, que pode ser enviesada pelo próprio conjunto de dados (GERRITS, 2001). Se o experimento não for apresentado de modo aleatório, as respostas dos sujeitos podem ser enviesadas pela própria apresentação do conjunto de dados, fazendo com que eles julguem como iguais vários segmentos numa sequência próxima. Alguns experimentos se debruçaram sobre ferramentas que evitassem o possível viés (GREEN e SWEETS 1966, *apud* BRASILEIRO, 2009). Pisoni (1975, *apud* BRASILEIRO, 2009) realiza um estudo no qual procura trabalhar com outro tipo de tarefa de discriminação, o 4IAX, o qual, segundo ele, é mais sensível a algumas diferenças acústicas.

Ambos os tipos de tarefas são considerados “não pesados” em relação ao tipo de ativação realizado da memória. No entanto, possuem abordagens e procedimentos de coleta de dados extremamente distintos. É preciso levar em conta o que se pretende com o estudo para que a escolha de um tipo ou outro seja calculado. Nesse sentido, existem muitos experimentos que optam por usar as duas tarefas (LIBERMAN *et al.*, 1961, *apud* BRASILEIRO, 2009), pois segundo Heeren (2006, *apud* BRASILEIRO, 2009:38): “a identificação sem a discriminação tende a superestimar as habilidades fonéticas dos aprendizes de L2”.

6.2 ESTUDOS SOBRE IDENTIFICAÇÃO E DISCRIMINAÇÃO

A partir do que foi exposto na seção acima, passa-se a um breve relato de alguns estudos que conduziram, em diversas línguas, experimentos de identificação e discriminação. O objetivo de alguns deles é descrever a distinção entre os resultados de um experimento e outro (identificação e

discriminação) e de que modo isso pode relativizar resultados encontrados em alguns experimentos de percepção e produção.

Na seção 5, discutiu-se sobre o processo de assimilação e categorização dos dados de L2 nos modelos SLM, PAM e PAM-L2. De acordo com o PAM-L2, se dois sons da L2 não possuem a mesma correspondência na L1, ou seja, não se encontram em duas categorias distintas, mas na mesma, esses não serão facilmente discriminados. Por outro lado, se os mesmos sons são assimilados em duas categorias existentes da L1, existe uma alta taxa de discriminação.

A partir do trabalho de Goto (1971, *apud* GUION, FLEGE, AKAHANE-YAMADA e PRUITT, 2000), sobre a influência da L1 (japonês) sobre a aquisição de segmentos da L2 (inglês) do par /ɹ/ e /r/, foram desenvolvidos outros trabalhos, em especial na última década, que buscavam confirmar os achados do modelo PAM. É o caso, por exemplo, dos trabalhos de Guion *et al.* (2000); Aoyama, Flege, Guion, Akahane-Yamada, e Yamada, (2004). Os resultados desses estudos apontam para uma das previsões de categorização preconizadas pelo PAM-L2, ou seja, os dois segmentos presentes em inglês seriam assimilados por uma única categoria fonêmica do japonês e não por duas distintas, como no inglês. De modo análogo ao fenômeno da DT, a literatura da área (*e.g.*, AOYAMA *et al.*, 2004) explica que o contraste entre /ɹ/ e /r/ é neutralizado, uma vez que os aprendizes relacionam ambos os sons do inglês como o /r/ do japonês, não realizando, portanto, nenhuma discriminação quando precisam usá-los no inglês. Apesar disso, como já foi mencionado (capítulo 5), embora seja possível que o contraste não seja produzido ou percebido, o aprendiz possui a capacidade de desenvolver a habilidade de aprender a identificar tal contraste de maneira a realizar distinções, diferentes das de um falante nativo, mas ainda assim adequadas (FLEGE, 1995, discutido no capítulo 5 deste trabalho).

Outro trabalho que merece ser citado é o de Guion *et al.* (2000), que objetiva analisar que modelo – SLM ou PAM – explica melhor as produções da

distinção /ɹ/, /r/ do inglês por falantes nativos de japonês. Esses falantes foram divididos em três grupos, de acordo com seu nível de proficiência, e os achados levam os autores a observar que o PAM seria mais adequado para explicar os resultados encontrados, uma vez que o modelo permite o acompanhamento desde os níveis de proficiência mais iniciais e procura explicar como os sons que não existem no inventário da L1 são assimilados.

Para além desses resultados, um dos principais pontos do trabalho é investigar a distinção feita por falantes de japonês como L1 através de experimentos de identificação e discriminação. Os pesquisadores realizaram dois testes: 1) de identificação: falantes de japonês como L1, que não haviam tido muito contato com inglês fora da sala de aula⁵⁰, realizaram um teste para julgar um conjunto de dados (segmentos em inglês e japonês); 2) de discriminação: com falantes de japonês como L1 e inglês como L2, de 3 níveis de proficiência distintos. Os estímulos foram os mesmos do teste 1 e, além disso, um grupo de 10 americanos realizou o experimento para que se pudesse comparar os dados de falantes nativos e não nativos. Nove sujeitos foram selecionados para participar do primeiro teste, sendo que esses nunca haviam morado fora do Japão e o contato com o inglês se deu apenas a partir da prática em sala de aula. Além dos sujeitos do teste 1, mais trinta sujeitos participaram do teste 2: 10 falantes de japonês que possuíam um nível alto de Inglês (além de terem feito aulas de inglês no Japão, esses constantemente usavam inglês no seu trabalho, mas nunca haviam viajado para os Estados Unidos); 10 tinham um nível de inglês considerado baixo (estudantes japoneses universitários, que tinham apenas o inglês de sala de aula, mas que haviam viajado para os USA); 10 eram falantes nativos de inglês. Entre os achados do teste 1, os autores encontraram, conforme era esperado, que os informantes identificaram mais os dados que possuíam contrastes do japonês, do que do inglês. Alguns segmentos, como o /θ/ do inglês que foi ouvido como

⁵⁰ Não possuir “muito contato com o inglês” significa aqui, conforme mencionado pelos autores, não ter viajado para nenhum país cuja língua nativa é o inglês e não ter tido contato com a língua fora do ambiente controlado de sala de aula.

/s/ do japonês, foram identificados desse modo 39% das vezes. Esses achados do experimento 1 fizeram com que os autores pensassem em um segundo experimento, agora de discriminação. Decidiu-se separar os contrastes em 3 grupos: inglês-inglês (I-I), japonês-japonês (J-J) e inglês-japonês (I-J). O teste 2 se mostrou importante, pois se verificou, assim como prevê o PAM, que o fator aprendido influencia na distinção dos contrastes, ou seja, os sujeitos com mais experiência na língua, tiveram percentuais de discriminação mais altos. Isso demonstra que é possível adquirir novos padrões contrastivos, não existentes na L1. Como era esperado, os informantes japoneses discriminaram mais corretamente os contrastes J-J do que I-J e I-I. Os autores encontraram, também, que alguns tipos de contrastes parecem apresentar mais dificuldade do que outros, como é o exemplo do par /ɹ/ e ///. O baixo percentual de discriminação sugere que “o // do inglês está mais próximo (e, assim, é mais difícil de distinguir) do japonês /r/ do que do /ɹ/ do inglês” (GUION *et al.*, 2000:2720). Esse achado confirma uma das previsões feitas pelo PAM, de que os segmentos da L2 que são mais semelhantes aos das L1 serão mais difíceis de serem identificados do que os que não são.

Além desses achados, é importante ressaltar que o modo como os autores trabalharam com os tipos de tarefa (identificação e discriminação) permitiu a organização do segundo experimento no formato como foi conduzido e, ao realizar os dois experimentos, os autores puderam refinar sua análise. Outros autores a partir de dados de línguas como o Malaio, também verificaram a relação entre experimentos de identificação e discriminação. Harnsberger (2001, *apud* BETTONI-TECHIO, 2008: 25), que realizou experimentos sobre a percepção de alguns segmentos do Malaio por falantes de Marata (idioma indiano), Oriya (idioma persa-árabe) e Americanos, investigou a capacidade dos falantes de distinguirem consoantes nasais, com pontos de articulação distintos, por sete grupos de falantes monolíngues de 7 línguas diferentes. Embora não se vá fazer uma descrição detalhada do experimento de Harnsberger, esse se destaca por apresentar resultados que vão de encontro a alguns tipos de categorização previstos pelo PAM, a saber, alguns dados

pertencentes ao tipo CG (*Category Goodness Difference*) não funcionavam de modo distinto dos itens que eram identificados como TC (*Two-category Assimilation*). A partir disso, o autor formulou outro tipo de sistema que daria conta dos sons serem categorizáveis ou não. A partir de um sistema percentual, Harnsberger (2001) estipulou que os sons da fala seriam considerados não-categorizáveis quando as respostas do grupo de sujeitos pesquisados fossem equivalentes a menos de 90%. Com isso, o autor procurou tornar a categorização em algo gradiente. Contudo, uma limitação encontrada nesse tipo de processo é saber como tratar dados que possuam valores próximos a 90%, ou seja, ainda que o sistema amplie as possibilidades de encaixar os resultados em categorias menos fixas, elas ainda existem. Segundo o autor, uma maneira de contorná-las, é criar “fronteiras” em mais de um processo de assimilação proposto pelo modelo PAM. Do mesmo modo que no estudo de Guion *et al.*(2000), o autor realizou experimentos de identificação e discriminação e, a partir deles, pôde perceber que nem todos os resultados pareciam se encaixar nas previsões dos processos de categorização do PAM.

Experimentos como o mencionado acima levantam algumas questões de ordem metodológica, que nem sempre são observadas como possíveis variáveis de um experimento, mas que podem revelar detalhes que não puderam ser verificados levando em conta a metodologia empregada. Além do tipo de tarefa apresentada ao sujeito-informante, a quantidade de estímulos e o modo como são apresentados podem influenciar nos resultados finais de um estudo. Wayland (2007), ao investigar como consoantes oclusivas do coreano e tailandês são percebidas por grupos de falantes de cada uma das línguas, discute os resultados de dois experimentos de identificação e discriminação cujos estímulos foram apresentados de modo distinto.

No experimento 1, falantes de coreano receberam a tarefa de identificar consoantes oclusivas do tailandês a partir de estímulos isolados ou em formato de “trio”. Assim, inicialmente os sujeitos receberam um único estímulo, cada consoante foi apresentada sem nenhum outro dado de apoio para testar o percentual de identificação e assimilação do Tipo Afinidade-Catagórica (CG

Type – Category Goodness Difference). Na sequência, os sujeitos do experimento recebiam três estímulos ao mesmo tempo, sendo o estímulo-alvo apresentado entre outros dois. Embora os dois formatos aos quais os informantes foram expostos tenham demonstrado resultados semelhantes, o estímulo no modo tríade foi reconhecido, mesmo que com uma diferença percentual pequena, melhor do que quando o estímulo se apresentava isolado. Para o experimento 2, que contou com o mesmo formato de apresentação de dados, mas que funcionou de modo inverso – informantes tailandeses foram escolhidos para identificar e discriminar as consoantes oclusivas do coreano –, os resultados foram diferentes. O percentual de acerto para as consoantes oclusivas do coreano, /p/, /t/ e /k/ variou entre um tipo de formato e outro: estímulo isolado e em tríade. Diferentemente do experimento 1, o formato tríade se mostrou um dispositivo mais acurado para a identificação dos pares contrastivos do experimento 2. Apesar da variação observada, assim como no experimento 1, os informantes tailandeses não tiveram dificuldades em identificar e discriminar os dados do coreano.

Até o momento, foram relatados estudos que procuraram verificar, entre outros aspectos, a relação das tarefas, sejam elas de discriminação ou identificação, o modo como os dados são apresentados, isolados, estímulo-alvo junto com outros, em palavras isoladas ou sentenças, com o resultado perceptual final. Tais estudos lançaram mão de estímulos que foram apresentados na tela do computador ou a partir de áudios. Assim, também é importante decidir a natureza dos estímulos: visuais ou auditivos.

Nesse sentido, experimentos recentes, como os de Kluge, Reis, Nobre-Oliveira e Bettoni-Techio (2009) Kluge (2009, 2010), com brasileiros, dão ênfase para a influência da natureza dos estímulos na percepção do aprendiz: não só o meio auditivo, mas também o visual. Rosenblum (2005, *apud* KLUGE, 2010: 261) acredita que “[...] está se tornando cada vez mais claro que a fala humana é uma função multimodal, geralmente apreendida por meios visuais

(leitura labial) assim como por meios auditivos (audição)” (tradução minha)⁵¹. É a partir de estudos como esses que Kluge (2009) lança a hipótese sobre os resultados do teste.

Em ambos os estudos mencionados, Kluge *et al* (2009) e Kluge (2009) as autoras analisaram o efeito de pistas visuais na percepção de consoantes nasais em posição final do inglês, /m/ (bilabial) e /n/ (alveolar), por sujeitos brasileiros. Kluge *et al.* (*op. cit*) contaram com a participação de 10 aprendizes de inglês de nível intermediário. Esses informantes realizaram três testes de identificação: AV (estímulo áudio-visual e visual), somente V (visual) e A (auditivo). Embora as autoras relatem que experimentos como o de Hazan *et al.* (2006, *apud* KLUGE *et al.*, 2009) encontraram que os informantes tiveram mais dificuldade em identificar os dados cujo estímulo era somente visual, tal fato não se mostrou o mesmo no experimento com os brasileiros. Entre os resultados desse experimento, as autoras relatam que os sujeitos se beneficiaram da combinação de pistas, ou seja, as identificações foram mais acuradas quando os informantes receberam a pista visual e auditiva. Além disso, o experimento analisou a influência da vogal que precedia o segmento nasal final. Entre as vogais pesquisadas, /ɪ/, /ɛ/ e /æ/, achou-se que as vogais baixas influenciaram a identificação de /n/; por outro lado, as vogais altas não auxiliaram na identificação de /m/. Contudo, apesar desses achados, a amostra se mostrou pequena. Assim, Kluge (2009) realizou um novo experimento, tomando como base o mesmo conjunto de dados de Kluge *et.al.* (2009), com 42 aprendizes brasileiros de nível Intermediário, sendo 21 homens e 21 mulheres, e 10 falantes-nativos Americanos (5 homens e 5 mulheres) para comparação de dados. Os sujeitos foram expostos a um conjunto de 3 pares de palavras (*Tim-tin*, *gem-gen*, e *cam-can*), produzidas por um falante de inglês do sexo masculino. O estudo de Kluge *et al.* (2009) que contou com 18 itens,

⁵¹ Trecho original: “[...] it is becoming increasingly clear that human speech is a multimodal function, usually apprehended by visual (lipreading) as well as auditory (hearing) means” (Rosenblum , 2005:51).

com 3 repetições (totalizando 108 palavras por sujeito), passou, em Kluge (2009) para 48 itens, com 3 repetições (num total de 144 palavras por sujeito). Os resultados apontam para além dos achados de Kluge *et al.* (2009), que relataram que o percentual de identificação acurada dos itens havia sido maior quando os informantes recebiam o estímulo visual. Kluge (2009) encontrou que o estímulo visual, sozinho, também foi considerado uma pista relevante na identificação dos dados. Além disso, a autora também investigou a influência da vogal antecedente ao segmento oclusivo, mas diferentemente de Kluge *et al.* (2009), o efeito das vogais foi analisado em cada uma das condições (estímulo AV, V e A) e não para as três condições. Os efeitos foram positivos para as condições AV e A, pois a vogal média auxiliou na identificação de /m/ enquanto que a vogal alta não; para /n/ a vogal média precedente não favoreceu a identificação nas condições A e AV. Contudo, para o grupo de americanos, utilizados como grupo controle, as vogais não influenciaram o processo de distinção em nenhuma das condições verificadas: AV, V e A.

Ambos os estudos mostraram que os informantes-aprendizes se beneficiam, em tarefas de identificação, da pista áudio-visual. Além desses dois experimentos resenhados, outros estudos de L2 mencionados por Kluge (2010:262-263), como o de Hayashi Sekiyama (1998, *apud* KLUGE, 2010) – que investigou a percepção de sílabas chinesas e japonesas por falantes de chinês e japonês – e Hazan *et al.*(2006, *apud* KLUGE *et al.*, 2009) – que analisaram a percepção de consoantes do inglês por falantes de várias línguas, como o espanhol e o japonês – corroboram o achado de que os informantes dependem de pistas visuais para a identificação adequada de segmentos das línguas-alvos que estão adquirindo.

Esses trabalhos tomam como base o de McGurk e MacDonald (1976), que se tornou referência por trabalhar com a natureza bimodal da percepção de fala de L1 a partir de estímulos auditivos cruzados com visuais. O estudo é dividido em dois experimentos e conta com informantes de idades variadas, a saber: 21 crianças em idade pré-escolar (3-4 anos), 28 crianças do ensino fundamental/primário (7-8 anos) e 54 adultos (18-40 anos). Em ambos os

experimentos, os informantes foram expostos a um conjunto de 4 sequências auditivas, /ba-ba/, /ga-ga/, /pa-pa/ e /ka-ka/⁵², as quais foram apresentadas aleatoriamente aos participantes. Ambos os experimentos contaram com uma combinação de dados como, por exemplo, sequência ba (áudio) e ga (movimento labial), ga (áudio) e ba (movimento labial), etc. No primeiro teste rodado, os informantes possuíam o estímulo audiovisual e tinham que assistir e repetir o estímulo, já no segundo teste, os sujeitos recebiam apenas o estímulo auditivo e, como tarefa, precisavam escutar e repetir o dado. Para o experimento 2, quando os participantes possuíam apenas o estímulo auditivo, os autores acharam respectivamente para os grupos de crianças na pré-escola, ensino primário e adultos, acertos equivalentes a 91%, 97% e 99%. Já para o primeiro experimento, no qual os informantes possuíam o estímulo auditivo e visual, a porcentagem de acertos foi de 41% (para crianças na pré-escola), 48% (para crianças do primário) e 8% (para adultos). Isso se explicaria a partir da natureza dos estímulos recebidos pelos sujeitos. O reconhecimento não apresenta “desvios” quando os informantes possuem apenas um estímulo, por exemplo, quando só têm o *input* auditivo. Além disso, os autores notaram que quando os informantes ouviam (ba), mas com movimentos labiais de (ga), como resultado, os informantes percebem o dado como (da). Os autores explicam os achados do experimento a partir de características acústicas, sendo que a informação acústica de [ba] se assemelha à de [da], mas não à de [ga]. Analogamente, pode-se pensar que os aprendizes de uma L2 possuem outras pistas, como o movimento de lábios e de outros articuladores, além do dado auditivo, primariamente utilizado em pesquisas de percepção em L1 e L2. Assim, a combinação entre características acústicas e articulatórias também pode auxiliar na identificação de aprendizes de L2. Em condições não manipuladas, diferente do experimento realizado por McGurk e McDonald, talvez os resultados fossem os mesmos observados por Kluge *et al.* (2009) e Kluge (2009): a natureza do estímulo influencia a percepção dos sujeitos, de

⁵² Cada sequência foi repetida para os informantes num intervalo de 0.5 segundos. Para maiores detalhes, ver McGurk e MacDonald (1976:746).

modo que esses passam a identificar mais acuradamente os segmentos que são expostos uma vez que possuem um *input* visual e auditivo.

Contudo, voltando ao início desta seção, quando se discutiu o possível enviesamento dos resultados a depender de como os dados são apresentados (escritos ou por meio de sons) ou a depender da sua experiência como falantes da língua-alvo, pode-se pensar que o percentual de erros encontrados por McGurk e McDonald (1976) pode estar relacionado ao processo de aprendizagem de língua em que se encontram os sujeitos do experimento: crianças em fase pré-escolar, crianças que já foram alfabetizadas – mas que ainda estão na fase de aquisição da língua escrita e falados – e adultos, que já concluíram seu processo de aquisição de L1. Têm-se aqui três grupos com perfis distintos. O primeiro grupo (crianças de 3-4 anos) se encontra numa fase de aquisição de L1, ou seja, os sujeitos estão adquirindo os gestos necessários para produzir e perceber certos sons. Ao receber os dois estímulos, auditivo (ba) e visual (ga), é esperado que os informantes tivessem mais dificuldade para identificar os sons, sendo que ainda estão em fase de aquisição. Quando se encontram em outra fase da aquisição e, portanto, já são capazes de organizar os gestos articulatorios, o percentual de erro seria menor⁵³. McGurk e McDonald (1976) realizaram um estudo sobre percepção em L1 (a extensão dos achados para outros estudos é possível, mas deve se entender que os alunos fazem algum tipo de ponderação sobre dados de aquisição de L2). Assim, a partir da exposição feita pelos autores parece ficar evidente que cada faixa etária tende a privilegiar uma pista para a percepção dos sons de sua língua, de modo que enquanto os mais jovens parecem privilegiar pistas auditivas, os mais velhos parecem privilegiar pistas visuais. Tentando ir um pouco adiante com a interpretação, talvez seja possível levantar a hipótese de que, à medida que “envelhecemos”, a percepção passa por um processo de

⁵³ Com isso, não se está afirmando que os resultados foram significativos. Não é possível verificar tal fato uma vez que os autores não submeteram seus dados a nenhuma análise estatística. Além disso, é importante mencionar que a interpretação de associar a semelhança de segmentos como (ba) e (ga) com a Fonologia Gestual parte da autora do texto, não podendo ser atribuída, portanto, à McGurk e McDonald (1976).

que vai do gradiente ao fato discreto. Nesse sentido, as crianças ainda conseguem perceber detalhes finos no sinal acústico, mas os adultos percebem apenas categorias de sons, por isso o estímulo visual acaba tendo, para eles, um peso distinto daquele que tem para as crianças.

O fato de o experimento manipular a combinação de dois estímulos, auditivo e visual, parece influenciar os grupos de informantes de modo distinto. Ainda que o estudo de McGurk e McDonald (1976) não realize um cruzamento de seus dados com os de outras línguas e com outros tipos de sons ou trace uma análise estatística para que se tenha conhecimento se a diferença percentual é significativa entre os grupos analisados e entre o conjunto de dados, é a partir de experimentos como esse que se fomentou a necessidade de enxergar a aquisição como um processo bimodal e que se começa a discutir sobre a influência da metodologia experimental nos resultados de experimentos de produção e percepção.

Este capítulo se propôs a levantar uma discussão acerca de aspectos responsáveis pela construção de um experimento de percepção e de que maneira tais fatores podem interferir nos resultados de um experimento. Ainda que outros estudos pudessem ter sido mencionados, procurou-se traçar um histórico com alguns trabalhos que tiveram como foco a análise de fenômenos distintos, mas, principalmente, a elaboração de experimentos que contassem com diferentes metodologias experimentais. O objetivo desses experimentos foi, em última instância, observar como os sujeitos respondiam aos estímulos, sejam eles visuais, auditivos ou ambos e às tarefas propostas, identificação e/ou discriminação.

As resenhas feitas nesta seção do capítulo objetivavam colher subsídios para a elaboração do experimento de percepção piloto e definitivo, os quais serão descritos no capítulo 7. Apesar de experimentos como os de McGurk e McDonald (1976) e Kluge (2009), dentre outros, lançarem mão de pistas visuais para verificar como se dá a percepção para sujeitos que estão adquirindo sua L1 ou L2 – mostrando a natureza multimodal da percepção –, o estudo realizado por este trabalho será baseado apenas em pistas auditivas.

7 METODOLOGIA DO PILOTO E EXPERIMENTO

Como discutido no capítulo 6, existem diversos fatores envolvidos no *design* de um experimento, que vão desde o embasamento teórico adotado até o perfil dos informantes, passando pela escolha do modelo experimental e a elaboração do *corpus*. A L1 dos informantes é outro fator que deve ser considerado, pois conforme observado no capítulo 5 a propósito dos experimentos resenhados, os aprendizes utilizam pistas contrastivas existentes em sua L1, ainda que tais pistas se encontrem em contextos distintos na L1 e na L2. Com isso e a partir dos questionamentos levantados por Zimmer e Alves (2008) e Albuquerque (2010), passa-se ao experimento de percepção piloto realizado com base nos resultados encontrados por Albuquerque (*op.cit*).

A divisão deste capítulo será feita da seguinte forma: primeiro se apresentará o método e resultados do piloto (de percepção), o qual contou com os mesmos dados produzidos pela falante nativa do experimento realizado por Albuquerque (2010) e com a mesma sentença-veículo (“say_____ promptly”); em um segundo momento será apresentado o método e resultados do experimento, construído a partir de problemas do piloto e do aumento da amostra (quantidade de informantes).

7.1 PILOTO

O piloto teve como objetivo analisar a utilização de pistas descritas ao longo do trabalho – duração do vozeamento, duração da vogal adjacente à oclusiva e a presença de *burst* – na percepção de oclusivas finais do inglês por falantes do PB como L1. Ele se insere na continuidade de experimentos de percepção que procuram contribuir para o tratamento do fenômeno da DT. Assim como outros experimentos que investigaram a DT em outras línguas e com aprendizes brasileiros de inglês como L2, (ZIMMER, 2004; ZIMMER e ALVES, 2007; ZIMMER e ALVES, 2008; BETTONI-TECHIO, 2005; ALBUQUERQUE, 2010) este piloto seguiu os seguintes objetivos e hipóteses:

a) Objetivos:

- 1) Verificar a percepção de falantes brasileiros, aprendizes de inglês como L2, das pistas acústicas (duração de vogal, duração do vozeamento e *burst*) relatadas pela literatura (WARNER *et al.*, 2004; KLEBER *et al.*, 2010).
- 2) Observar a influência do ponto de articulação (bilabial, alveolar, velar) no percentual de discriminação do contraste surdo-sonoro em posição final pelos informantes-aprendizes.
- 3) Verificar se há um maior percentual de acertos quando as pistas não são manipuladas e se os informantes-aprendizes reconhecem mais uma determinada pista em detrimento das outras.
- 4) Observar se os informantes-aprendizes são capazes de reconhecer o contraste surdo-sonoro em posição final como em (WARNER *et al.*, 2004; KLEBER *et al.*, 2010) mesmo que tal discriminação não seja feita na L1 deles.

b) Hipóteses

- Como encontrado em outros estudos (WARNER *et al.*, 2004; KLEBER *et al.*, 2010), espera-se que a duração de vogal seja uma pista mais robusta para a discriminação de oclusivas surdas e sonoras em posição final;
- Espera-se que, assim como no experimento de produção realizado por Albuquerque (2010), o percentual de acerto entre as oclusivas pesquisadas (bilabiais, alveolares e velares) varie de acordo com o ponto de articulação;
- Espera-se que os informantes discriminem com mais acurácia os dados sem manipulação e que as pistas manipuladas (duração do vozeamento, duração de vogal e *burst*) tenham percentuais de discriminação distintos, uma vez que o PB não possui oclusivas em fim de palavra. Assim, os sujeitos podem apresentar um desempenho melhor em reconhecer uma ou outra pista.
- Prevê-se que os aprendizes de inglês como L2, falantes de PB como L1, sejam capazes de utilizar as pistas analisadas como ferramentas de

discriminação do contraste surdo-sonoro mesmo que essas sejam utilizadas em um contexto distinto em inglês (início e fim de palavra) e em PB (início e meio de palavra).

7.1.1 Os informantes

Inicialmente, o piloto foi aplicado a 10 informantes, com idades entre 15 e 30 anos, dos quais 2 precisaram ser descartados, uma vez que as distinções para ambos foram realizadas de modo aleatório: os sujeitos deram uma resposta padrão “=” para metade dos dados e “*” (diferente) para outra metade. Desta forma, o piloto contou com 8 informantes falantes de português brasileiro, que estudavam inglês no Celin-UFPR (Centro de Línguas e Interculturalidade da Universidade Federal do Paraná).

Os informantes foram selecionados a partir do tempo de instrução formal de inglês recebido ao longo das suas vidas como aprendizes. Todos os sujeitos selecionados encontravam-se no nível Intermediário e haviam completado cerca de 400 horas de instrução formal⁵⁴.

As informações sobre os informantes se encontram organizadas na Tabela 1:

⁵⁴ Por instrução formal entende-se a realização de um curso de língua em um instituto de línguas ou escola de línguas que não o ensino regular completado, geralmente, no período escolar.

Participantes	Idade	Naturalidade	Grau de escolaridade	Tempo de Instrução formal (em horas)	Conhecimento de outra língua estrangeira
LU	24	Pato Branco-PR	Superior completo	420	Francês.
JU	25	Curitiba-PR	Superior Completo	400	Não
PA	26	Curitiba-PR	Superior Completo	450	Não
TH	25	Curitiba-PR	Superior Completo	380	Não
ER	17	Curitiba-PR	Concluindo o ensino médio	400	Não
DU	26	Curitiba-PR	Superior incompleto	370	Não
BI	25	Curitiba-PR	Superior Incompleto	420	Italiano
MA	28	Curitiba-PR	Superior Completo	480	Não.

Tabela 1 Perfil dos informantes falantes de PB como L1 que participaram do piloto.

7.1.2 O *corpus* e a gravação de estímulos do teste de percepção

Conforme exposto no final do capítulo 4, Albuquerque (2010) realizou um estudo de produção que contou com um *corpus* composto de três contextos adjacentes, a saber: inexistente (sem a presença de nenhum elemento após a palavra-alvo), vocálico e consonantal (plosivo). Entre esses três contextos, observou-se que algumas pistas, como o *burst*, eram realizadas

majoritariamente apenas quando o contexto seguinte era a oclusiva (como “*promptly*”). Optou-se, então, para o piloto, por inserir as palavras-alvo (expostas no quadro 1) em uma sentença-veículo como “say_____ *promptly*”⁵⁵. Tal contexto foi escolhido, pois: a) como observam Zimmer e Alves (2007), a presença de um contexto adjacente tem influência na produção e percepção dos dados de aquisição, uma vez que os aprendizes não estão acostumados a ouvir os dados isolados; b) um contexto vazio adjacente (conforme foi utilizado no experimento de produção de ALBUQUERQUE, 2010) poderia comprometer a presença, em especial, do *burst* (elemento que será analisado na pesquisa). Além disso, ter como contexto adjacente uma vogal, como na sequência “say _____ *again*”, eliminaria as chances de, novamente, se encontrar o *burst*, por se tratar de um contexto no qual a transição entre a palavra-alvo e a seguinte é tênue e, também, pela sonoridade da vogal se mesclar à da oclusiva sonora, sendo difícil, nesse sentido, de manipular o dado. Apesar do contexto em que se tem uma oclusiva, como na palavra “*promptly*”, não excluir a possibilidade de que não se tenha a presença do *burst* em final de palavra, em todas as 5 repetições dos dados da falante nativa PM, foi nesse contexto que se encontrou o *burst*. Não houve produção de *burst* quando o contexto era vocálico ou inexistente.

Os estímulos foram selecionados de modo que se formassem pares mínimos e houvesse a mesma vogal, uma vez que a sua qualidade poderia inserir outra variável no experimento. Assim, os dados selecionados foram:

Oclusivas surdas	Oclusivas sonoras
cap [kæp]	cab [kæb]
bat [bæt]	bad [bæd]
back [bæk]	bag [bæg]

Quadro 1 Estímulos utilizados no experimento piloto de percepção com oclusivas finais do inglês.

⁵⁵ Os estímulos deste teste foram retirados de Albuquerque (2010)

É importante ressaltar que se optou por trabalhar com todos os pontos de articulação das oclusivas, pois a literatura (PYE, 1986) apresenta indícios da influência do ponto de articulação na identificação da sonoridade da consoante final. Além disso, Albuquerque (2010) verificou que existe uma gradiência entre o percentual de sonoridade a depender do ponto de articulação.

Os estímulos no teste de percepção foram gravados em um laboratório com tratamento acústico, LEFON (microfone Shure KSM 27, com a taxa de amostragem do sinal a 44100hz) com a informante nativa PM (do experimento conduzido por Albuquerque, 2010). A informante PM era da Carolina do Norte e tinha 35 anos na época das gravações. PM não possuía conhecimento de outra língua que não o inglês.

Os estímulos foram manipulados, de modo a trabalhar com os parâmetros descritos previamente: duração de vogal, burst e vozeamento. Para realizar a manipulação dos dados, foi extraída a duração média dos parâmetros (dos dados gravados pela informante nativa PM). Isso foi feito para que a pista manipulada tivesse a duração adequada em relação ao dado sem manipulação. Abaixo é possível ver a manipulação do dado “*say bad promptly*” (Figura 16), no qual foi extraído *burst* e o vozeamento, restando apenas a duração de vogal (Figura 17).

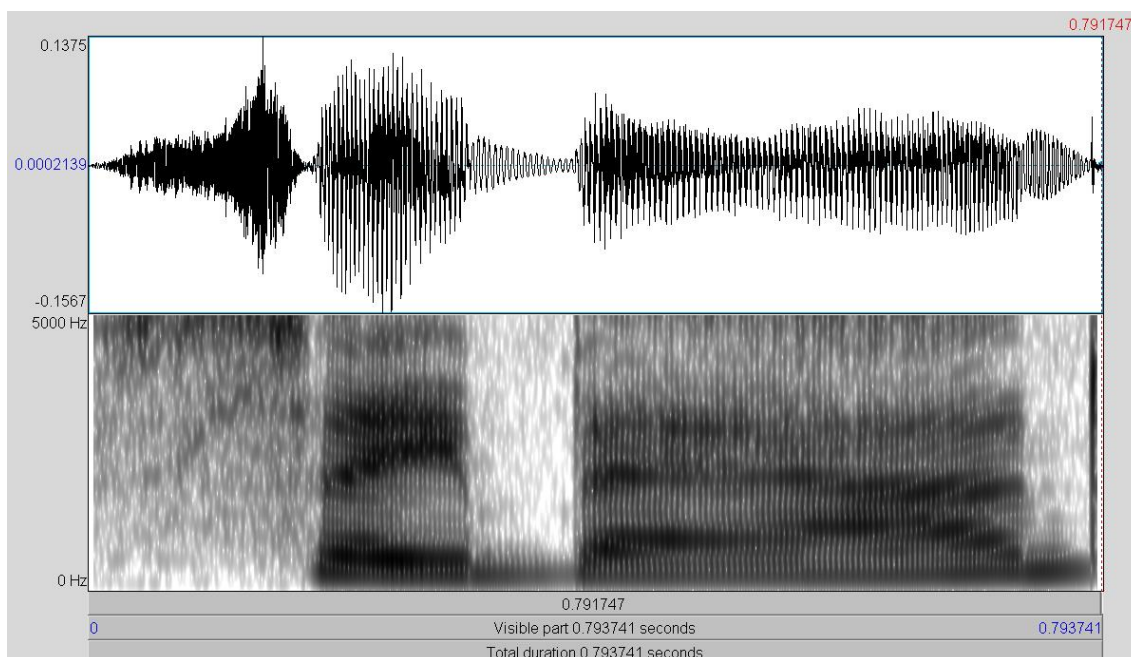


Figura 16 O espectrograma da sentença “say bad promptly” produzida pela informante PM (nativa), sem a manipulação da palavra-alvo “bad”.

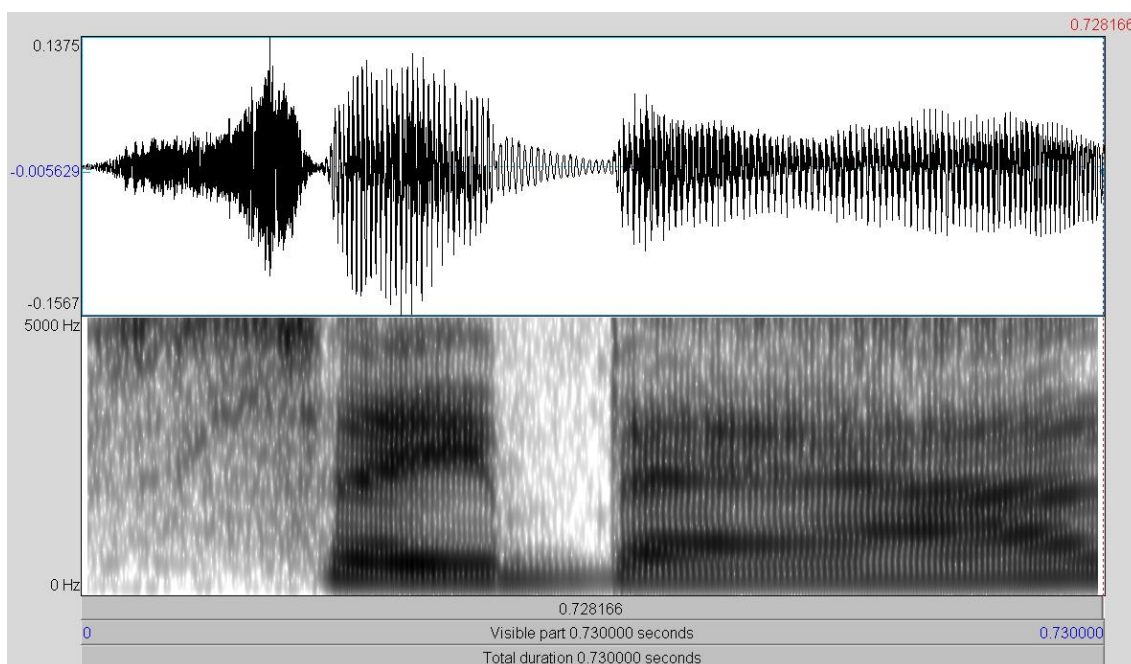


Figura 17 O espectrograma da sentença “say bad promptly” produzida pela informante PM (nativa), com a manipulação da palavra-alvo “bad”, sem burst, sem barra de sonoridade (indicação visual do parâmetro vozeamento).

Do mesmo modo que a pista *burst* foi manipulada, as pistas duração da vogal antecedente e vozeamento receberam o mesmo tratamento: calculou-se a média da duração da sentença em relação à palavra. Posteriormente, também foram calculadas as médias da duração do vozeamento e duração da vogal, afim de que se obtivesse um valor adequado para realizar a manipulação, não obtendo, portanto, sentenças com tempo distinto do tempo médio produzido pela informante PM. Outro aspecto importante para a manipulação se encontra no ponto do corte. É importante que se preserve, para a duração da vogal, a transição entre os formantes, para que o dado manipulado não possua um efeito artificial. Abaixo, na Figura 20, pode-se notar a duração da vogal manipulada. Nas Figuras 18 e 19, para que se compare a manipulação feita na vogal do segmento “*bat*”, tem-se a duração da vogal antecedida de oclusiva surda (Figura 18) e da sua contraparte sonora (Figura 19), respectivamente.

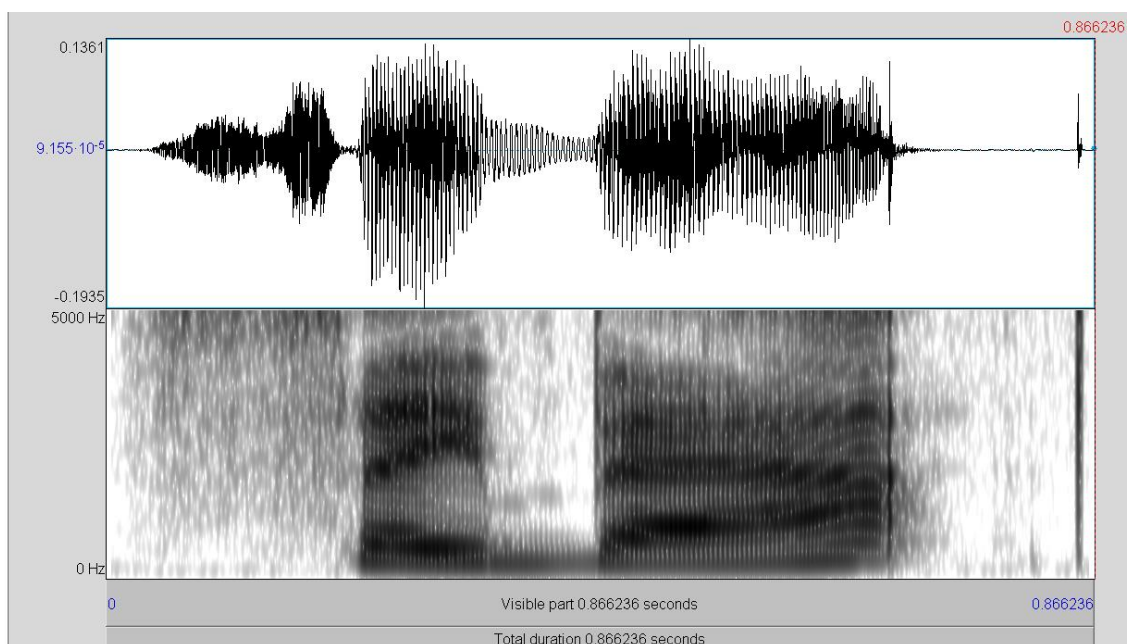


Figura 18 O espectrograma da sentença “say bat promptly”, sem manipulação, produzida pela informante PM (nativa). Duração da vogal da sequência “bat”: 0,2345 ms.

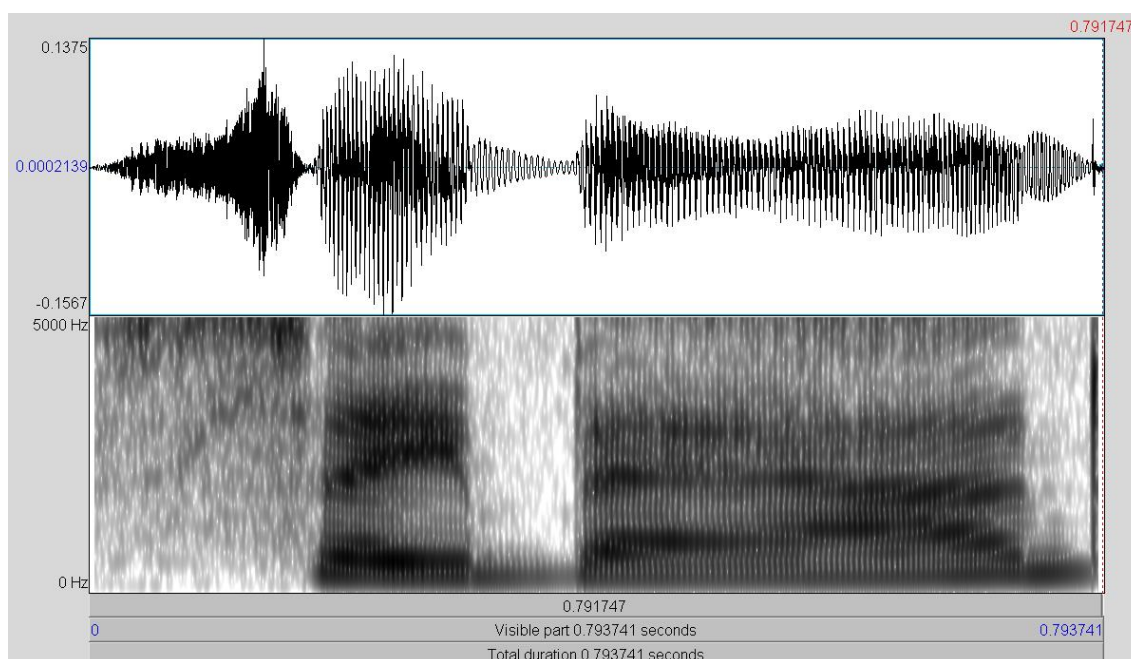


Figura 19 O espectrograma da sentença "say bad promptly", sem manipulação, produzida pela informante PM (nativa). A duração da vogal é da sequência "bad" é de 0,3459 ms.

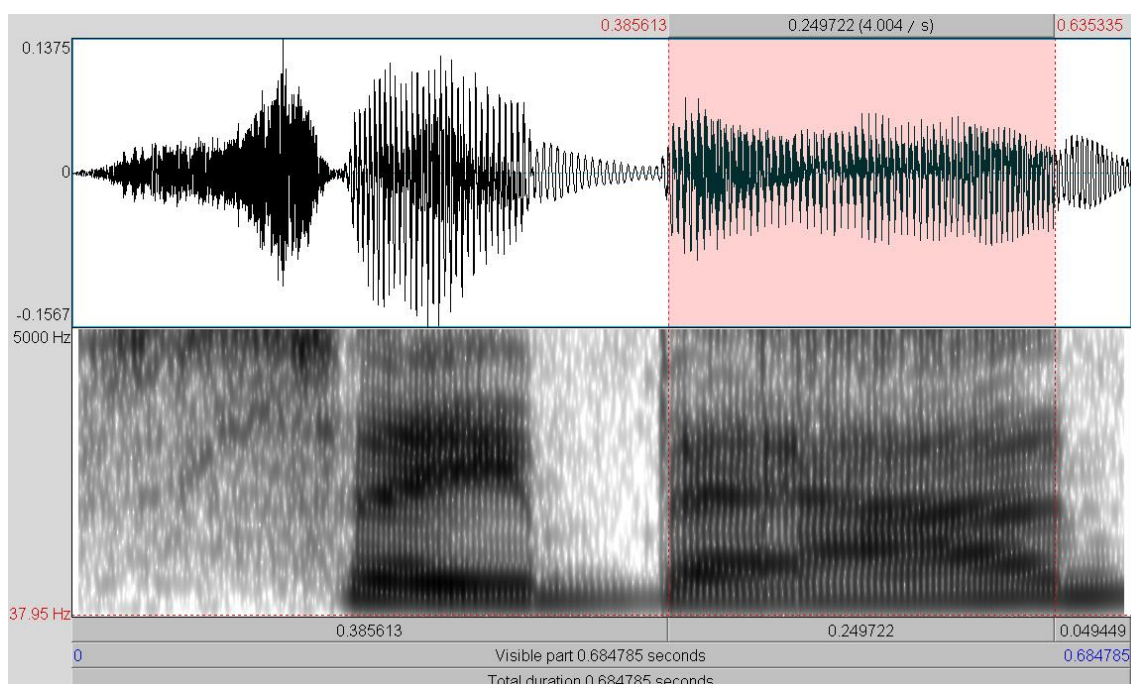


Figura 20 O espectrograma da sentença "say bad promptly", com a duração da vogal manipulada, produzida pela informante PM (nativa). A duração da vogal é da sequência "bad" é de 0,2497 ms.

7.1.3 O teste de discriminação perceptual

7.1.3.1 Procedimentos

Todos os participantes eram alunos do curso de inglês, nível intermediário, no Celin-UFPR. Foram realizadas duas seções de gravação, sendo que cada uma durou em média 20 minutos. A coleta de dados ocorreu no laboratório de informática do Celin-UFPR pela pesquisadora e as instruções foram dadas em português. Embora as instruções tenham sido dadas coletivamente, cada informante realizou o teste em um computador com fone de ouvido. A cada grupo de 12 dados, o informante recebia um aviso “Descanse um pouco. Clique para continuar”. Tal sentença foi incluída a fim de que se evitasse o efeito de cansaço, o qual poderia enviesar os resultados. É mister dizer que não foi realizado nenhum teste de familiarização.

Antes de realizarem o teste, os informantes receberam um questionário para que completassem com informações pessoais e outras relacionadas ao tempo de aprendizado da língua e contato com outros idiomas (ver Anexo 10.1).

7.1.3.2 O teste

Como se tratou de um piloto seguiu-se o modelo experimental utilizado em outros estudos sobre a DT em outras línguas (WARNER *et al.*, 2004; BROERSMA, 2005; SHRAGER, 2002): o discriminatório. O piloto foi montado de forma a oferecer aos informantes, na tela do computador (a partir do programa Praat), dois sinais, = e *. Além disto, os informantes ouviam dois estímulos auditivos e deveriam, então, pressionar = na tela do computador, se concluíssem que os dois dados ouvidos eram iguais, ou *, se considerassem os dois dados diferentes. Para descartar um possível viés ortográfico, os dados não eram visualizados na tela, os informantes tinham apenas o estímulo sonoro, conforme pode ser verificado nas Figuras 21 e 22. Após apertarem o

botão “ok”, os informantes eram direcionados para uma tela semelhante à Figura 21:

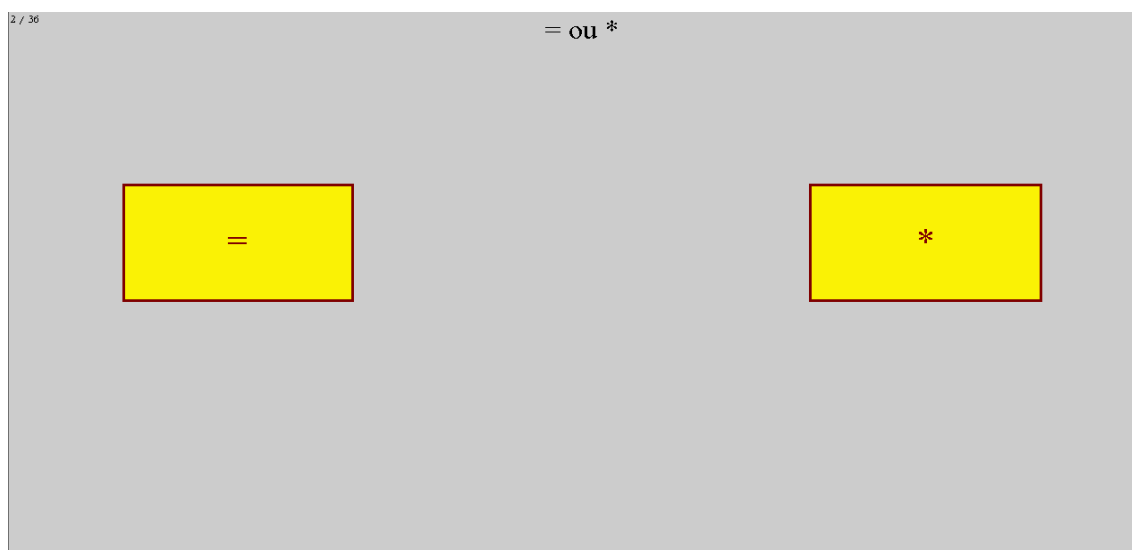


Figura 21 Tela do Praat visualizada pelos informantes ao receberem o estímulo auditivo.

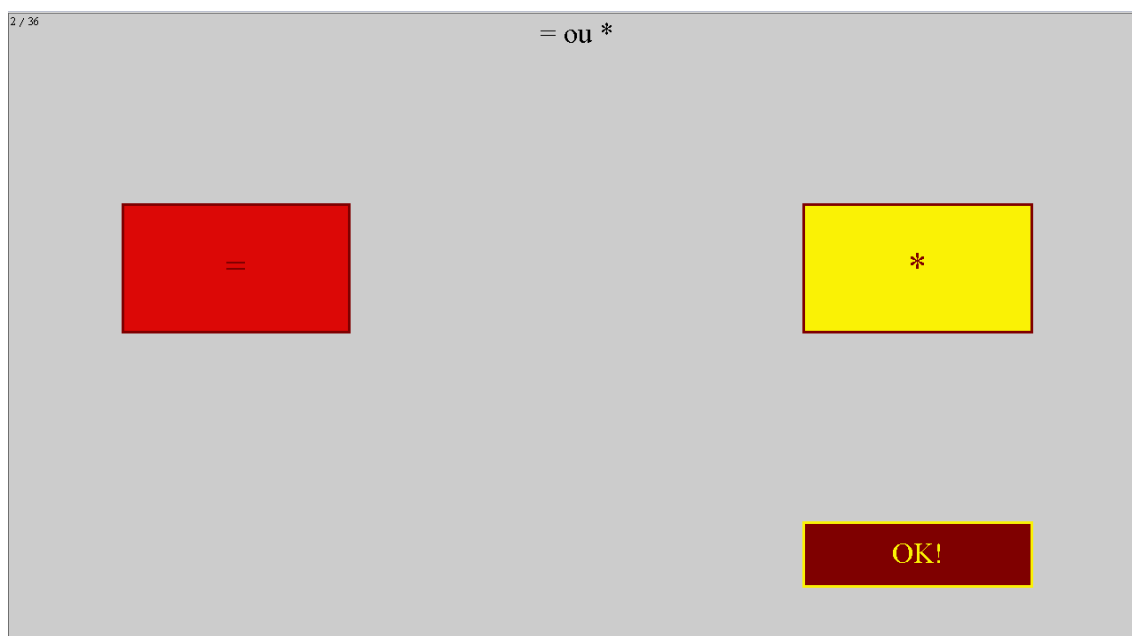


Figura 22 Tela do Praat visualizada pelos informantes ao clicarem na resposta que julgassem correta.

O teste foi composto de um total de 288 dados, sendo que cada informante avaliou 36 dados, conforme o esquema abaixo. Entre estes dados, os sujeitos ouviam: o dado sem manipulação, com manipulação da duração de vogal, com manipulação de vozeamento e com manipulação de *burst* (presença ou não).

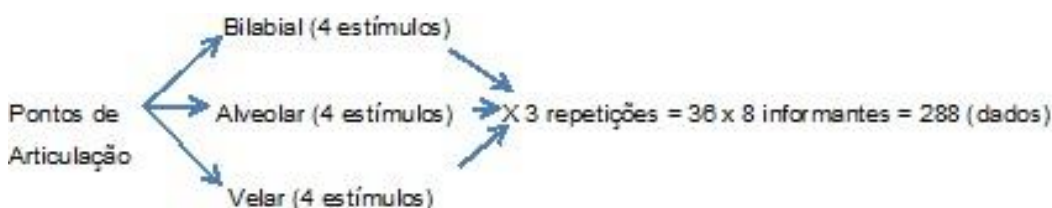


Figura 23 Esquema informativo da quantidade de dados obtidos.

7.1.4 Resultados do piloto

Ainda que os resultados deste piloto ofereçam poucos indícios sobre a seleção de pistas primárias e secundárias na discriminação de oclusivas finais do inglês, algumas conclusões podem ser traçadas.

O piloto contou com 8 informantes. Um dos problemas encontrados pela pesquisadora foi o tamanho da amostragem. Embora os achados do estudo tenham apontado na mesma direção de outras pesquisas já aqui resenhadas (e.g. FLEGE, 1989; WARNER *et al.*, 2004; BROERSMA, 2005), como a amostra foi pequena (poucos informantes), uma nova investigação será necessária, com um maior número de informantes, para que se averiguem os resultados deste piloto.

Os primeiros resultados revelaram que quando os dados não haviam sido manipulados, ou seja, os informantes contavam com todas as pistas – duração de vogal, vozeamento e *burst* – o percentual de acerto foi maior. No entanto, para os dados manipulados, o cenário é um pouco distinto. Como pode ser visto na Tabela 2, que contempla a percepção de oclusivas bilabiais,

o percentual de discriminação das pistas duração de vogal e vozeamento foi semelhante: 75% dos sujeitos foram capazes de distinguir os dados utilizando esta pista quando precisavam confrontar “*cab* manipulado” e “*cap*”. Por outro lado, a porcentagem de discriminação correta foi de 100% quando a pista era o *burst*. Ainda que não seja possível mencionar a significância de tal achado, uma vez que não se conduziu uma análise estatística, é possível que se comece a pensar que os falantes parecem estar separando as pistas em primárias e secundárias, uma vez que embora a discriminação das pistas vozeamento e duração de vogal não tenha sido de 100%, a discriminação de ambas as pistas não foi nula, elas foram consideradas como parâmetros para a distinção de oclusivas surdas e sonoras em posição final.

Dado	N	Número de acertos	Porcentagem de acertos
Cab = cap	24	24	100%
Cab burst * cap	24	24	100%
Cab vogal *cap	24	18	75%
Cab vozeamento *cap	24	18	75%

Tabela 2 Resultados do experimento de discriminação para as oclusivas bilabiais. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais ou diferentes.

Com relação às oclusivas alveolares, Tabela 3, pode-se verificar o mesmo achado. Neste conjunto de dados, tem-se que as pistas vozeamento e duração de vogal também obtiveram um percentual de discriminação de 75%. No entanto, quando a pista era o *burst*, todos os informantes pesquisados discriminaram o dado “*bad*” como sendo distinto de “*bat*”. É possível dizer que a pista *burst* parece ter sido escolhida, ao menos neste conjunto de dados, como um parâmetro mais saliente para a identificação da oclusiva alveolar sonora [d].

Dado	N	Número de acertos	Porcentagem de acertos
Bad = bad	24	24	100%
bad burst * bat	24	24	100%
bad vogal * bat	24	18	75%
bad vozeamento * bat	24	18	75%

Tabela 3 Resultados do experimento de discriminação para as oclusivas alveolares. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais e diferentes.

Apesar dos achados serem iguais para as oclusivas bilabiais e alveolares, os valores oscilam para as velares. Nota-se na Tabela 4 que as três pistas *burst*, duração de vogal e vozeamento foram discriminadas de modo distinto pelos informantes. Novamente, não é possível afirmar que existe uma diferença significativa entre os dados, mas a pista *burst* parece ter sido utilizada como pista primária na discriminação do contraste surdo-sonoro.

Dado	N	Número de acertos	Porcentagem de acertos
bag =bag	24	24	100%
bag burst * back	24	24	100%
bag vogal * back	24	21	88%
bag vozeamento * back	24	18	75%

Tabela 4 Resultados do experimento de discriminação para as oclusivas velares. Os sinais “=” e “*” correspondem, respectivamente, aos dados percebidos como iguais e diferentes.

Em linhas gerais, o piloto corroborou a hipótese lançada por outros estudos, nos quais os falantes são capazes de reconhecer uma pista existente na sua L1 no sistema da L2, mesmo que essa se encontre em um contexto distinto de sua língua materna (FLEGE, 1989; WARNER *et al.*, 2004 e BROERSMA, 2005). Um exemplo disto é a utilização do *burst* como pista que existe na distinção de oclusivas iniciais e em posição intervocálica no PB, mas não em posição final. Além disso, o estudo também verificou que os dados sem manipulação foram discriminados corretamente pelos informantes. Tal fato pode ser observado para todas as oclusivas, independentemente do ponto de articulação.

No entanto, ao contrário da previsão feita no experimento de produção com oclusivas finais do inglês por falantes de PB como L1 (ALBUQUERQUE, 2010), não se observou uma relação entre as pistas utilizadas na discriminação surdo-sonora (*burst*, duração de vogal e vozeamento) e o ponto de articulação. Isto pode ter ocorrido, pois a amostra (número de informantes) foi pequena. É importante ressaltar que tal fato precisa ser verificado a partir de um novo experimento, que conte com um maior número de informantes. Do mesmo modo, o piloto não encontrou os mesmos resultados que Warner *et al.*, (2004) e Kleber *et al.* (2010), os quais relataram ser a duração da vogal a pista mais robusta para a discriminação do contraste surdo-sonoro em posição final. O piloto revelou, no entanto, o *burst* como a pista que os informantes mais discriminaram corretamente. Isto poderia ser explicado através de um sistema de pesos, da abordagem conexionista, uma vez que a pista *burst*, por ser algo que não é utilizado no reconhecimento das oclusivas em PB, pode se tornar uma pista mais saliente, com maior peso na discriminação de dados de L2. Novamente, os resultados não foram conclusivos, uma vez que a amostragem foi pequena e a taxa de acertos entre os parâmetros não mostrou uma variação maior do que 2 informantes para os parâmetros duração de vogal e vozeamento para todos os pontos de articulação.

São necessários ajustes no tamanho da amostragem para que se verifique se os resultados são robustos. No entanto, outro passo foi dado na

direção do processo de aquisição de percepção de oclusivas finais por aprendizes brasileiros no sentido de oferecer dados que se inserem nos estudos sobre a DT. O fato dos informantes terem sido capazes de discriminar os dados independentemente da manipulação feita pode significar que não há uma neutralização do contraste surdo-sonoro, não só para a produção, mas também para a percepção.

7.2 EXPERIMENTO

A partir dos questionamentos levantados no capítulo anterior, sobre a metodologia dos experimentos de percepção e os resultados do piloto neste capítulo, formulou-se um novo experimento. É importante salientar que este experimento é fruto dos testes metodológicos realizados no piloto e de problemas advindos dele⁵⁶. Assim, os objetivos deste experimento são:

- 1) a influência do ponto de articulação na discriminação e identificação do contraste surdo-sonoro para oclusivas finais do inglês para falantes de PB como L1;
- 2) se existe uma diferença significativa entre um maior percentual de acerto entre diferentes níveis de fluência (nível básico, pré-intermediário e intermediário);
- 3) a influência do tipo de tarefa (identificação e discriminação) dada ao sujeito na identificação do contraste em posição final;
- 4) a correlação entre o tempo de reação para responder aos estímulos e os dados tidos como “errados”;

⁵⁶ Conforme exposto na descrição do experimento piloto, alguns problemas tiveram de ser revistos para a elaboração do experimento definitivo, a saber: tamanho da amostra, quantidade de manipulações, heterogeneidade do grupo-alvo (com relação ao nível de inglês), impossibilidade de se relatar a significância dos resultados.

5) semelhanças e diferenças na produção da DT para o grupo experimental (GE) e o grupo controle (GC), através da comparação entre os resultados obtidos para ambos os grupos.

7.2.1 Os informantes

Com base no piloto descrito neste capítulo, observou-se que a quantidade de informantes, assim como o perfil do aprendizado desses não havia sido controlado satisfatoriamente. Tal fator pode ter enviesado consideravelmente os resultados do piloto. Para este novo experimento, a amostra foi aumentada e foram selecionados informantes de níveis de proficiência distintos, os quais fizeram parte do GE (Grupo Experimental), e um grupo de falantes nativos de inglês como GC (Grupo Controle). Pensou-se em pesquisar mais de um nível de proficiência para testar a hipótese lançada por modelos como o PAM-L2, de que existe uma relação de aprendizado dos fones da L2 e este aprendizado não se restringe ao nível fonético.

Assim, foram selecionados 60 informantes para o GE, 20 de cada nível de proficiência: básico, pré-intermediário e intermediário. Estes possuíam idades entre 25 e 38 anos e eram estudantes do Celin–UFPR. Tomou-se o cuidado para que os informantes de básico estivessem no segundo semestre de instrução formal. Tal controle foi realizado para que se garantisse que os informantes teriam um conhecimento mínimo sobre a fonologia do inglês, de modo a saberem da existência do contraste surdo-sonoro em fim de palavra no inglês. Os informantes do GE foram selecionados a partir de um questionário (ver Anexo 10.1) que, entre outras informações, descartou o conhecimento prévio de outra língua que não o inglês, um grande intervalo no aprendizado do inglês como língua estrangeira e o fato de terem estudado inglês apenas na escola.

Duas ferramentas foram utilizadas para classificar os aprendizes nos 3 níveis de proficiência previamente mencionados: horas de instrução formal na língua inglesa e o teste TOEIC (*Test of English for International*

Communication)⁵⁷. Assim como no piloto, os informantes do GE foram primariamente selecionados de acordo com a quantidade de horas de instrução formal na língua inglesa. Dessa forma, os alunos classificados como inseridos em um nível básico de instrução tinham entre 90 - 120 horas de aula, os do pré-intermediário entre 240-340 e os do intermediário entre 480 e 580 horas. Essas horas correspondem à carga horária ministrada em cada semestre letivo do curso de línguas do Celin-UFPR (Centro de Línguas e Interculturalidade): cada semestre possui 60 horas. Dentro desta carga-horária, os aprendizes são expostos às quatro habilidades: expressão escrita (escrita), compreensão escrita (leitura), compreensão oral (prática auditiva) e expressão oral (fala/expressão oral). Tomou-se o cuidado de que todos os 60 informantes do GE tivessem passado pela mesma metodologia de ensino, ou seja, todos haviam realizado o curso na instituição mencionada. A justificativa para a utilização da quantidade de horas de instrução como uma ferramenta primária na classificação e seleção dos informantes reside no fato de que a quantidade de horas atribuída a uma mesma metodologia de ensino, com uma carga-horária regular de exercícios de pronúncia e de outras habilidades, pode se mostrar mais adequada ao tipo de experimento que foi realizado aqui (sobre a percepção de determinados sons). No entanto, como já foi mencionado no piloto, a ferramenta mais comumente utilizada para checar o nível de proficiência dos informantes são os testes de proficiência, como o IELTS (*International English Language Testing System*), o CPE (*Certificate of Proficiency in English*), TOEFL, TOEIC, entre outros. Cada um desses testes possui um tipo de avaliação sobre as habilidades adquiridas no processo de aquisição de uma L2 e um entendimento sobre o que é ser proficiente em uma língua estrangeira. Como ferramenta auxiliar, após a seleção primária a partir da quantidade de horas de instrução formal, utilizou-se o TOEIC. Tal teste foi selecionado, pois entre as atividades utilizadas para avaliar a proficiência do aluno, se encontram exercício de leitura (para checar a pronúncia de

⁵⁷ Para maiores informações sobre a base pedagógica e linguística do teste ver: <http://www.ets.org/Media/Research/pdf/TC-10-11.pdf>

segmentos isolados) e atividades comunicativas, que permitem avaliar a competência oral do informante. Ainda que nenhuma ferramenta de classificação do nível de proficiência dos informantes seja completamente eficiente, pois não é possível avaliar de modo concreto todas as habilidades cognitivas dos informantes na escolha do que e como produzir um determinado som, procurou-se uma classificação adequada ao propósito do experimento. Assim, a partir do TOEIC, as classificações se mantiveram para todos os níveis. Contudo, é preciso salientar que os resultados no TOEIC de alguns informantes do grupo intermediário (nomenclatura do Celin-UFPR) ficaram próximos à pontuação indicada como divisória entre pré-intermediário e intermediário. Dessa forma, é importante apontar que, embora tenha sido conduzido um teste de nivelamento dos alunos, a partir do exame TOEIC, para este experimento, parece ser mais adequado classificar os alunos nos níveis da instituição de ensino (básico, pré-intermediário e intermediário), pois nesses módulos está assentado um tipo de trabalho com as habilidades, a pronúncia especialmente, que é imprescindível como variável a ser controlada no experimento. Assim, os aprendizes serão referidos como GE e pelos níveis básico, pré-intermediário e intermediário. O Quadro 2 mostra alguns dados dos informantes: sexo, idade média, lugar onde nasceu, a média de horas de ensino formal, a média de informantes que viajaram para um país onde se fala inglês como L1, conhecimento de outra LE (ver Anexo 10.3 para verificar o perfil detalhado dos informantes).

	Nível Básico	Nível Pré-Intermediário	Nível Intermediário
Nº de informantes	20	20	20
Homens	11	8	9
Mulheres	9	12	11
\bar{X} idade	30	30	30
\bar{X} Horas de ensino formal	103	318	510
Conhecimento de outra LE (Língua Estrangeira)	Não.	Não.	Sim (1 informante – conhecimento de Espanhol)
Viajou para país de língua inglesa	Não.	Não.	Sim (1 informante – Estados Unidos – Chicago)

Quadro 2 Resumo dos dados dos informantes do GE. X: média; LE: Língua Estrangeira.

Para o GC, foram selecionados 8 americanos da Carolina do Norte (ver Anexo 10.2 para observar o questionário aplicados aos informantes nativos). A formação desse grupo se deu pela disponibilidade dos participantes: todos estavam realizando curso básico de português brasileiro para estrangeiros no Celin. Os informantes possuíam idades entre 20 e 35 anos e ficaram no Brasil por 1 mês. É necessário salientar a importância de o GC ser da Carolina do Norte, pois os estímulos foram gravados por um informante do mesmo lugar, o que evitou que outra variável fosse inserida no experimento: sotaque. Nenhum dos informantes do GE e GC receberam pagamento por terem participado do experimento⁵⁸.

⁵⁸ Não há, de modo geral, “pagamento a informantes” de experimentos deste tipo no Brasil, ao contrário do que acontece em outros países, como EUA, e.g. Ainda assim, como forma de agradecimento pela participação, ofereceu-se um chocolate a cada participante.

Informantes	Idade	Sexo	Nível de Português no Celin	Horas de estudo no Celin	Naturalidade	Grau de escolaridade	Outra LE (Língua Estrangeira)
N1	29	M	Básico	80	EUA/Carolina do Norte	ESC	Não.
N2	25	M	Básico	80	EUA/Carolina do Norte	ESC	Não.
N3	26	M	Básico	80	EUA/Carolina do Norte	ESC	Espanhol.
N4	26	M	Básico	80	EUA/Carolina do Norte	ESC	Francês.
N5	35	M	Básico	80	EUA/Carolina do Norte	ESC	Não.
N6	33	F	Básico	80	EUA/Carolina do Norte	ESC	Não.
N7	28	F	Básico	80	EUA/Carolina do Norte	ESC	Não.
N8	28	F	Básico	80	EUA/Carolina do Norte	ESC	Não.

Quadro 3 Perfil dos informantes do GC.

Notação: N1=Falante nativo nº 1; PB=Português Brasileiro; ESC= Ensino Superior Completo.

7.2.2 Método

7.2.2.1 Gravação dos estímulos e coleta dos dados

A partir dos resultados do piloto, não foi possível chegar a conclusões mais precisas sobre que pista se mostrou mais relevante ou eficaz na discriminação do contraste surdo-sonoro. Além disso, a sentença-veículo utilizada no experimento piloto “say ___ *promptly*” poderia oferecer alguma

dificuldade para informantes com menos experiência no inglês, como é o caso do grupo de alunos do nível básico, que não conhecem a palavra “promptly”. Para que isso não viesse a comprometer o experimento, optou-se por utilizar a sentença-veículo “say ___ to me”. Para a produção, ou seja, para a gravação dos estímulos, inserir a palavra-alvo em uma sentença faz com que se tenha mais controle da duração dos segmentos envolvidos (GUION, 1995), um dos dados essenciais para o presente experimento. Para a percepção, como relatam Zimmer e Alves (2007), a presença de um contexto adjacente, como em “to me”, fornece aos informantes mais pistas para a identificação do segmento alvo e se configura como uma situação comunicativa mais próxima da experienciada pelos informantes em situações como aulas de inglês. Nas aulas de língua, embora existam exercícios de pronúncia que primem pela repetição isolada de palavras e sons, tal prática tem como objetivo final que os aspectos praticados isoladamente sejam incorporados às palavras-alvo e sentenças-alvo produzidas pelos informantes.

Com a mudança da sentença-veículo, foi escolhido outro locutor (diferente do piloto) para realizar a gravação dos dados. O locutor era da Carolina da Norte, como o GC, tinha 30 anos na época da gravação e assim como os outros informantes do GC, estava estudando português como língua estrangeira no CELIN-UFPR, completando 80 horas de curso. Ele não possuía conhecimento de outra língua estrangeira além do português. Foram gravadas 3 repetições das sentenças-veículo afim de que se pudesse realizar a duração relativa dos parâmetros utilizados para o experimento, medidas importantes para a realizar a manipulação dos dados. Os estímulos gravados pelo locutor foram coletados em cabine com tratamento acústico localizado na Universidade Federal do Paraná (microfone Shure KSM 27, com a taxa de amostragem do sinal a 44100Hz).

7.2.2.2 Coleta dos dados dos informantes do GE E GC.

Antes de realizar os testes (Identificação e Discriminação), os informantes do GE e GC preencheram os questionários (ver Anexo 10.1 e 10.2)

para que algumas informações pudessem ser controladas: idade, naturalidade, escolaridade, conhecimento de outra LE, horas de ensino formal no Celin-UFPR. Após a obtenção das informações pessoais, foram realizadas 10 seções de coleta. A coleta dos dados para os testes de percepção foi feita no laboratório multimídia do Celin-UFPR (centro de línguas onde eram ministradas as aulas dos aprendizes que fizeram parte do estudo). A sala contava com 10 computadores com fones. Devido o tamanho da sala, só foi possível realizar a coleta com grupos de 10 informantes. A partir da disponibilidade de horário dos informantes, foram realizadas 3 seções com o GE básico, 2 com o GE Pré-Intermediário, 4 com o GE Intermediário e 1 com o GC. As instruções sobre os testes foram dadas em português para o GE e em inglês para o GC. Em média, cada informante levou 25 minutos para realizar todos os testes. Ambos os testes foram realizados na mesma seção (primeiramente, foi realizado o teste de discriminação e, em seguida, o teste de identificação). Antes de todas as seções foi conduzido um teste de familiarização, para evitar que os informantes desempenhassem algum problema na realização dos testes principais. Mais informações sobre o teste de familiarização serão discutidas na seção 7.2.2.4. É importante mencionar que toda a coleta foi realizada pela pesquisadora.

7.2.2.3 Corpus

Do mesmo modo que no piloto, este experimento conta com seis palavras do inglês. As palavras do *corpus* obedecem ao padrão CVC, sendo V a vogal [æ] e a C final sendo o segmento oclusivo surdo ou sonoro, a saber: *cap* e *cab*; *bat* e *bad*; *back* e *bag*. Tais dados passaram por uma manipulação, uma vez que um dos objetivos do experimento é verificar se existe uma pista acústica que seja mais saliente, ou seja, que auxilie maximamente na percepção de oclusivas surdas e sonoras em fim de palavra. Assim, para este experimento foram utilizados os parâmetros duração de vogal, [æ], que antecede o segmento oclusivo final e o vozeamento. Para este experimento, optou-se por não verificar a influência do *burst* como parâmetro acústico

utilizado na discriminação do contraste surdo-sonoro, pois: a) conforme relata a literatura (KENT e READ, 1992), o *burst* não é uma pista imprescindível na caracterização das oclusivas e não possui uma distribuição homogênea em fim de palavra, ou seja, o *burst* nem sempre é realizado em contexto final; b) embora o *burst* tenha se mostrado uma pista importante na percepção do contraste surdo-sonoro de alguns pontos de articulação em algumas línguas (ERNESTUS e BAAYEN, 2006; WARNER *et al.*, 2004; BROERSMA, 2005), ele não foi encontrado com frequência no conjunto de dados de produção deste experimento. Ou seja, na análise dos dados obtidos do locutor americano, não se verificou a presença do *burst*. A presença do *burst* se deu em poucos dados, o que não justificou a inserção dele como um parâmetro a ser estudado.

Semelhante ao piloto, os parâmetros vozeamento e duração de vogal foram manipulados de modo a se obter a duração média de ambas as pistas. As durações médias foram extraídas de todas as repetições. A manipulação dos dados foi realizada para que os sujeitos pudessem identificar e discriminar surdas e sonoras. Cada teste, de identificação e de discriminação, contou com um total de 1.620 respostas do GE (3 manipulações X 3 repetições X 3 pontos de articulação X 60 informantes do GE) e 216 respostas (3 manipulações X 3 repetições X 3 pontos de articulação X 8 informantes do GC). O esquema da Figura 24 mostra a organização dos dados.

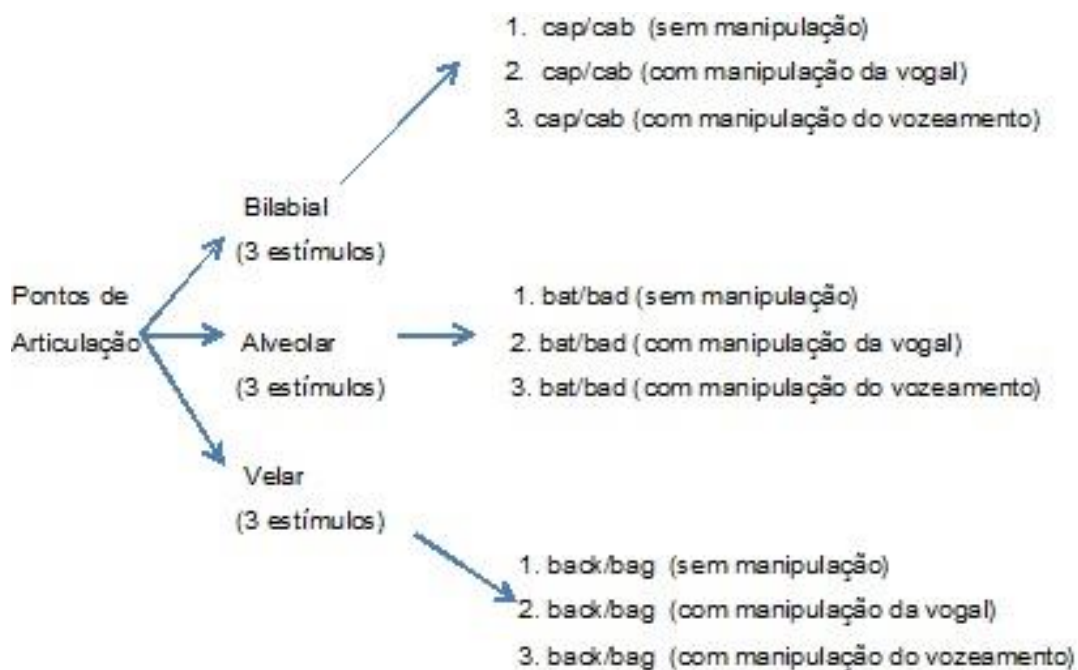


Figura 24 Esquema informativo dos dados dos experimentos de identificação e discriminação.

Os dados da Figura 24 que apresentam os números 2 e 3 possuem manipulação. Ao dizer “manipulação de vogal”, entende-se que o informante só escuta a vogal, sem o vozeamento e, por “manipulação de vozeamento”, compreende-se a apresentação do dado com a manipulação da vogal, para que ela se aproxime à duração relativa de uma vogal que antecede uma oclusiva surda, e o vozeamento.

Nas Figuras 25 e 26 é possível visualizar os dados sem a manipulação. No espectograma da Figura 27 nota-se o estímulo “bad” com a manipulação da duração de vogal. A duração da vogal é semelhante à da Figura 25, que apresenta o estímulo “bad”, o qual não sofreu manipulação.

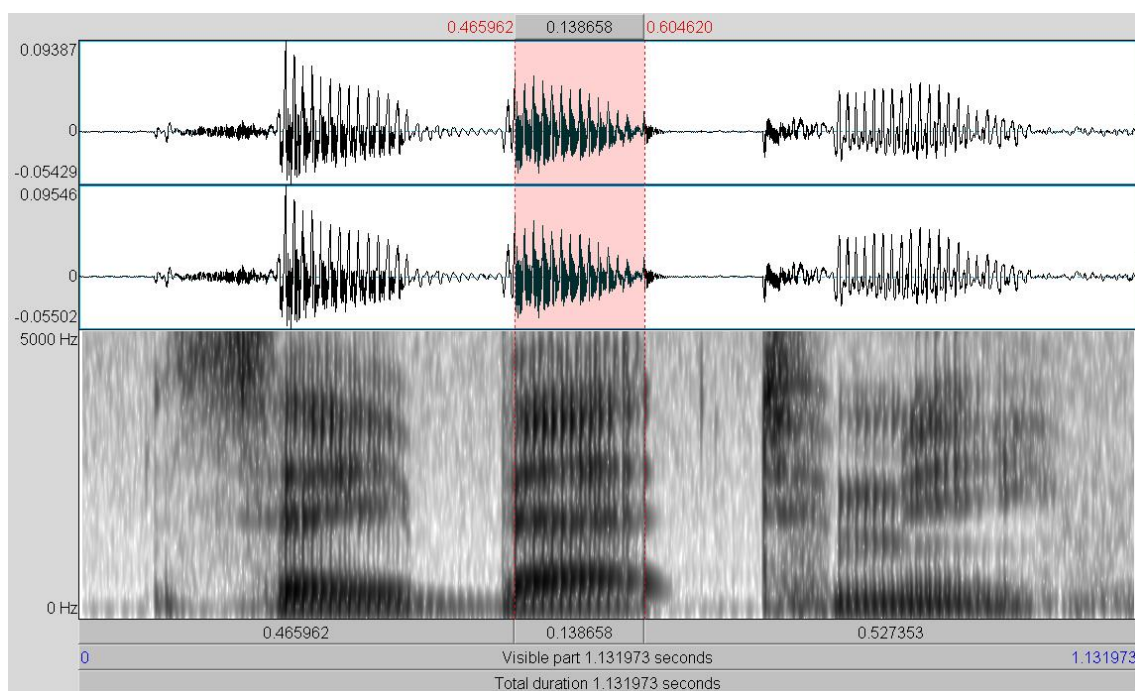


Figura 25 Espectrograma do estímulo bat. Entre as linhas pontilhadas a duração da vogal: 0.1386 ms.

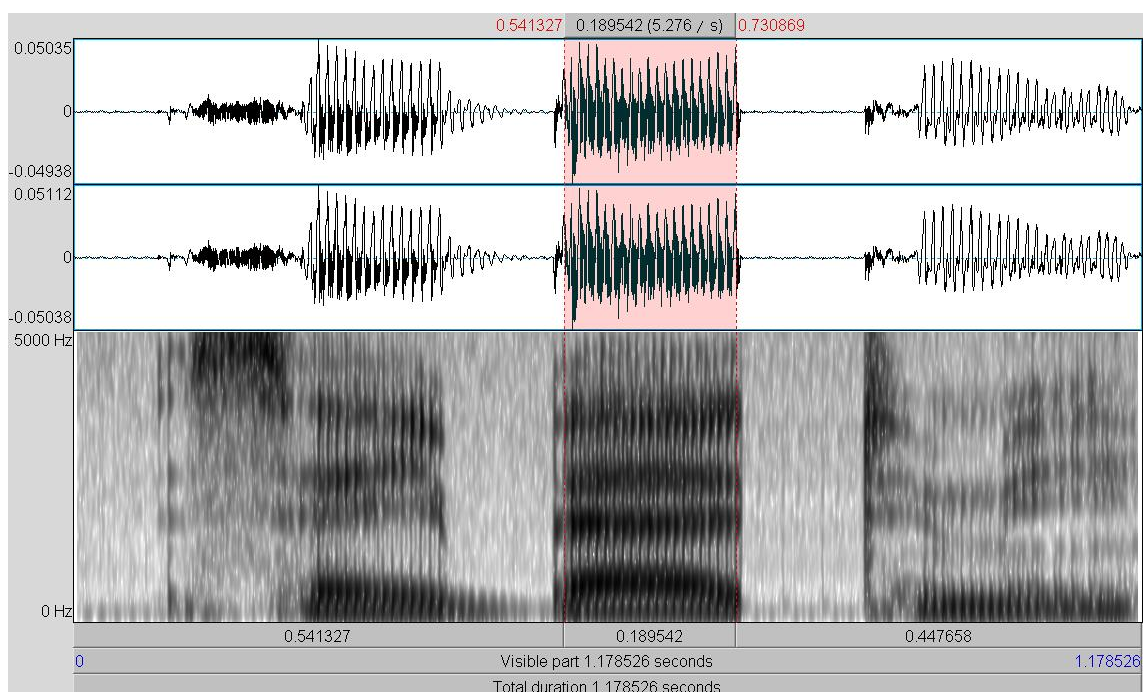


Figura 26 Espectrograma do estímulo bad sem a manipulação. Entre as linhas pontilhadas a duração da vogal: 0.1895 ms.

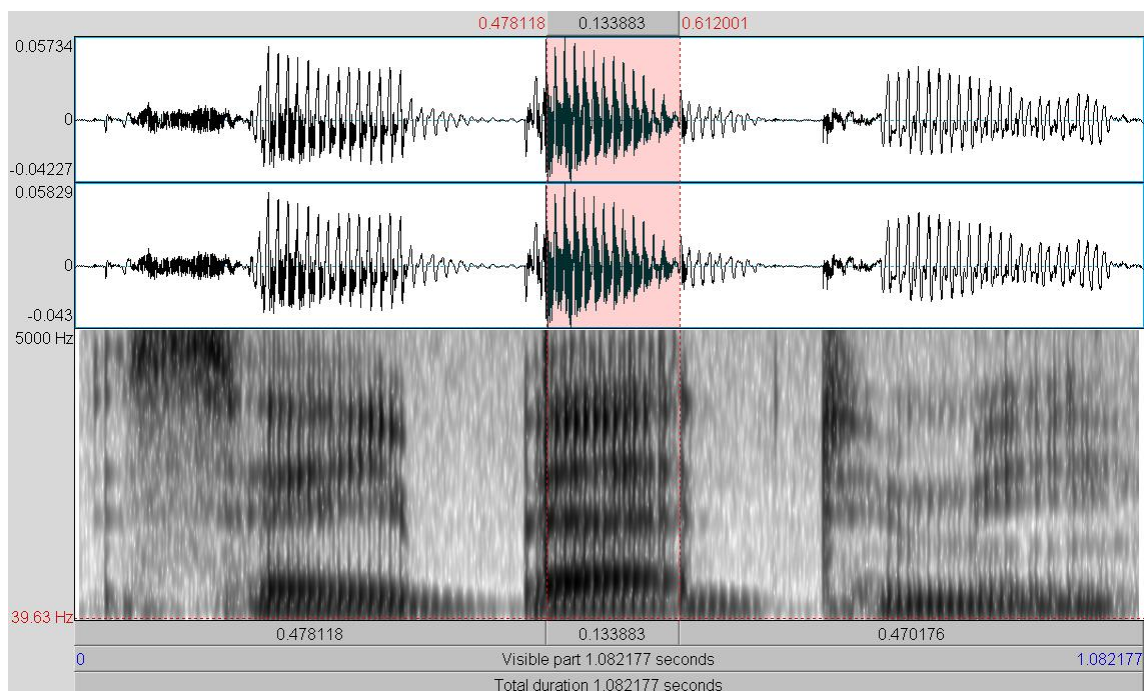


Figura 27 Espectrograma do estímulo bad com a manipulação da duração da vogal. Entre as linhas pontilhadas a duração da vogal: 0.1338 ms.

7.2.2.4 Teste de Discriminação e Identificação

Os sujeitos foram submetidos a um pequeno teste de familiarização com outro conjunto de palavras que também formam pares mínimos em inglês (*bop* e *bob*; *pat* e *pad*; *rack* e *rag*) para que se familiarizassem com o experimento. Nenhum sujeito demonstrou qualquer tipo de estranhamento com relação ao teste ou com o procedimento de escolha dos itens.

Após o teste de familiarização, os sujeitos foram submetidos aos testes de discriminação e identificação, nesta ordem. Ambos os testes contaram com o mesmo conjunto de dados: *cap* e *cab*; *bat* e *bad*; *back* e *bag*. O teste de discriminação teve o mesmo tipo de apresentação de dados que o piloto. Optou-se por manter o experimento de discriminação, pois ele trabalha apenas com o estímulo auditivo, sem auxílio de nenhum outro recurso (e.g. visual). Além disso, nesse tipo de teste, o informante deve decidir se os dados que lhe são apresentados são “=” (iguais) ou “*” (diferentes). Cada informante ouve 2

sentenças-veículos. Neste tipo de teste, o sujeito possui duas categorias que, embora não sejam completamente fechadas, oferecem duas opções possíveis (ou seja, ao discriminar o dado como igual ou diferente o informante procura relacionar os dados e notar alguma semelhança ou diferença entre eles, utilizando pistas diferentes). O informante não possui uma variedade de itens a serem selecionados. O modo de apresentação do teste de discriminação pode ser visualizado abaixo nas Figuras 28 e 29.

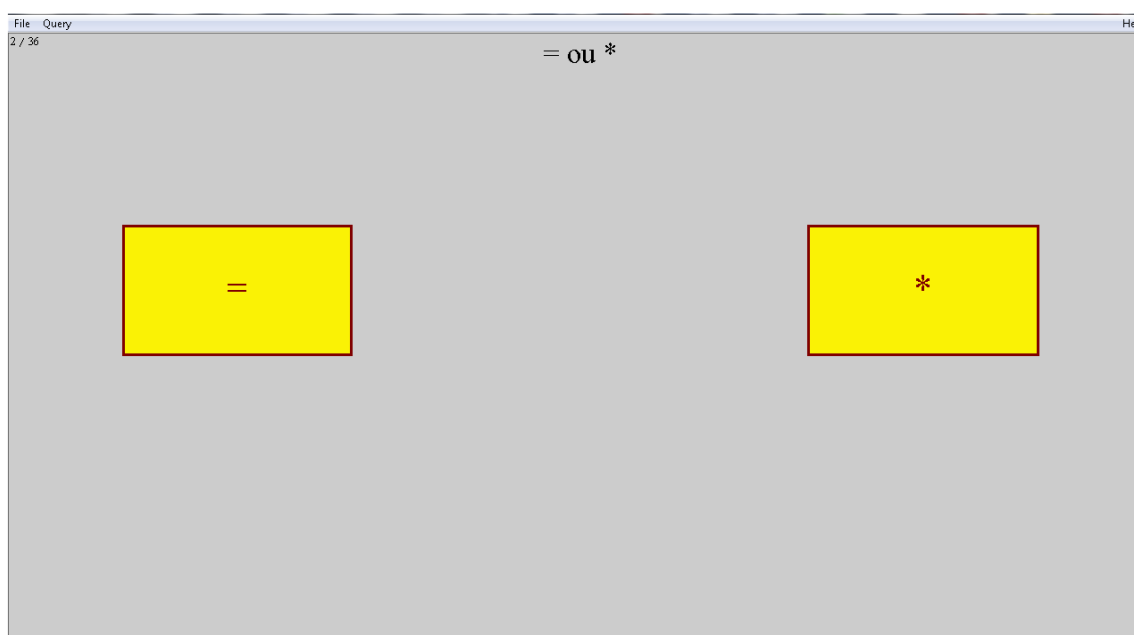


Figura 28 Tela do Praat visualizada pelos informantes no teste de discriminação.

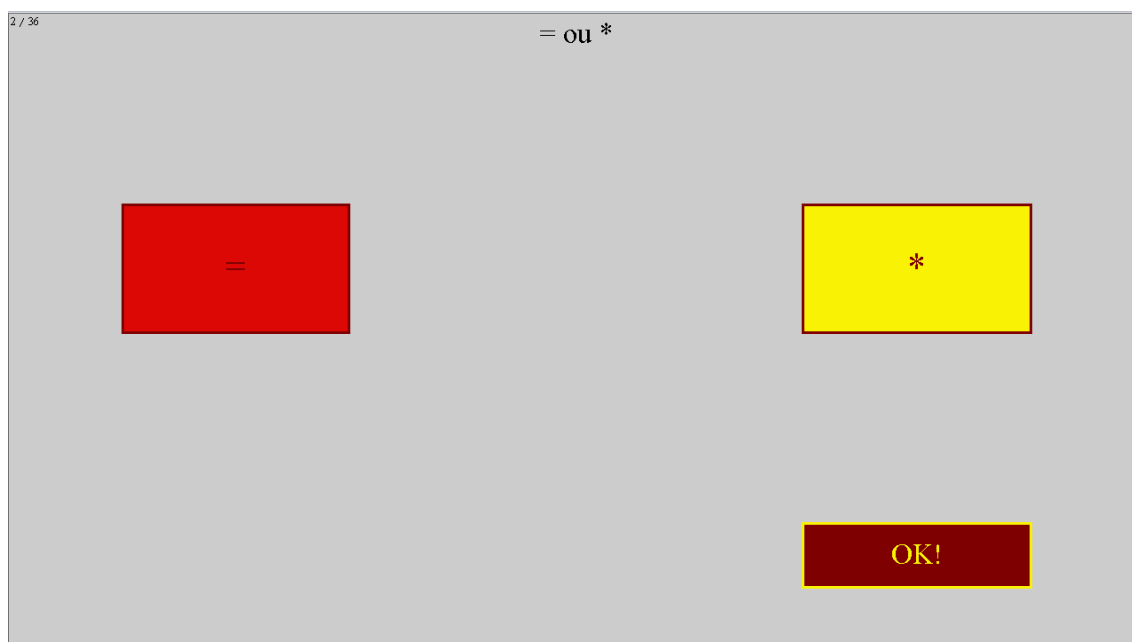


Figura 29 Tela do Praat visualizada pelos informantes no teste de discriminação ao clicarem na resposta que julgassem correta.

Para o teste de identificação, também se utilizou o programa Praat. Ao contrário do teste de discriminação, no qual os informantes recebiam dois estímulos auditivos, os sujeitos receberam apenas um estímulo e foram orientados a clicar na opção (na tela do programa) que considerassem mais adequada. Decidiu-se inserir mais um tipo de teste (ao contrário do piloto que contou apenas com o teste de discriminação) para observar se, quando os informantes tinham a possibilidade de visualizar as palavras-alvo na tela, a percepção do contraste era feita de modo mais acurado, ou seja, se o tipo de teste tem, de fato, alguma influência nos resultados perceptuais. Conforme já foi mencionado anteriormente, embora o teste de identificação seja superficialmente semelhante ao de discriminação (o informante recebe um estímulo auditivo e precisa optar por uma das opções que lhe são apresentadas na tela), o teste de identificação introduz mais uma variável ao experimento: o dado escrito. Assim, o informante não recebe apenas o dado auditivo, mas também o escrito, o que pode auxiliar na identificação dos dados, mas, ao mesmo tempo, pode enviesar os resultados uma vez que os

informantes não devem perceber apenas o som, mas também precisam estabelecer uma relação entre esse e sua ortografia correspondente. Apesar da possibilidade de enviesamento, este experimento se propôs a realizar uma investigação desse tipo de teste justamente para levantar questionamentos sobre metodologias experimentais e a influência dessas nos resultados finais de um estudo. Nas Figuras 30 e 31, é possível ver o modo como os informantes visualizaram os estímulos no teste de identificação.

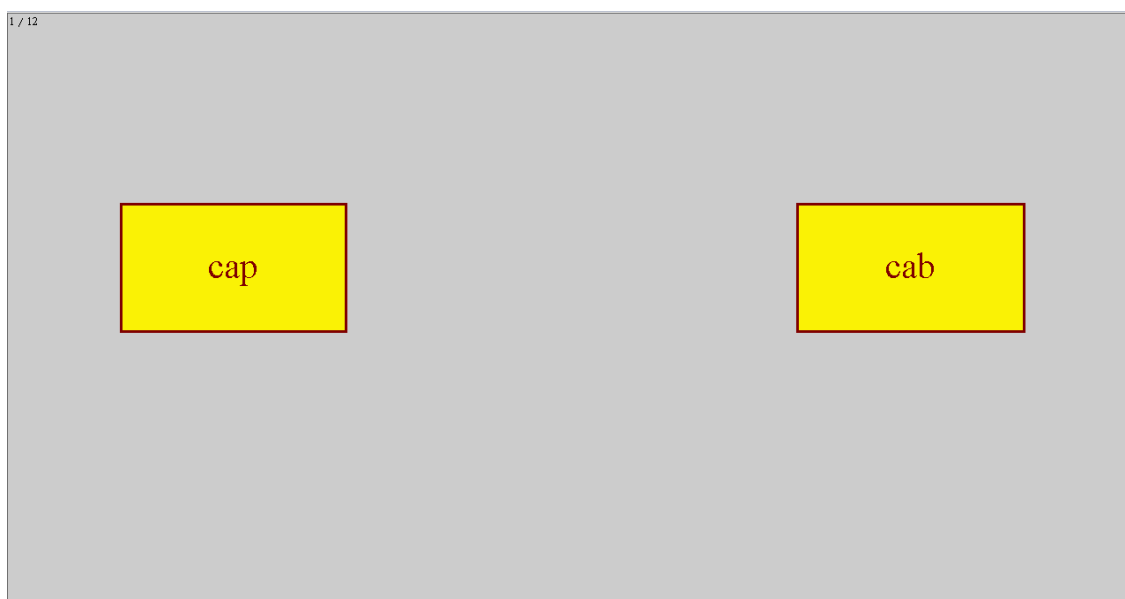


Figura 30 Tela do Praat visualizada pelos informantes ao receberem o estímulo auditivo no experimento de identificação.

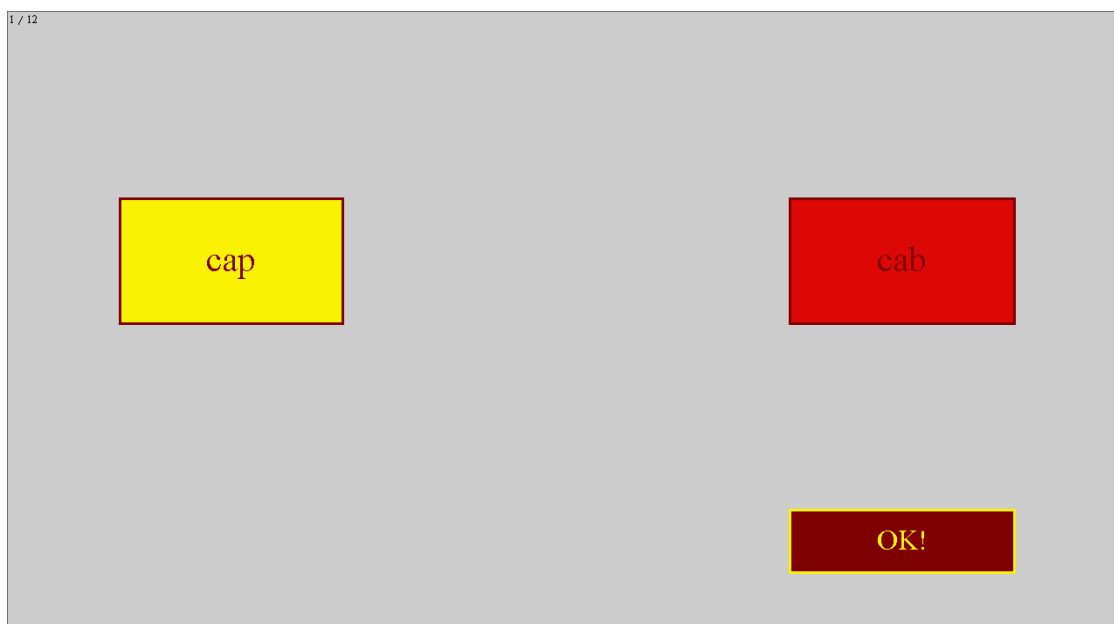


Figura 31 Tela do Praat visualizada pelos informantes no experimento de identificação ao clicarem na resposta que julgassem correta.

Ao final de cada conjunto de dados (3 repetições de cada par: cap e cab; bat e bad; back e bag), uma tela com a instrução “Descanse um pouco. Clique para continuar”⁵⁹ era apresentada aos sujeitos. Esta tela foi inserida para que os informantes não fossem sobrecarregados pela sequência dos dados, fazendo com que isso se tornasse outra variável no experimento.

Os sujeitos realizaram os dois testes em sequência, primeiro o de discriminação e, em segundo lugar, o de identificação.

7.2.3 Análise Estatística

A análise estatística para este experimento foi rodada no programa SPSS 18.0 (SPSS, Chicago, IL), sendo que o nível de significância estabelecido foi de $\alpha = .05$.

Foi feita uma inspeção inicial para verificar se os dados obedeciam a uma distribuição normal ou não. A partir da análise das medidas de Simetria

⁵⁹ Como o experimento continha dados do inglês e falantes nativos também fizeram parte do experimento, optou-se por colocar a instrução em inglês “*Rest a bit, Click to continue*”.

(*Swedess*) e Curtose (*Curtosis*), pode-se observar que apenas parte dos dados cumpriam os pressupostos de normalidade da distribuição. Assim, testes não-paramétricos foram utilizados. O detalhamento sobre a natureza dos testes e de como eles se encaixam no experimento pode ser visto abaixo:

- Teste de Wilcoxon: para designs intrassujeitos, para amostras relacionadas (isto é, comparações de resultados dentro de um mesmo grupo) com variáveis de dois níveis;
- Teste de Mann-Whitney: para comparações intersujeitos, para amostras independentes (grupos diferentes) com variáveis de dois níveis (dois grupos);
- Teste de Kruskal-Wallis: para comparações intersujeitos para amostras independentes (grupos diferentes) com variáveis de mais de 3 níveis.
- Teste de Spearman: teste não paramétrico que avalia o quão boa é a relação dentre duas variáveis.

7.2.3.1 Resultados obtidos para o teste de discriminação

Serão comentados abaixo os resultados para o teste de discriminação a partir dos resultados do teste Kruskal-Wallis, Friedman e Wilcoxon. O teste Kruskal-Wallis foi utilizado para verificar se os três grupos do GE se comportavam da mesma maneira. Descobriu-se que tal fato se verifica para o teste de discriminação, mas não para o de identificação. Desta maneira, para que os dados recebessem a mesma análise, optou-se por analisar os grupos separadamente.

Como o experimento contou com 3 manipulações por ponto de articulação, rodou-se primeiramente, o teste Friedman para verificar se havia alguma relação entre os resultados. Uma vez que os resultados foram

significativos para todos os níveis de proficiência e pontos de articulação, passou-se à aplicação do Wilcoxon, para verificar em quais pares a diferença na porcentagem de acertos era significativa. Abaixo, serão apresentadas as tabelas com os resultados do teste de discriminação por nível de proficiência. Para que se mantenha um padrão, os dados não manipulados serão nomeados de padrão DC; os que possuem a duração de vogal preservada serão denominados de padrão DV; aqueles que possuem apenas o parâmetro duração de vozeamento receberão o título de padrão V.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	44	18	29	43	38	16	45	38	21
% \bar{X}	73	30	48	72	63	27	75	63	35
% DP	27	31	40	32	31	27	36	31	31

Tabela 5 Resultados da média e percentual de acertos para o teste de discriminação para o GE Básico. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.

Percebe-se, na Tabela 5, que no GE Básico o padrão sem manipulação obteve o maior percentual de acerto (em comparação com os dados manipulados) para todos os pontos de articulação: bilabial, alveolar e velar. O Teste de Wilcoxon, utilizado para verificar diferenças estatísticas pareadas entre os padrões, demonstra que a diferença entre o percentual de acerto entre DC (73%) e DV (30%) e DC (73%) e V (48%) para as bilabiais é significativa, tendo como valores ($Z = -3,291^a$, $p = 0,001$) e ($Z = -2,391^a$, $p = 0,017$), respectivamente. Assim, é possível dizer que os informantes do GE Básico, para as bilabiais, têm um maior percentual de discriminação quando o dado não sofre manipulação do que quando o dado passa por alterações na duração da vogal ou no vozeamento. No entanto, o teste mostrou que a diferença entre

o padrão DV (30%) e V (48%) não é significativa ($Z=-1,565^b$, $p=0,118$), não sendo possível afirmar, portanto, que os informantes desse grupo discriminam melhor quando percebem o dado com o vozeamento preservado ao invés da duração da vogal.

Para as alveolares e velares, o cenário é semelhante no que diz respeito à diferença entre os percentuais de acerto de DC e V. Para as alveolares, DC (72%) e V (27%), correspondendo a ($Z= -2,991^a$, $p= 0,003$) e, para as velares, DC (75%) e V (35%), obtendo valores de ($Z= -2,935^a$, $p=0,003$). Além disso, é interessante observar que a diferença entre os padrões DV e V é significativa tanto para alveolares quanto para velares, sendo que para as alveolares DV (63%) e V (27%), tendo valores ($Z=-3,131^a$, $p=0,002$) e para as velares o percentual de acerto para DV (75%) e V (35%), possuindo valores ($Z=-2,474^a$, $p=0,013$). A partir do Teste de Wilcoxon é possível dizer que não só os informantes pareceram ter mais dificuldade em discriminar surdas e sonoras quando ouviam o padrão V em relação à DC, mas também que aquele padrão foi discriminado com menos acurácia quando comparado com DV, ou seja, a pista duração de vogal é importante na discriminação do contraste sonoro para oclusivas alveolares e velares.

Os resultados para o grupo GE Pré-Intermediário se mostram distribuídos de maneira semelhante.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	53	28	39	50	45	19	49	39	27
% \bar{X}	88	47	65	83	75	32	82	65	45
% DP	24	31	35	31	30	22	29	31	32

Tabela 6 Resultados da média e percentual de acertos para o teste de discriminação para o GE Pré-Intermediário. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração de Vogal; V: Vozeamento.

Novamente, os resultados apontam que o padrão DC foi o melhor discriminado para todos os pontos de articulação. O Teste de Wilcoxon revelou, para as bilabiais, que a diferença entre o padrão DC (88%) e DV (47%) é significativa ($Z=-3,167^a$, $p=0,002$) e que a distinção entre os percentuais de DC (88%) e V (65%) também é ($Z= -2,508^a$, $p=0,012$) e, confirma, portanto, que o padrão DC foi mais facilmente discriminado do que DV e V. Mas, do mesmo modo que para o GE Básico, não foi possível encontrar a diferença entre DV e V não é significativa, demonstrando que os informantes do GE Pré-Intermediário discriminaram ambos os padrões de modo semelhante.

Os resultados para alveolares e velares também foi parecido com os do GE Básico. Para as alveolares, a diferença entre DC (83%) e V (32%), correspondendo a ($Z=-3,700^a$, $p= 0,000$) e, para as velares, DC (82%) e V (45%), obtendo valores de ($Z= -3,458^a$, $p=0,001$). Somando-se a isso, observa-se também que diferença entre os padrões DV e V é significativa tanto para alveolares quanto para velares, sendo que para as alveolares DV (75%) e V (32%), tendo valores ($Z=-3,572^a$, $p=0,000$) e para as velares o percentual de acerto para DV (65%) e V (45%), possuindo valores ($Z=-2,142^a$, $p=0,032$). Diferentemente do grupo GE Básico, a relação entre os resultados para as velares do padrão DC (82%) e DV (65%) é significativa ($Z=-2,500^a$, $p=0,012$). Assim, é possível apontar que os informantes do GE Pré-Intermediário pareceram ter mais dificuldade em discriminar surdas e sonoras quando ouviam o padrão V em relação a DC, mas também que aquele padrão foi discriminado com menos acurácia quando comparado com DV, ou seja, a pista duração de vogal é importante na discriminação do contraste sonoro para oclusivas alveolares e velares.

De modo distinto, os resultados não são os mesmos para o GE Intermediário, conforme será relatado na sequência.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	52	25	36	48	41	23	51	47	23
% \bar{X}	87	42	60	80	68	38	85	78	38
% DP	28	33	24	22	37	41	27	26	40

Tabela 7 Resultados da média e percentual de acertos para o teste de discriminação para o GE Intermediário. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.

Começando pelas bilabiais, é possível dizer que todas as diferenças entre os percentuais de acerto para todos os padrões, a saber, DC e DV, DC e V, DV e V, são significativas. Ou seja, os informantes do GE Intermediário foram capazes de discriminar com mais acurácia o padrão DC, do que DV e V. De modo distinto dos outros GE, quando se compara os acertos entre DV (42%) e V (60%), a diferença é significativa ($Z = -1,977^b$, $p = 0,048$) e isso revela, portanto, que os informantes deste grupo discriminaram oclusivas surdas e sonoras melhor quando eram expostas à pista vozeamento do que duração de vogal.

No entanto, para as alveolares, o único padrão que apresentou uma diferença significativa foi V (38%) em relação à DC (80%), cujos valores no teste de Wilcoxon foram $Z (-2,865^a)$ e $p (0,004)$. O teste de Wilcoxon ($Z = -1,645^a$, $p = 0,100$), não revelou nenhuma diferença significativa entre DC (80%) e DV (68%); do mesmo modo, o teste ($Z = -1,952^a$, $p = 0,051$) revelou que não há diferença significativa entre DV (68%) e V (38%). Ou seja, só é possível dizer que os informantes do GE Intermediário discriminaram melhor o padrão DC em relação ao V.

Para as velares, a diferença entre o percentual de acerto entre os padrões DC (85%) e V (38%) é significativa ($Z = -2,914^a$, $p = 0,004$) e entre DV (78%) e V (38%) também ($Z = -3,042^a$, $p = 0,002$). No entanto, de modo semelhante ao GE Básico, a diferença entre os padrões DC (85%) e DV (78%)

não é significativa ($Z = -0,775^a$, $p = 0,439$), ou seja, os informantes acertaram o DC e o DV na mesma proporção.

Parte-se agora para a descrição dos resultados do teste de discriminação para o GC.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	24	24	24	24	24	24	24	24	24
\bar{X}	24	24	21	24	23	13	23	23	6
% \bar{X}	100	100	87	100	95	54	95	95	25
% DP	0	0	17	0	12	25	12	12	33

Tabela 8 Resultados da média e percentual de acertos para o teste de discriminação para o GC. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.

	Bilabial			Alveolar			Velar		
	DC x DV	DC x V	DV x V	DC x DV	DC x V	DV x V	DC x DV	DC x V	DV x V
Z	,000	-1,732	-1,732	-1,000	-2,232	-2,232	,000	-2,414	-2,333
P	1,000	,083	,083	,317	,026	,026	1,000	,016	,020

Tabela 9 Resultados do teste de Wilcoxon para o teste de discriminação para o GC. Z: Wilcoxon; p: p valor.

A partir dos resultados expostos nas Tabelas 8 e 9, pode-se perceber que a diferença entre o percentual de acerto para as alveolares entre DC (100%) e v (54%) é significativa ($Z = -2,232$, $p = 0,026$) e que a distinção entre os valores de DV (95%) e V (54%) também é ($Z = -2,232$, $p = 0,026$). Um cenário semelhante pode ser articulado para as velares, sendo que a diferença entre o percentual de acerto também foi maior para o padrão DC e DV. A diferença entre DC (95%) e V (25%) é significativa ($Z = -2,414$, $p = 0,016$). Do mesmo modo, o teste de Wilcoxon destaca a distinção entre os valores de DV (95%) e V (25%) como significativa ($Z = -2,333$, $p = 0,020$). Isso significa que o padrão V

dificultou a discriminação do contraste surdo-sonoro em posição final tanto para alveolares como para velares.

7.2.3.2 Resumo dos resultados do teste de discriminação: relação entre GE e GC e pontos de articulação.

É possível observar que embora os resultados sejam diferentes para entre os 3 níveis de proficiência do GE e para o GC, algumas considerações são passíveis de serem feitas. É importante ressaltar que a nomenclatura “significativo” e “não significativo” no Quadro 4 está relacionada com o fato do padrão ter ou não sido melhor reconhecido pelos informantes. Assim, serão analisados os padrões que passaram por manipulação: DV (duração de vogal preservada) e V (vozeamento).

Ponto de articulação	Padrão DC x DV	Padrão DC x V	DV X V
Bilabiais	Não Significativo: GC	Não Significativo: GC	Significativo: GE Intermediário
Alveolares	Não Significativo: todos os níveis	Significativo: Todos os níveis	Não Significativo: GE Intermediário
Velares	Significativo: GE Pré-Intermediário	Significativo: Todos os níveis	Significativo: Todos os níveis

Quadro 4 Resumo dos resultados para o teste de discriminação.

É possível notar que, de modo geral, o padrão DV, que representa o dado cuja duração da vogal foi mantida, foi mais bem discriminado quando comparado com o padrão V, o qual se encontra relacionado apenas à presença do vozeamento. O padrão V se mostrou significativo somente para o GE Intermediário para as bilabiais.

Entre os objetivos do experimento esperava-se que os informantes discriminassem melhor os dados quando esses não tivessem passado por nenhuma manipulação. Tal objetivo foi confirmado para o teste de discriminação para alguns pontos de articulação e níveis, conforme pode-se notar no Quadro 4. Foi observado, também, que os padrões DV e V foram discriminados, o que corrobora estudos como os de Warner *et.al.* (2004) e Kleber *et.al.* (2010): os informantes são capazes de utilizar pistas acústicas existentes nas suas L1, mas em posições diferentes das existentes na L2 (começo, meio ou fim de palavra) para discriminar os dados da L2 em questão. Além disso, os autores encontraram em seus estudos a duração de vogal como sendo a pista mais utilizada pelos informantes para discriminar o contraste surdo-sonoro. Embora não seja possível afirmar isso para todos os dados, tal fato se verifica para as alveolares e velares na maioria dos grupos.

Além disso, também foi verificado que o ponto de articulação parece influenciar o percentual de acerto do contraste surdo-sonoro em posição final. É importante ressaltar que todos os níveis do GE e o GC demonstram ter uma maior dificuldade de reconhecer o padrão V para os dados que apresentavam velares em posição final. Como pode ser visto nas Tabelas 5, 6, 7 e 8, os dados que continham o padrão V obtiveram uma maior porcentagem de desvio padrão. Isso pode ressaltar, por um lado, a dificuldade encontrada pelos informantes em utilizar este padrão para a discriminação do padrão surdo-sonoro. Embora o desvio padrão tenha sido alto para alguns dados do padrão DV, as ocorrências incidiram sobre uma menor quantidade de dados quando comparadas ao padrão V. Nas Figuras 32, 33 e 34 pode-se ter uma melhor visualização da média de acerto dos grupos GE e GC para o teste de discriminação.

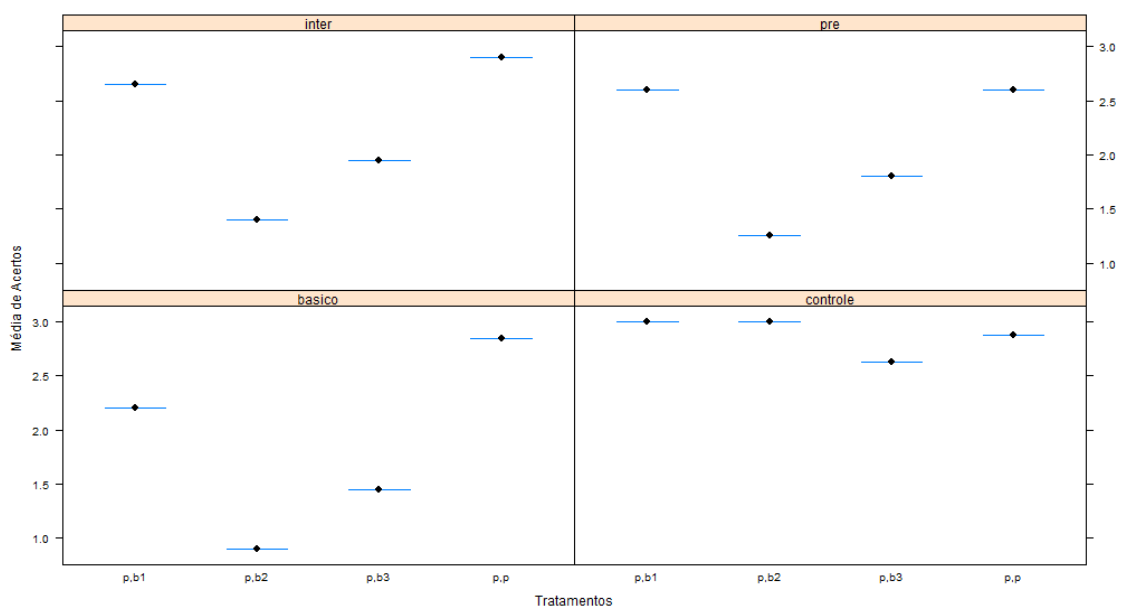


Figura 32 Gráfico com a média de acertos das bilabiais do teste de discriminação. Notação: p,b1= Dado Completo ;p,b2= padrão Duração de Vogal; p,b3= Vozeamento.

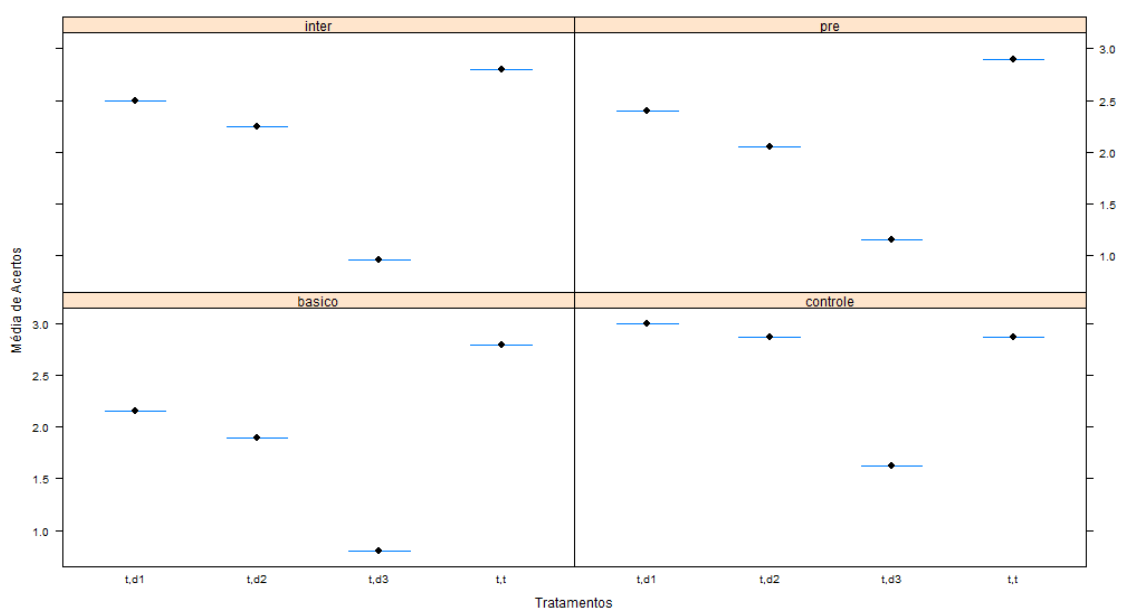


Figura 33 Gráfico com a média de acertos das alveolares do teste de discriminação. Notação: t,d1= Dado Completo ;t,d2= padrão Duração de Vogal; t,d3= Vozeamento.

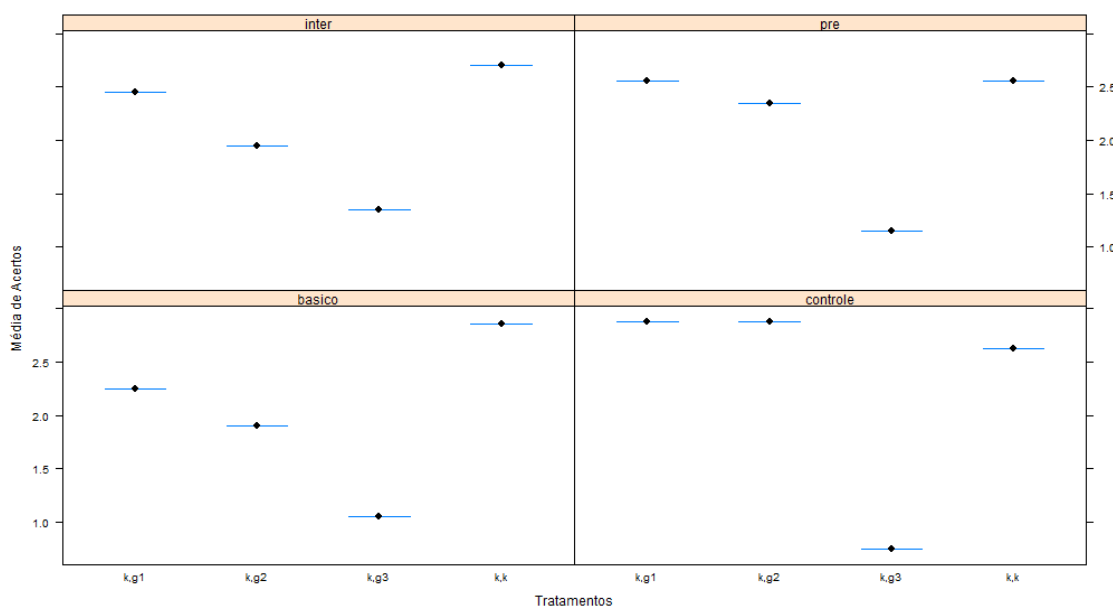


Figura 34 Gráfico com a média de acertos das velares do teste de discriminação. Notação: k,g1= Dado Completo ;k,g2= padrão Duração de Vogal; k,g3= Vozeamento.

Passa-se agora ao relato dos resultados para o teste de identificação.

7.2.3.4 Resultados obtidos para o teste de identificação.

Após a discussão dos resultados obtidos para o teste de discriminação, que tinha como metodologia a apresentação de dois estímulos auditivos e o informante deveria classificar os dados em igual ou diferente, passa-se à análise dos resultados do teste de identificação.

É importante lembrar como esse teste foi conduzido. Diferentemente do teste de discriminação, os sujeitos recebiam um estímulo, mas precisavam optar por uma das duas palavras que apareciam na tela do programa Praat, ou seja, trata-se de uma espécie de “etiquetamento” dos dados.

Do mesmo modo que para o teste de discriminação, os dados que não sofreram manipulação tiveram um percentual de acerto alto. Isso pode indicar que os dados não manipulados não foram afetados pela metodologia, uma vez que tal fato foi verificado para o teste de discriminação também.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	50	23	41	55	40	36	59	48	25
% \bar{X}	83	38	68	91	66	60	98	80	42
% DP	27	35	31	23	32	31	7	29	36

Tabela 10 Resultados da média e percentual de acertos para o teste de identificação para o GE Básico. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.

Percebe-se, na Tabela 10, que no GE Básico o padrão sem manipulação obteve o maior percentual de acerto (em comparação com os dados manipulados) para todos os pontos de articulação: bilabial, alveolar e velar. O Teste de Wilcoxon, utilizado para verificar diferenças estatísticas pareadas entre os padrões, demonstra que a diferença entre o percentual de acerto entre DC (83%) e DV (38%) para as bilabiais é significativa, tendo como valores ($Z = -3.094^b$, $p = 0,002$). Assim, é possível dizer que os informantes do GE Básico, para as bilabiais, apresentaram um maior percentual de identificação quando o dado não havia sofrido manipulação do que quando o dado passa por alterações na duração da vogal. Além disso, ao contrário do teste de discriminação, o Teste de Wilcoxon para o teste de identificação mostrou que a diferença entre o padrão DV (38%) e V (68%) também é significativa ($Z = -2.738^c$, $p = 0,006$), sendo possível afirmar, portanto, que os informantes desse grupo identificam melhor quando percebem o dado com o vozeamento preservado ao invés da duração da vogal.

Para as alveolares e velares, o cenário é semelhante ao das bilabiais com relação à diferença no percentual de acerto ser significativa entre DC (91%) e DV (66%), correspondendo a ($Z = -2.411^b$, $p = 0,016$) e, para as velares, DC (98%) e DV (80%), obtendo valores de ($Z = -2.636^b$, $p = 0,008$). Além disso, é interessante observar que a diferença entre os padrões DC e V é significativa

tanto para alveolares quanto para velares, sendo que para as alveolares DC (91%) e V (60%), tendo valores ($Z = -2.823^b$, $p = 0,005$) e para as velares o percentual de acerto para DC (98%) e V (42%), apresenta valores ($Z = -3.682^b$, $p = 0,000$). Por outro lado, as velares parecem ter um comportamento mais parecido com as bilabiais quando se observa que a diferença entre o percentual de acerto entre DV (80%) e V (42%) é significativa ($Z = -2.759^b$, $p = 0,006$). Por outro lado, a diferença entre DV (66%) e V (60%) não é significativa para as alveolares ($Z = -.544^b$, $p = 0.586$).

A partir do Teste de Wilcoxon é possível dizer que não só os informantes pareceram ter mais dificuldade em identificar surdas e sonoras quando ouviam o padrão V em relação a DC, mas também que o padrão V foi identificado com menos acurácia quando comparado com DV, ou seja, a pista duração de vogal é importante na identificação do contraste sonoro, ao menos para oclusivas velares. De modo curioso, o cenário é distinto para as bilabiais, uma vez que os informantes sentiram mais dificuldade em reconhecer a distinção surdo-sonora quando não tinham a pista vozeamento.

Os resultados para o grupo GE Pré-Intermediário se mostram distribuídos de maneira semelhante ao GE Básico.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	60	39	54	59	50	44	58	56	30
% \bar{X}	100	65	90	98	83	73	96	93	50
% DP	0	34	26	7	25	27	10	13	39

Tabela 11 Resultados da média e percentual de acertos para o teste de identificação para o GE Pré-Intermediário. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; V: Vozeamento.

Ao contrário do teste de discriminação, os informantes não identificaram o padrão DC, no GE Pré-Intermediário, com mais acurácia do que os demais

padrões em todos os pontos de articulação. O Teste de Wilcoxon revelou, para as bilabiais, que a diferença entre o padrão DC (100%) e DV (65%) é significativa ($Z = -3.109^b$, $p = 0,002$), mas que a distinção entre os padrões DC (100%) e V (90%) não é significativa ($Z = -1.604^b$, $p = 0,109$). Mas, do mesmo modo que para o GE Básico, a diferença entre DV (65%) e V (90%) é significativa ($Z = -2.491^c$, $p = 0,013$). Assim como no GE Básico, os informantes parecem identificar com mais acurácia os dados com o padrão V do que DV.

Para as alveolares, o cenário também é semelhante ao do GE Básico. A diferença entre os padrões DC (98%) e DV (83%), correspondendo a ($Z = -2.460^b$, $p = 0,014$) e, entre o DC (98%) e V (73%), obtendo valores de ($Z = -2.879^b$, $p = 0,004$) é significativa. Para as velares, a distinção entre os padrões DC (96%) e V (50%) é significativa ($Z = -3.493^b$, $p = 0,000$), mas entre DC (96%) e DV (93%) não é ($Z = -1.000^b$, $p = 0,317$), uma vez que os falantes identificaram os dois padrões de modo semelhante. É interessante notar que do mesmo modo que para as bilabiais, a diferença entre os padrões DV (93%) e V (50%) é significativa ($Z = -3.250^b$, $p = 0,001$). Para as velares tal resultado parece estar atrelado ao fato dos falantes terem identificado melhor os dados quando possuíam a pista duração de vogal, em detrimento do vozeamento.

Diferentemente dos resultados do GE Básico e GE Pré-Intermediário, o GE Intermediário demonstrou novos direcionamentos, ao menos com relação às alveolares.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	60	60	60	60	60	60	60	60	60
\bar{X}	57	34	56	58	53	53	59	58	45
% \bar{X}	95	57	93	97	88	88	98	97	75
% DP	22	35	23	10	19	22	7	15	33

Tabela 12 Resultados da média e percentual de acertos para o teste de identificação para o GE Intermediário. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração de Vogal; V: Vozeamento.

Os informantes do GE Intermediário não identificaram o padrão DC com mais acurácia do que os demais padrões em todos os pontos de articulação. O Teste de Wilcoxon revelou, para as bilabiais, que a diferença entre o padrão DC (100%) e DV (65%) é significativa ($Z = -3.126^b$, $p = 0,002$), mas que a distinção entre os padrões DC (95%) e V (93%) não é significativa ($Z = -272^b$, $p = 0,785$). Mas, do mesmo modo que para o GE Básico, a diferença entre DV (57%) e V (93%) é significativa ($Z = -2.814^c$, $p = 0,005$). Assim como no GE Básico e Pré-Intermediário, os informantes parecem identificar com mais acurácia os dados com o padrão V do que DV para as bilabiais.

Para as velares, a distinção entre os padrões DC (98%) e V (75%) é significativa ($Z = -2.558^b$, $p = 0,011$), mas entre DC (98%) e DV (97%) não é ($Z = -.447^b$, $p = 0,655$), uma vez que os falantes identificaram os dois padrões de modo semelhante. É interessante notar, contudo, que do mesmo modo que para as bilabiais, a diferença entre os padrões DV (97%) e V (75%) é significativa ($Z = -2.754^b$, $p = 0,006$), mas para as velares esse fato está atrelado aos falantes terem identificado melhor os dados quando possuíam a pista duração de vogal, em detrimento do vozeamento.

Contudo, as alveolares apresentaram um resultado distinto do teste de discriminação e de identificação para todos os GE. Nenhum dos padrões apresentou diferença significativa entre si. Entre DC (97%) e DV (88%) os valores foram ($Z = -1.667^b$, $p = 0,096$) e entre DC (97%) e DC (88%) foi de ($Z = -1.518^b$, $p = 0,129$). Por fim, ambos os padrões DV (88%) e V (88%) obtiveram percentuais iguais e, assim, não são distintos entre si.

Encaminha-se agora para a descrição dos resultados do teste de identificação para o GC.

	Bilabial			Alveolar			Velar		
	DC	DV	V	DC	DV	V	DC	DV	V
N	24	24	24	24	24	24	24	24	24
\bar{X}	24	24	22	24	24	12	24	24	8
% \bar{X}	100	100	92	100	100	50	100	100	33
% DP	0	0	24	0	0	18	0	0	27

Tabela 13 Resultados da média e percentual de acertos para o teste de identificação para o GC. Notação: N = número total de dados; \bar{X} = média de acertos para cada padrão; % \bar{X} = média de acertos em porcentagem para cada padrão; % DP = porcentagem do desvio padrão; DC: Dado Completo; DV: Duração Vogal; Vozeamento.

	Bilabial			Alveolar			Velar		
	DC x DV	DC x V	DV x V	DC x DV	DC x V	DV x V	DC x DV	DC x V	DV x V
Z	.000 ^b	-1.000 ^c	-1.000 ^c	.000 ^b	-2.251 ^c	-2.251 ^c	.000 ^b	-2.414 ^c	-2.414 ^c
P	1.000	.317	.317	1.000	.024	.024	1.000	.016	.016

Tabela 14 Resultados do teste de Wilcoxon para o teste de discriminação para o GC. Z: Wilcoxon; p: p valor.

A partir dos resultados expostos nas Tabelas 13 e 14, percebe-se que a diferença entre o percentual de acerto para as alveolares entre DC (100%) e V (50%) é significativa ($Z = -2.251^c$, $p = 0,024$) e que a distinção entre os valores de DV (100%) e V (50%) também é ($Z = -2.251^c$, $p = 0,024$). Um cenário semelhante pode ser articulado para as velares, sendo que a diferença entre o percentual de acerto também foi maior para o padrão DC e DV. A diferença entre DC (100%) e V (33%) é significativa ($Z = -2.414^c$, $p = 0,016$). Do mesmo modo, o teste de Wilcoxon destaca a distinção entre os valores de DV (100%) e V (33%) como significativa ($Z = -2.414^c$, $p = 0,016$). Isso significa que o padrão V dificultou a discriminação do contraste surdo-sonoro em posição final tanto para alveolares como para velares, assim como no teste de discriminação. Assim como no teste de discriminação, nenhuma diferença no percentual de acerto entre os padrões para as bilabiais foi verificada.

7.2.3.5 Resumo dos resultados do teste de identificação: relação entre GE e GC e pontos de articulação.

Vários dos resultados encontrados para o teste de discriminação foram verificados também para o teste de identificação. Entre eles se encontra uma maior média de acertos para o padrão DC em relação ao DV, mas nem sempre em relação à V. É importante ressaltar que a nomenclatura “significativo” e “não significativo” no Quadro 5 está relacionada com o fato do padrão ter ou não sido melhor reconhecido pelos informantes. Assim, serão analisados mais detidamente os padrões que passaram por manipulação: DV (duração de vogal preservada) e V (vozeamento).

Ponto de articulação	Padrão DC x DV	Padrão DC x V	DV X V
Bilabiais	Não Significativo: GC	Não significativo: Todos os grupos.	Não Significativo: GC
Alveolares	Significativo: GE Básico e Pré-Intermediário	Significativo: Todos os níveis	Significativo: GC
Velares	Significativo: GE Básico	Significativo: Todos os níveis.	Significativo: Todos os níveis.

Quadro 5 Resumo dos resultados para o teste de identificação.

Do mesmo modo que no teste de discriminação, observou-se uma relação entre a média de acerto do padrão manipulado e o ponto de articulação. O padrão V obteve uma média de acerto maior em relação ao padrão DV, e apresentou uma média semelhante ao padrão DC para as bilabiais para todos os GE, i.e., os informantes, com exceção dos pertencentes ao GC, identificaram melhor os dados quando possuíam a pista vozeamento. Para as alveolares, no entanto, de modo geral, a média de acerto entre o padrão DC ser maior do que DV foi semelhante para alguns grupos, GE Básico

e GE Pré-Intermediário e não para GE Intermediário e GC. Para as velares, o padrão V foi considerado o menos saliente na identificação das oclusivas sonoras alveolares e velares para todos os níveis, i.e. , é possível que o ponto de articulação tenha dificultado a percepção da distinção surdo-sonora. Se tomarmos como base a fonologia gestual (Browman & Goldstein, 1986 e os seguintes; Albano, 2001), pode-se tentar explicar a proximidade e distância nos resultados de alguns parâmetros acústicos, como por exemplo, bilabiais e velares⁶⁰. Uma especulação possível seria prever que as oclusivas localizadas nos extremos do trato selecionam padrões distintos para serem mais facilmente percebidas. Por outro lado, os valores para os padrões DV e V para as alveolares, se mostram semelhantes para quase todos os GE, o que poderia indicar que quando a oclusiva se encontra na posição média do trato, ambas as pistas parecem ser importantes na distinção do contraste surdo-sonoro. Assim como no teste de discriminação, é importante ressaltar a importância da porcentagem do desvio padrão para o padrão V. Da mesma maneira que no teste de discriminação, é visível que os informantes demonstraram uma maior dificuldade de identificar os dados que continham esse padrão. No entanto, ao contrário do teste de discriminação, é possível observar que uma alta porcentagem de desvio padrão não ocorre na mesma quantidade de dados que no teste de discriminação. Poderia se inferir que os informantes tiveram menos dificuldade de identificar o padrão V no teste de identificação do que no teste de discriminação.

Em geral, a análise da média de acerto dos dados do teste de identificação não parece corroborar os achados de Warner *et.al.* (2004) e Kleber *et.al.* (2010), que a duração da vogal se mostra como uma melhor pista na identificação do contraste surdo-sonoro em oclusivas em posição final.

⁶⁰ Este trabalho não tem como objetivo sugerir uma representação do contraste surdos-sonoro a partir da Fonologia Gestual e, portanto, não nos deteremos na elucidação de qual seria a configuração gestual que aproximaria as oclusivas de diferentes pontos de articulação. Contudo, existe uma preocupação a ser investigada em pesquisas futuras sobre os efeitos articulatórios observados para algumas oclusivas (e.g. bilabiais e velares).

Por fim, é interessante notar que alguns resultados se assemelham ao teste de discriminação e que os resultados dos informantes do GC não se distanciam dos obtidos pelos GE, i.e., os informantes do GE não estão percebendo os dados de maneira aleatória e estão muito próximos das dificuldades apresentadas pelos falantes nativos do GC.

Uma melhor visualização das considerações feitas para os resultados dos GE e GC pode ser feita a partir das Figuras 35, 36 e 37.

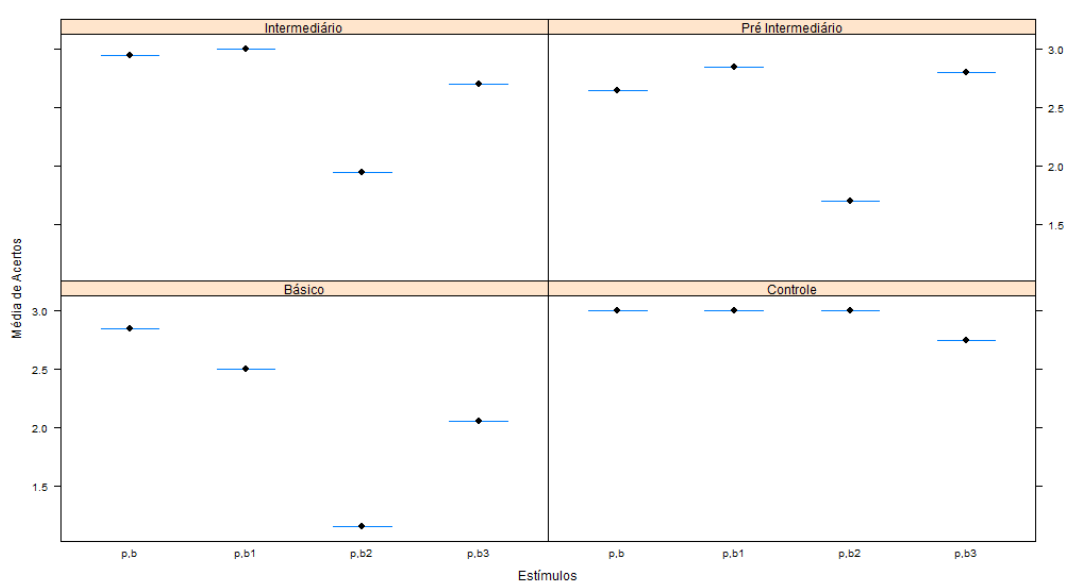


Figura 35 Gráfico com a média de acertos das velares do teste de identificação.
 Notação: p,b1= Dado Completo ;p,b2= padrão Duração de Vogal; p,b3= Vozeamento.

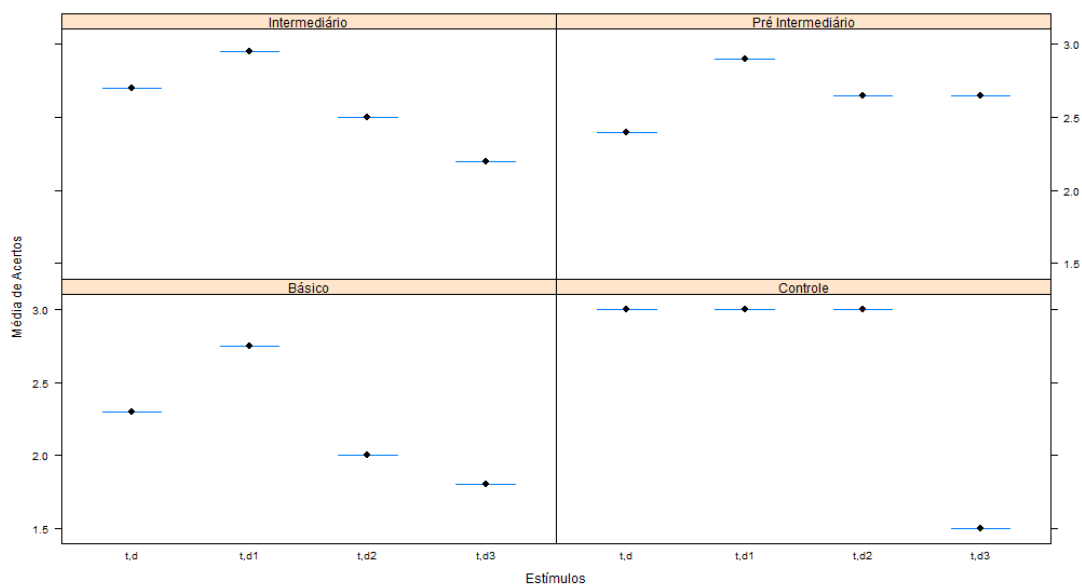


Figura 36 Gráfico com a média de acertos das velares do teste de identificação. Notação: t,d1= Dado Completo ;t,d2= padrão Duração de Vogal; t,d3= Vozeamento.

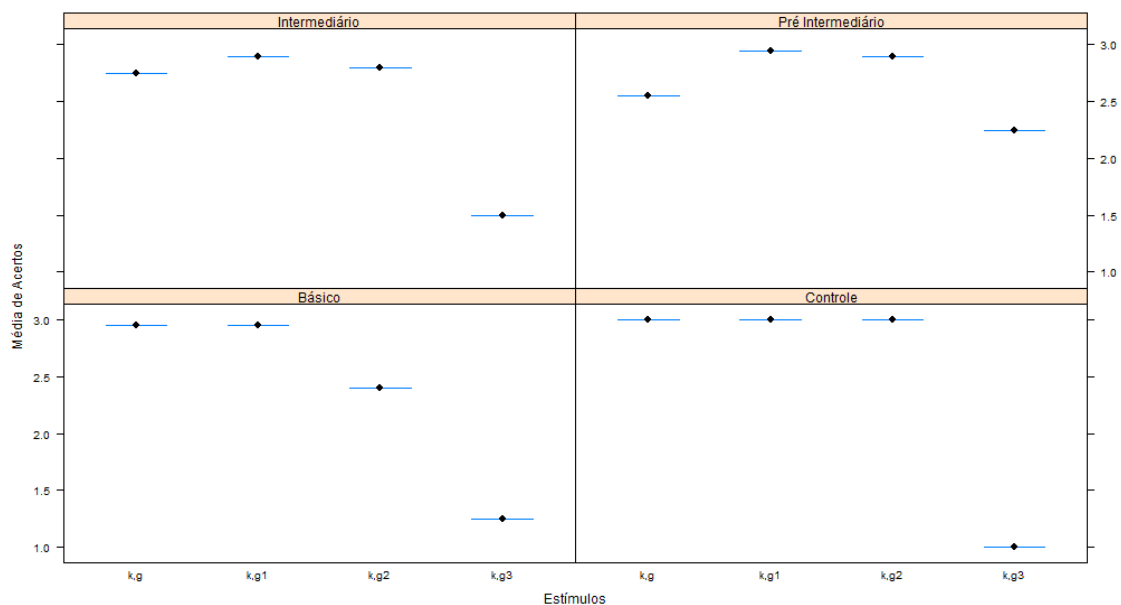


Figura 37 Gráfico com a média de acertos das velares do teste de identificação. Notação: k,g1= Dado Completo ;k,g2= padrão Duração de Vogal; k,g3= Vozeamento

7.2.3.6 Análise do Tempo de Reação

Conforme havia sido mencionado nos objetivos deste experimento, foi rodado o teste de correlação de Pearson, para verificar se existia alguma

correlação significativa entre as covariáveis, no caso deste estudo o Tempo de Reação e cada uma das covariáveis: Nível de proficiência, Estímulo, Repetição e Resultado. Em ambos os testes, todas as correlações se mostraram fracas com coeficiente de correlação ρ em torno de 0.10. Entretanto, entre o Tempo de Reação e a Repetição o valor de ρ foi maior, em torno de 0.22. Apesar de ainda se mostrar uma correlação fraca, pode-se dizer que o Tempo de Reação tende a diminuir à medida que o indivíduo ouve pela segunda ou terceira vez os estímulos. As Figuras 38 e 39 mostram o ρ valor para os dois testes.

\$correlation

	Tempo	Nivel	Estimulo	Repeticao	Resultado
Tempo	1.00	0.10	0.11	-0.22	-0.13
Nivel	0.10	1.00	0.00	0.00	0.12
Estimulo	0.11	0.00	1.00	0.00	0.01
Repeticao	-0.22	0.00	0.00	1.00	0.04
Resultado	-0.13	0.12	0.01	0.04	1.00

\$pvalue

	Tempo	Nivel	Estimulo	Repeticao	Resultado
Tempo	1.000000e+00	3.856306e-07	2.221008e-08	0.00000000	1.255103e-10
Nivel	3.856306e-07	1.000000e+00	1.000000e+00	1.00000000	8.015506e-10
Estimulo	2.221008e-08	1.000000e+00	1.000000e+00	0.86342982	6.430098e-01
Repeticao	0.000000e+00	1.000000e+00	8.634298e-01	1.00000000	4.392278e-02
Resultado	1.255103e-10	8.015506e-10	6.430098e-01	0.04392278	1.000000e+00

Figura 38 Valores do teste de Spearman para o teste de discriminação.

\$correlation

	Tempo	Resultado	Nivel	Estimulo	Repeticao
Tempo	1.00	0.04	0.00	0.16	-0.15
Resultado	0.04	1.00	0.07	-0.05	0.01
Nivel	0.00	0.07	1.00	0.00	0.00
Estimulo	0.16	-0.05	0.00	1.00	-0.01
Repeticao	-0.15	0.01	0.00	-0.01	1.00

\$pvalue

	Tempo	Resultado	Nivel	Estimulo	Repeticao
Tempo	1.00000000	1.156389e-02	7.934144e-01	0.00000000	0.00000000
Resultado	0.01156389	1.000000e+00	3.775921e-05	0.002829974	0.6610215
Nivel	0.79341438	3.775921e-05	1.000000e+00	0.986579090	1.0000000
Estimulo	0.00000000	2.829974e-03	9.865791e-01	1.000000000	0.5203315
Repeticao	0.00000000	6.610215e-01	1.000000e+00	0.520331461	1.0000000

Figura 39 Valores do teste de Spearman para o teste de identificação.

7.2.3.7 Análise da Semelhança entre os Níveis de Proficiência

Ao conduzir o testes experimento de percepção, notou-se que tanto para o teste de discriminação, como para o de identificação, os acertos gerais para os GE e GC mostraram que o GE Básico se distancia do GE Pré-Intermediário e GE Intermediário e GC; o GE Pré-Intermediário e GE Intermediário possuem médias de acertos semelhantes, por vezes próximas às do GC. Na Figura 40 é possível notar que há uma concentração maior dos dados para os sujeitos do GC, diferentemente do que acontece para os GE, cujos resultados se espalham mais. Isso demonstra que as médias de acertos do grupo controle são mais altas e consistentes do que as dos grupos-alvo, assim como era esperado.

As Figuras 40 e 41 mostram uma visão geral do resultado da análise com o intuito de avaliar a distribuição dos dados. Assim, utilizou-se o gráfico

box plot, no qual se tem valor máximo e mínimo, o primeiro e terceiro quartil⁶¹ e a mediana. Nota-se na Figura 40 um resumo gráfico da média geral de acertos por nível de proficiência, sem que se discriminem os estímulos. É possível perceber que o GC se destaca dos outros GE, como era esperado. O distanciamento do grupo de nativos ocorre pelo valor da sua mediana superior (terceiro quartil) e pela distribuição dos dados. Ao mesmo tempo, é possível observar que os GE Pré-Intermediário e Intermediário possuem valores de mediana muito próximos. Contudo, apesar de a distribuição parecer ordenada, ainda existem alguns casos discrepantes, como é o exemplo de dois informantes do GC que exibiram valores diferentes do resto do grupo (no gráfico da Figura 40, esses sujeitos são duas “bolinhas” que aparecem com acertos abaixo de 1, ou seja, tal informante errou alguns dos dados, e o outro obteve acertos em torno de 1,5, o que significa que o outro informante obteve, para alguns dados, entre 1 e 2 acertos). Como a amostra do GC foi pequena (8 informantes) e dentro dessa 2 sujeitos apresentaram dados divergentes, optou-se, como pôde ser visto nas análises para o teste de discriminação e para o teste de identificação, por analisar os dados desse grupo separadamente.

Salienta-se que a Figura 40 representa o número de acertos plotados em uma escala de 1 a 3, que são os valores que cada indivíduo poderia acertar (acertando 1, 2 ou 3 vezes) pelo Nível de Proficiência.

⁶¹ O quartil é um tipo de medida em dados que possuem uma distribuição ordenada. Ele representa um dos três valores quando os dados são divididos em 4 partes. O primeiro quartil equivale a 25%, o segundo designa até o percentual de 50% (que seria a mediana), o terceiro ao valor de 75%. Para o experimento aqui descrito, é importante que se saiba o quanto os acertos se aproximam da mediana. Como houve alguns casos de informantes que se mostram isolados nas Figuras 21, e 22, realizar a quantificação dos acertos conforme a divisão mencionada acima confere um tratamento mais adequado às produções discrepantes e à concentração dos valores em cada grupo.

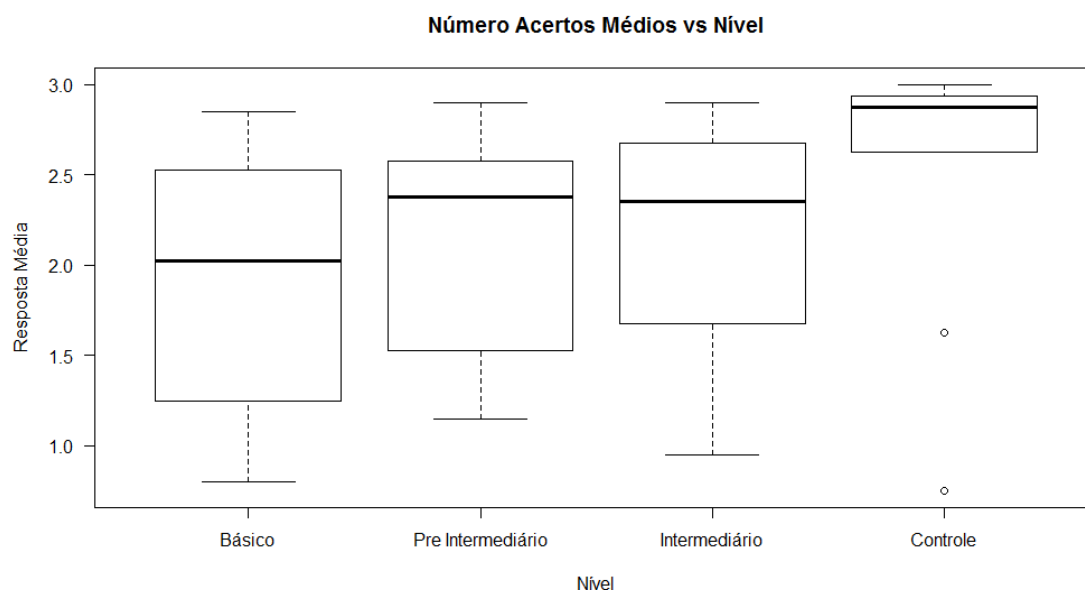


Figura 40 Box plot da média de acertos por nível para o teste de discriminação.

As mesmas observações feitas com relação ao teste de discriminação podem ser assumidas para o teste de identificação. Ambos os GE, Pré-Intermediário e Intermediário, apresentam valores próximos. Contudo, tais valores são maiores (entre 2,5 e 3 acertos) do que os do teste de discriminação (entre 2 e 2,5 acertos). A média de acertos para o GC ainda continua mais alta quando comparada à dos outros grupos, mas o valor também aumentou.

Em ambos os testes, pode-se observar que dois informantes do GC apresentam alguns resultados divergentes e se encontram deslocados do resto do grupo. De modo distinto do teste de discriminação, o GE Pré-Intermediário apresentou um informante com resultados muito diferentes do grupo, obtendo uma média de acertos entre 1,5 e 2. A Figura 41 representa o número de acertos plotados em uma escala de 1 a 3, que são os valores que cada indivíduo poderia acertar (acertando 1, 2 ou 3 vezes) pelo Nível de Proficiência.

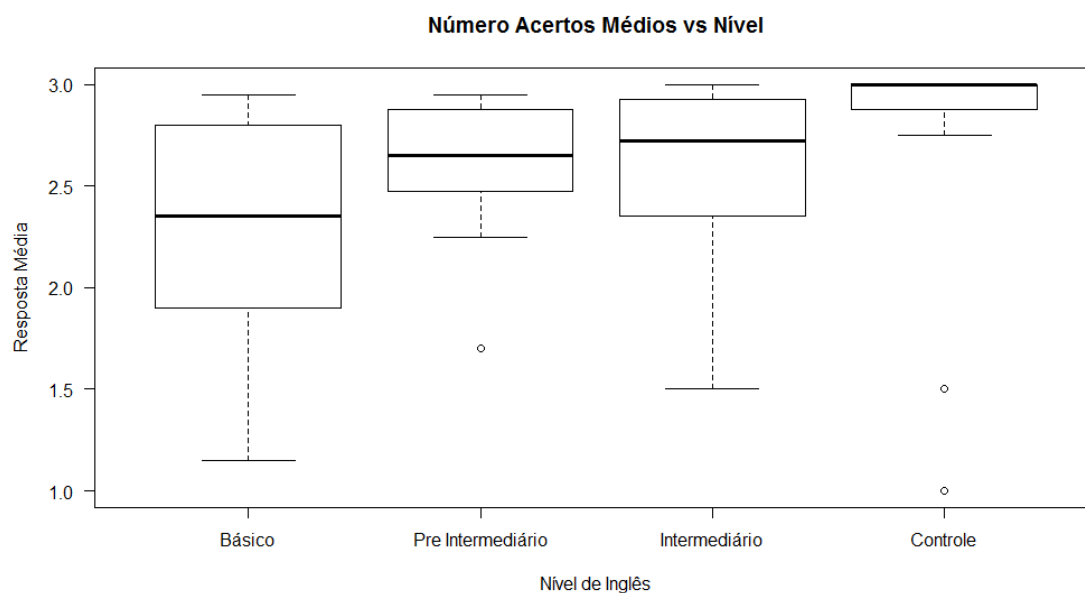


Figura 41 Box plot da média de acertos por nível para o teste de identificação.

Embora as Figuras 40 e 41 não revelem informações mais pontuais sobre a relação entre as pistas e parâmetros acústicos utilizados (já analisados em seções prévias) na distinção do contraste surdo-sonoro, elas são importantes para uma compreensão geral das relações entre os níveis. Conforme pôde ser visualizado, essas figuras permitem a aproximação dos GE Pré-Intermediário e Intermediário. Apesar desses GE possuírem quantidades distintas de horas de ensino formal, os resultados apontam uma semelhança na capacidade perceptiva do contraste sonoro em posição final em comparação com o GC.

Com relação aos dados não manipulados, é possível dizer que ambos os pares, oclusiva surda e sonora (e.g., p,b1) e surda e surda (e.g., p,p), foram discriminados com médias mais altas e desvios padrão mais baixos do que os dados manipulados, conforme era esperado. Tal fato se verificou para todas as oclusivas analisadas: bilabiais, alveolares e velares.

A partir dos resultados obtidos dos testes de discriminação e identificação, mais pistas puderam ser lançadas para o tratamento da DT. Conforme foi pontuado ao longo da descrição da descrição dos resultados dos

testes, notou-se para os GE que, embora se tenha utilizado o padrão DV, para alveolares e velares, como uma pista melhor na distinção surdo-sonora em posição final, a média de acertos variou entre os níveis de proficiência. Em nenhum dos GE, os padrões DC, DV e V foram discriminados ou identificados categoricamente, ou seja, de modo que não se percebesse, por exemplo, a existência de manipulações nas produções que resultassem em dados perceptualmente distintos. Embora não se possa afirmar, ainda, que a percepção para o fenômeno da DT seja estritamente categórica ou gradiente, os modelos dinâmicos, como é o caso a Fonologia Gestual, talvez possam dar conta de explicar a interação entre as pistas (duração de vogal e duração de vozeamento). Ou seja, que tipo de configuração gestual é necessária para a percepção de uma oclusiva alveolar, por um lado, e de bilabiais e velares, por outro?

Verificou-se que, uma vez que uma das pistas manipuladas não estava presente, os sujeitos precisavam lançar mão de outros parâmetros para que percebessem os dados acuradamente. Não se observou, portanto, que na falta de um dos parâmetros acústicos, os GE e o GC tenham errado boa parte dos dados, neutralizando, assim, o contraste surdo-sonoro.

Pôde-se notar, também, que parece existir uma gradiência em relação ao aprendizado das pistas, ao menos do GE Básico em relação aos outros GE e com relação ao GC. Embora o teste de *Spearman* tenha registrado uma significância pequena entre a média de acertos e os níveis de proficiência, é possível notar que a percepção dos parâmetros acústicos pelos GE é semelhante à do GC. Ou seja, os aprendizes possuem dificuldades de reconhecer o padrão V para as alveolares e velares, assim como os falantes nativos. Em última instância, isso não só revela que os informantes-aprendizes estão selecionando padrões semelhantes aos dos nativos para realizar a distinção do contraste sonoro em posição final, mas também que a discriminação e identificação de oclusivas surdas e sonoras se dá num *continuum*. Tal afirmação procura mostrar que a discriminação e identificação das oclusivas em posição final não ocorre de forma categórica, não é possível

assumir que os aprendizes não sejam capazes de discriminar ou identificar, a *priori*, segmentos que não fazem parte de seu sistema fonológico. Os sujeitos vão adquirindo, aos poucos, a percepção das pistas acústicas, as quais são responsáveis para a diferenciação da sonoridade em posição final.

8 CONSIDERAÇÕES FINAIS

8.1 DISCUSSÃO DOS RESULTADOS E IMPLICAÇÕES TEÓRICAS

Conforme pôde ser observado ao longo deste estudo, a DT, fenômeno analisado por este trabalho, é tratada por uma vasta literatura, que conta tanto com estudos de produção quanto de percepção. Dessa forma, esta pesquisa se propôs a realizar um experimento que pudesse lançar mais pistas no tratamento da DT a partir de dados de aprendizes brasileiros. Embora alguns estudos já tivessem realizado experimentos de percepção com brasileiros, (Zimmer e Alves, 2008, Bettoni-Techio, 2005), este trabalho buscou observar outros aspectos envolvidos na percepção das oclusivas finais do inglês, a saber: influência entre nível de proficiência e percentual de acertos, ponto de articulação e parâmetros acústicos (duração de vogal e vozeamento), a relação entre tempo de reação.

Foi possível realizar algumas generalizações a partir dos resultados obtidos pelo experimento. O primeiro ponto a ser discutido é com relação à questão experimental. A partir do exposto pela literatura Guion *et.al.*(2000), Harnsberger (2001), Wayland (2007), dentre outros, o modo como os dados são apresentados para os informantes, seja por um par de estímulos auditivos (experimentos descritivos) ou a partir de um único estímulo auditivo (experimentos de identificação), influencia no percentual de acertos dos sujeitos em testes perceptuais. Para este estudo em especial, embora se observe que as generalizações obtidas foram semelhantes para um e outro tipo de teste, a porcentagem de acertos foi maior para o teste de identificação do que para o teste de discriminação. Assim, outros estudos que contem com um maior número de dados talvez possam iluminar a discussão, verificando se há ou não a interferência do tipo de teste na percepção.

Com relação aos parâmetros acústicos, verificou-se que parece existir uma influência do ponto articulatorio e os padrões escolhidos para discriminar e identificar as consoantes, assim como o experimento de Albuquerque (2010) relatou para dados da produção de oclusivas finais do inglês por informantes

brasileiros. Dentre os resultados do experimento, notou-se que os informantes discriminaram e identificaram com mais acurácia o contraste surdo-sonoro para as bilabiais a partir do padrão V, que corresponde à presença de vozeamento, do que a partir do padrão DV, quando possuíam a duração da vogal antecedente preservada. É importante mencionar, no entanto, que essa afirmação vale para o GE Intermediário no teste de discriminação e para todos os GE no teste de identificação. Em última instância, tal achado fornece uma pista interessante no que diz respeito aos níveis de proficiência e a aquisição do contraste surdo-sonoro. Embora em muitos resultados os GE Pré-Intermediário e Intermediário se comportem de modo semelhante, em alguns momentos o GE Intermediário parece estar lançando mão de pistas diferentes para estabelecer o contraste surdo-sonoro em posição final. Por outro lado, de modo geral, é preciso que se investigue mais fundo a relação entre os GE Pré-Intermediário e Intermediário, pois conforme foi relatado na seção sobre a constituição do grupo de informantes, mencionou-se que a ferramenta de agrupamento utilizada havia sido a quantidade de horas de ensino formal dos alunos. Tal ferramenta pode ter enviesado a formação dos grupos, pois se os alunos tivessem sido classificados a partir de um teste de proficiência, talvez alguns alunos do GE Intermediário tivessem ficado no GE Pré-Intermediário e vice-versa. No entanto, é importante que se entenda que o agrupamento dos alunos foi feito de maneira a respeitar uma determinada metodologia com relação à prática de pronúncia em cada nível do centro de língua pesquisado.

Para as oclusivas alveolares e velares o resultado foi inverso: o padrão DV se apresentou como melhor na distinção surdo-sonora de fim de palavra. Tanto para os GE quanto para o GC, o maior índice de identificações e discriminações não adequadas recaiu sobre o padrão V para as velares. Novamente, é possível observar a interação destas duas variáveis: ponto de articulação e parâmetro acústico. Tal constatação é passível de ser explicada a partir de modelos gestuais como o de Browman e Goldstein (1986, 1989, 1990, 1992), a partir da semelhança entre os gestos responsáveis pela constituição de cada uma das oclusivas: bilabiais, alveolares e velares. Ao receberem os

estímulos relativos às bilabiais, os informantes possuem uma pista gestual diferenciada, o fechamento dos lábios ao final do segmento, gesto que não é incorporado para as alveolares e velares, não com os mesmos movimentos articulatórios ao menos.

Ao contrário do que se esperava, no entanto, não foi encontrada nenhuma relação distintiva entre os níveis de proficiência e o percentual de acertos. Em geral, o nível de proficiência dos aprendizes só obteve um percentual significativo no GE Básico em relação aos demais GE e o GC para as oclusivas bilabiais. Assim, para este conjunto de dados, não se observou a influência generalizada do nível no percentual de acertos. Por outro lado, alguns apresentaram um alto índice de desvio padrão. Isso pode significar que apesar da duração de vogal ter se mostrado como uma pista mais robusta para a discriminação e identificação de oclusivas surdas e sonoras em posição final na maioria dos casos, não é possível dizer que isso se dá para todos os casos, pois se observa uma flutuação nas escolhas dos informantes. Ou seja, em alguns dados os informantes apresentaram porcentagens de acerto e erro muito próximas, o que pode indicar que eles ainda estão em processo de aquisição das pistas que levam à discriminação e identificação do contraste surdo-sonoro em posição final.

Além disso, assim como havia sido observado por Albuquerque (2010) a partir de modelos dinâmicos, como é o caso da fonologia gestual de Browman e Goldstein (*op.cit.*), os resultados do desempenho dos informantes do GC não se distanciaram dos dados para os GE, com exceção da observação para os GE com relação às bilabiais. Não foi verificada nenhuma influência do parâmetro acústico nas bilabiais para o GC. Contudo, da mesma forma que para os GE, houve um percentual de acerto significativamente menor para as velares.

Apesar do tempo de reação não ter sido significativo para nenhum dos experimentos, pode-se dizer que ele obteve um valor de significância baixo, mas existente, para a variável repetição, ou seja, conforme era esperado, os sujeitos demonstraram uma relação de aprendizagem com a apresentação dos

dados, de forma que o percentual de acertos foi diretamente proporcional à ordem das repetições (a primeira repetição obteve um percentual de acerto menor em relação às outras duas repetições e a terceira foi a que apresentou um percentual maior). Isso se encontra de acordo com a previsão do PAM-L2 (1995) de que os informantes conseguem aprender a identificar pistas para discriminar dados da L2 que não são utilizadas da mesma maneira na sua L1.

Em relação aos resultados deste experimento e os modelos teóricos que se debruçaram sobre a DT, algumas considerações podem ser feitas. Conforme descrito anteriormente, a DT é um fenômeno descrito como fonológico, ou seja, tal classificação parte de modelos mais tradicionais que realizam a separação entre o nível fonético e fonológico, como foi o caso da análise feita por Oostendorp (2007) a partir dos traços articulatórios, dentro do modelo da OT⁶². É interessante observar, no entanto, que para conseguir dar conta do fenômeno da DT, o autor precisou lançar mão de outra teoria auxiliar, a *Turbidity Theory* (Teoria da Turbidez, conforme foi traduzido por este estudo). Tal teoria se propõe a incorporar na análise dados que apresentem variação, como é o caso da DT, mas ainda tendo como base ferramentas categóricas como os traços binários de Chomsky e Halle (1968). Tal informação se mostra contrastante com o fato relatado por autores como Port (1996), que descrevem a DT como um fenômeno gradiente e, como tal, prevê a adoção de modelos e/ou abordagens que ofereçam um tratamento a partir de primitivos que aceitem tal variação. Apesar de Oostendorp (*op.cit.*) não negar a existência da gradiência nos dados da DT, o tratamento dado pelo autor para os dados de neutralização incompleta preservam a existência de categorias [t], [d] e [d_o] (sonora desvozeada) e corroborar a análise de Port e Leary (2005) de que [t] e [d_o] não são iguais e nem diferentes entre si através de um tratamento formal da DT, ou seja, procura-se acomodar a variação presente no dado fonético criando-se mais uma categoria fonológica.

⁶² Embora a OT não tenha, necessariamente, como primitivo de base os traços binários de Chomsky e Halle (1968), o trabalho de Oostendorp (*op.cit.*) os utiliza para analisar a DT.

A discussão que está por trás das considerações feitas por Port e Leary (*op.cit.*) e Oostendorp (*op.cit.*) se concentra na distinção entre abordagens funcionalistas e formalistas, entre incorporar o dado fonético ou não, e como contemplar o dado fonético dentro de modelos formais. A saída de Oostendorp (*op.cit.*) parece se assemelhar à tomada por Keating (1985) sobre a duração extrínseca das vogais ser específico de língua, o que implica, portanto, que a duração dos segmentos vocálicos faria parte do inventário dos sujeitos, e não algo implementacional, fonético. Ao longo do trabalho, Keating (*op.cit.*) tece argumentos que lidam com a limitação do modelo com o qual a autora trabalha (o gerativismo), e o fenômeno que está analisando. Ao fim do estudo, a autora relata que a natureza dos dados aponta para a necessidade de se incorporar a variável tempo, presente em modelos dinâmicos. Contudo, isso não pode ser feito por um modelo que agrupa os dados em categorias fechadas, como é o caso do gerativismo. Ambos, Keating (*op.cit.*) e Oostendorp (*op.cit.*), embora trabalhem com fenômenos distintos e com teorias distintas, procuraram adequar os seus dados gradientes a partir de primitivos categóricos.

Além disso, é interessante notar que, ao contrário do que prevê o PAM L2, que de modo geral seria mais fácil para os aprendizes discriminarem do que identificarem, tal fato não pôde ser corroborado por esse trabalho. A princípio, os percentuais de acerto para o teste de identificação foram semelhantes ou maiores, para alguns dados, aos do teste de discriminação. Isso corrobora, mais uma vez, que os informantes parecem passar por um processo de amadurecimento gestual. Ao mesmo tempo, é possível analisar os resultados do estudo a partir das categorias estabelecidas por (Best e Tyler, 2007). Os resultados para os dois padrões manipulados, DV e V, não se encontram distribuídos de modo homogêneo. Assim, talvez a percepção do contraste surdo-sonoro a partir do padrão DV possa ser encaixada dentro do que os autores indicam como assimilação do tipo 1 (Assimilação do Tipo Duas-Categorias), pois a discriminação e identificação desta manipulação alcançou, na maioria dos dados, um percentual de acerto muito próximo ao dos dados que não sofreram manipulação, *i.e.*, a percepção foi de boa a ótima. Contudo, o

padrão V não parece ter uma distribuição tão regular, pois ao mesmo tempo em que se mostrou o padrão mais saliente para as bilabiais no teste de identificação, todos os grupos, inclusive o GC tiveram problemas para identificar o padrão para as velares. Portanto, a assimilação deste padrão parece se encaixar no tipo de assimilação 3 (Assimilação do Tipo Única-Categoria), na qual ambos os membros do contraste são percebidos como igualmente desviantes ou igualmente aceitáveis exemplares da categoria da L1. Não nos parece que a discriminação e identificação do padrão V recaia sobre a categoria 4 ou 5, de sons incategorizáveis, pois apesar do percentual de acerto ser baixo para as velares, os informantes, de modo geral, procuraram estabelecer o contraste surdo-sonoro em posição final.

Tanto os dados de produção de oclusivas finais do inglês por falantes brasileiros Zimmer (2007), Zimmer e Alves (2007), Zimmer e Alves (2008), Bettoni-Techio (2005), Albuquerque (2010) quanto os de outras línguas, como os de Piroth e Janker (2004), Ernestus e Baayen (2006), entre outros, apontam para a gradiência do fenômeno da DT na produção e, uma vez que se observa que não se trata de um fenômeno categórico, outros tipos de primitivos de base gradiente precisam ser adotados. A partir de experimentos como o que foi relatado por este trabalho, tomando como base estudos de percepção sobre a DT (Warner *et al.*, 2004, Kleber *et al.*, 2010, etc), nota-se que a DT parece se mostrar um fenômeno gradiente também para a percepção. Embora sejam necessários outros estudos, que levem em conta as variáveis aqui pesquisadas (nível de proficiência dos informantes, manipulação de parâmetros acústicos, diferentes pontos articulatórios das oclusivas em posição final, tempo de reação) aponta-se que modelos que assumem como primitivos de análise unidades híbridas (como os gestos) e não categorias (como os traços) parecem dar conta de fenômenos como a DT de modo mais parcimonioso. Assim, embora não se defenda, neste trabalho, a utilização de um modelo ou outro para dar conta da DT, tem-se que modelos que possuem um entendimento da natureza de um fenômeno como sendo gradiente se mostram mais adequados para tratar desse, como é o caso do conexionismo e a fonologia gestual. Não

se exclui, portanto, modelos como a OT, uma vez que esse pode trabalhar com primitivos que não sejam os traços distintivos.

A partir da adoção de modelos dinâmicos, a noção de “erros” de produção e percepção, – levantada ao longo do trabalho – dá lugar, dentro de modelos que levem em conta as diversas variáveis que fazem parte do processo de aquisição de L2, à noção de “ajuste” que os aprendizes de L2 precisam fazer para adequar os sons recebidos dentro do sistema da L1/L2.

8.2 LIMITAÇÕES DO ESTUDO E SUGESTÕES DE PESQUISAS FUTURAS.

Apesar de o experimento ter se mostrado extremamente distinto do piloto em termos de controle dos dados, número e proficiência dos informantes e outros aspectos, teria sido interessante possuir diferentes grupos entre os níveis de proficiência: mais de um grupo de cada nível (básico, pré-intermediário e intermediário). Assim, teria sido possível comparar se grupos distintos possuiriam diferentes percentuais de acerto.

Ainda que não se trate de uma limitação deste trabalho em especial, pode-se dizer que não foi possível realizar uma análise de correlação entre os GE e o GC, pois no segundo não havia o mesmo número de sujeitos que nos outros. A disponibilidade de sujeitos nativos com um perfil semelhante (Americanos que pertenciam a um mesmo estado nos Estados Unidos – Carolina do Norte) foi um fator decisivo para a seleção desse grupo e da quantidade de informantes. No entanto, um maior número de informantes nativos poderia fornecer uma análise comparativa com cada grupo alvo.

Da mesma maneira que foi feito para os brasileiros, analisar como eles percebem pistas existentes em sua língua a partir dos dados de um falante nativo de inglês, uma sugestão que se apresenta é rodar o experimento com americanos percebendo como os brasileiros produzem as oclusivas em posição final. Tal experimento investigaria se os americanos utilizam as mesmas pistas que os brasileiros para realizar a discriminação e identificação dos dados e o papel de pistas secundárias na percepção destes sons.

Assim como Kluge *et al.* (2009) e Kluge (2009) para oclusivas nasais em posição final, um estudo posterior poderia contar com a utilização de pistas visuais (o sujeito veria como estímulo uma imagem de alguém produzindo a palavra-alvo).

Embora este estudo não tenha tido como objetivo propor soluções pedagógicas para que se lide com dados de pronúncia de oclusivas em posição final do inglês, a motivação para o estudo partiu das observações da produção dos alunos com os quais a pesquisadora teve contato. É mister dizer que, apesar de não se poder tecer nenhuma observação mais concreta sobre como inserir tais achados no interior das aulas de língua, os resultados do estudo podem conferir ferramentas aos docentes para um melhor entendimento da produção dos seus alunos como algo dinâmico, e que se aproxima mais da produção de um falante nativo do que se distancia. Além disso, o docente pode começar a notar que os alunos percebem diferenças em produções de sons semelhantes e que a relação entre produção e percepção de sons não é categórica como indicam alguns materiais didáticos. A partir disso, é possível que a noção do “erro” ou desvio de produção e percepção seja rediscutida, de modo que os docentes possam entender que não se trata de perceber ou não um determinado som, mas sim de guiar os discentes sobre as especificidades dos gestos necessários para a produção e compreensão de sons que não possuem a mesma configuração dos existentes na L1 de seus alunos.

9 REFERÊNCIAS BIBLIOGRÁFICAS

ALBANO, Eleonora. *O gesto e suas bordas: esboço de Fonologia Acústico-Articulatória do Português Brasileiro*. Campinas: Mercado de Letras/ALB/FAPESP, 2001.

ALBUQUERQUE, J. I. A de. "Dessonorização Terminal (?): Discussão sobre os resultados de um experimento e direcionamentos para o papel da percepção na aquisição de L2", Anais do IX Encontro do CELSUL, Palhoça, SC, 2010.

AL-SHORMANI, M. Fossilization and Plateau Effect in Second Language Acquisition. CALTS: University of Hyderabad, 2010. Disponível em: <<http://pt.scribd.com/doc/25280138/Fossilisation-and-Plateau-Effect-in-Second-Language-Acquisition>>. Acesso em: 11.12..2011.

ALVES, U. *O papel da instrução explícita na aquisição fonológica da L2: evidências fornecidas pela Teoria da Otimidade*. 335 f. Dissertação de Mestrado. Universidade Católica de Pelotas. Pelotas, 2004.

_____. *A aquisição das seqüências finais de obstruintes do inglês (L2) por falantes do Sul do Brasil: análise via Teoria da Otimidade*. Tese – Doutorado em Letras. PUCRS, Porto Alegre, 2008.

AOYAMA, K.; FLEGE, J. E.; GUION, S.; AKAHANE-YAMADA, R.; YAMADA, T. "Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/", *Journal of Phonetics* 32: 233–250, 2004.

ASLIN, R. N., PISONI, D. B., & JUSCZYK, P.W. Auditory development and speech perception in infancy. In P. H. Mussen (Series Ed.), M. M. Haith, & J. J. Campos (Vol. Eds.), *Handbook of child psychology: Vol. 2. Infancy and developmental psychobiology* (4th ed., pp. 573–687). New York: Wiley, 1983.

BEDDOR, P. S., GOTTFRIED, T. L. Methodological Issues in Cross-Language Speech Perception Research with Adults. In Strange, W. (ed) *Speech*

Perception and Linguistic Experience: Theoretical and Methodological Issues. Timonium, MD: York Press. pp. 207-232, 1995.

BEST, C. T. "A direct realist view of cross-language speech perception". In: W. Strange (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research*, p. 171–203. Timonium, MD: York Press, 1995.

___ & Tyler, M. D. Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins, 2007.

BETTONI-TECHIO, M. *Production of final alveolar stops in Brazilian Portuguese/English interphonology*. 152 f. Dissertação: Mestrado em Letras. Universidade Federal de Santa Catarina, 2005.

_____. *Perceptual training and word-initial /s/-clusters in Brazilian Portuguese/English interphonology*. Tese de doutorado. Universidade Federal de Santa Catarina, 2008.

BLANK, C.; ZIMMER, M. C. "A influência da grafia em tarefa de acesso lexical envolvendo a L2 (francês) e a L3 (inglês) de um multilíngue: uma abordagem via sistemas dinâmicos". *Calidoscópico (UNISINOS)*, v. 9, p. 28-40, 2011.

BLOOM, L., HOOD, L., LIGHTBOWN, P. "Imitation in language development: If, when and why". *Cognitive Psychology* 6(3), p. 380-420, 1974.

BLOOMFIELD, L. *Language*. Holt, Rinehart, and Winston, New York, 1933.

BOERSMA, P.; WEENIK, D. "Praat: Doing phonetics by computer" (Versão 4.4.03). Disponível em: <http://www.praat.org/>, 2006.

BRASILEIRO, I. *The effects of bilingualism on children's perception of speech sounds*. Utrecht: LOT, 2009.

BROWMAN, C. & GOLDSTEIN, L. "Towards an articulatory phonology". In: *Phonology Yearbook*, 3, p. 219-252, 1986.

_____. "Articulatory gestures as phonological units". In: *Phonology*, 6, p. 201-251. 1989.

_____. "Tiers in articulatory phonology, with some implications for casual speech". In: KINGSTON, T.; BECKMAN, M.E. (Eds.). *Papers in Laboratory Phonology I: between grammar and physics of speech*. p. 341-376. Cambridge University Press, 1990.

_____. "Articulatory Phonology: an overview", In: *Phonetica*, 49, p. 155-180, 1992.

_____. "Gestural syllable position effects in American English". In: BELL-BERTI, F.; RAPHAEL, L. J. *Producing speech: Contemporary Issues*. For Katherine Safford Harris. p. 19-33. Nova York: AIP Press, 1996.

BROWN, H. D. "Cognitive and affective characteristics of good language learners". Paper presented at Los Angeles Second Language Acquisition Research Forum, UCLA, February 1977.

BROERSMA, M. "Perception of familiar contrasts in unfamiliar positions." *Journal of the Acoustical Society of America*, 117, 3890–3901, 2005.

CHOMSKY, N. *Aspects of the theory of syntax*, The MIT Press, Cambridge, Ma, 1965.

_____ & HALLE, Morris. *The sound pattern of English*. The MIT Press, 1995 [1968].

_____. *Regras e representações*, Rio de Janeiro: Zahar, 1981.

COOK, V. *Second language learning and language teaching*. London: Edward Arnold, 1991.

CORDER, P. "The Significance of Learners' Errors". *IRAL*, Heidelberg, v.5, n.4, p.161-170, 1967.

DAVIS, M. H., MARSLEN-WILSON, W. D., e GASKELLI, M. G. "Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition". *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), p. 218–244, 2002.

DE BOT, K; VERSPOOR, M.; LOWIE, W. "Dynamical Systems Theory and Applied Linguistics: the ultimate 'so what?'". *International Journal of Applied Linguistics*, 15, p. 116-118, 2005.

DIEHL, R.L., & KLUNDER, K.R. On the objects of speech perception. *Ecological Psychology*, 1, 121-144, 1989.

_____; LOTTO, A. J., HOLT, L. L. "Speech perception". *Annual Review of Psychology*, 55 (1): 149–179, 2004.

DINNSEN, D.; ECKMAN, F. "Some substantive universals in atomic phonology". *Lingua* 45, p. 1-14, 1978.

ECKMAN, F. R. "On the naturalness of interlanguage phonological rules". *Language Learning* 31, p. 195–216, 1981.

_____ & Fred R. "Markedness and the contrastive analysis hypothesis". In: IOUP, Georgette; WEINBERGER, Steven H. *Interlanguage phonology: the acquisition of a second language sound system*. p. 55-69. Newbury, 1987.

EDGE, J. *Mistakes and Correction*. London: Longman, 1989.

ELLIS, N. C. "Implicit and explicit knowledge about language". In: CENOZ, J.; HORNBERGER, N. H. (Eds.). *Encyclopedia of Language and Education*, v. 6, p. 119-132, 1997.

_____; SCHMIDT, R. Morphology and Longer Distance Dependencies: Laboratory Research Illuminating the A in SLA. *Studies in Second Language Acquisition*, 19:145-71, 1997.

_____. “Emergentism, connectionism and language learning”. *Language Learning*, v. 48, n. 4, p. 631-664, 1999.

_____. “Implicit and explicit knowledge about language”. In: CENOZ, J.; HORNBERGER, N. H. (Eds.). *Encyclopedia of Language and Education*, v. 6, p. 119-132, 2007.

ELMAN, J. L., BATES, E. A., JOHNSON, M. H., KARMILOFF-SMITH, A., PARISI, D., PLUNKETT, K. *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press, 1996.

ERNESTUS, M.; BAAYEN, H. “The functionality of incomplete neutralization in Dutch: the case of past-tense formation”. In: GOLDSTEIN, Louis; WHALEN, D. H.; BEST, Catherine D. (Ed). *Laboratory Phonology 8*. Berlin: De Gruyter, 2006.

ESCUDERO, P. *Linguistic Perception and Second Language Acquisition*, Vol. 113. Utrecht: LOT, 2005.

EWERT, A. *The French Language*. Faber and Faber, London, 1933.

FERREIRA-JUNIOR, F. “Uma Interlíngua Conexionista”. *Trabalhos em Linguística Aplicada*, v. 46, p. 219-231, 2007.

FLEGE, J.E. “The production of ‘new’ and ‘similar’ phones in a foreign language: evidence for the effect of equivalence classification”. *Journal of Phonetics*, 15. p. 47-65, 1987.

_____. English speakers learn to suppress stop devoicing. In *Papers from the Eighteenth Regional Meeting of the Chicago Linguistic Society*, ed. by Robert A. Chametzky, R. Hirzel and Kevin Tuite, pp. 111–122, 1982.

_____. "Chinese subjects' perception of the word-final English /t-/d/ contrast: Performance before and after training," *J. Acoust. Soc. Am.* 86, p. 1684–1697, 1989.

_____. Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech*, 1995.

perception and linguistic experience: Issues in cross-language research (pp. 233–277). Timonium, MD: York Press.

_____; MUNRO, M.J., E MACKAY, I. The effect of age of second language learning on the production of English consonants. *Speech Communication* 16, 1- 26, 1995.

FOWLER, C. A., BEST, C. T., e MCROBERTS, G. W. "Young infants' perception of liquid coarticulatory influences on following stop consonants," *Percept. Psychophys.* 48, 559–570, 1990.

FOX, R.A.; TERBEEK, D. "Dental flaps, vowel duration and rule ordering in American English", *Journal of Phonetics* 5, p.27-34, 1977.

FRIES, C. *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press, 1945.

FRISCH, S. A., LARGE, N. R., e PISONI, D. B. "Perception of wordlikeness: Effects of segment probability and length on processing of non word sound patterns". *Journal of Memory and Language*, 42, p. 481–496, 2000.

GAFOS, A. A grammar of gestural coordination. *Natural language and linguistic theory*, 20(2): 269-337, 2002.

GALASSO, J. *Interference in Second Language Learning: A Review of the Fundamental Difference Hypothesis*. Northridge: California State University, 2002.

GASS, S. M. & J.SCHACHTER (eds.). *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press, 1989.

GERRITS, E. *The categorization of speech sounds by adults and children: A study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues*. PhD thesis, University of Utrecht, 2001.

GIBSON, J. J. "A Theory of Direct Visual Perception". In: Alva Noë/Evan Thompson (Eds.), *Vision and Mind. Selected Readings in the Philosophy of Perception*, Cambridge, MIT Press, pp. 77–89, 2002.

GOLDRICK, M. "Turbid output representations and the unity of opacity". *Proceedings of NELS*, 30, p. 231–245. Also available as ROA 368-1299, 2000.

GUION, S. Word frequency effects among homonyms. *Texas Linguistic Forum*. 35.103–16, 1995.

_____; FLEGE, J. E., AKAHANE-YAMADA, R. e PRUITT, J. C. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107(5), 2711-2724, 2000.

HASKELL, T. R., MACDONALD, M. C., E SEIDENBERG, M. S. Language learning and innateness: Some implications of compounds research. *Cognitive Psychology*, 4, 119–163, 2003.

HAWKINS, S.; NGUYEN, N. "Effects on word recognition of syllable-onset cues to syllable-coda voicing". In: LOCAL, J., OGDEN, R., & TEMPLE, R. (Eds.), *Papers in laboratory phonology VI* (pp. 38–57). Cambridge: Cambridge University Press, 2003.

_____. "Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English". *Journal of Phonetics*, 32, p.199–231, 2004.

ITARD, J. M.G. *The Wild Boy of Aveyron*. New York: Meredith House, 1962.

JESPERSEN, O. *Lehrbuch der Phonetik*. Leipzig: B. G. Truebner, 1926.

JUSCZYK, P. W. From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics* [Special issue on phonetic development], 21(1–2), 3–28, 1993.

KEATING, P. “Universal phonetics and the organization of grammars”. In: FROMKIN, V. (Ed.). *Phonetic Linguistics: essays in honor of Peter Ladefoged*. New York: Academic Press, p. 115-132, 1985.

KELSO, J. A. S. *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge: Bradford Books, MIT Press, 1995.

KENT, Ray; READ, Charles. *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group Inc., 1992.

KLEBER, F., JOHN, T. and HARRINGTON, J. “The implications for speech perception of incomplete neutralization of final devoicing in German”. *Journal of Phonetics*, 38, p. 185-196, 2010.

KLUGE, D. C.; REIS, M. S.; NOBRE-OLIVEIRA, D.; BETTONI-TECHIO, M. The use of visual cues in the perception of English syllable-final nasals by Brazilian EFL learners. In: Watkins, M. A.; Rauber, A. S. & Baptista, B. O. (Eds.). *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Cambridge Scholars Publishing, p. 141-153, 2009.

KLUGE, D. C. Brazilian EFL learners'identification of word-final /m-n/: native/nonnative realizations and effect of visual cues. Doctoral dissertation. Universidade Federal de Santa Catarina: Florianópolis, 2009.

KLUGE, D. C. “The effect of visual cues in the perception of nonnative contrasts”. *Signotica* (UFG), v. 22, p. 257-276, 2010.

KRASHEN, S. *Principles and Practice in Second Language Acquisition*, Pergamon, 1982.

LADEFOGED, P. *Elements of acoustic phonetics*. Chicago: University of Chicago Press, 1962.

_____. *A course in Phonetics* – Third Edition. Harcourt Brace College Publishers, 1993.

_____; MADDIESON, I. *The Sounds of the World's Languages*. Oxford: Blackwell, 1996.

LADO, R.: *Linguistics across cultures*. Michigan: University of Michigan Press, 1957.

_____. *Language Teaching – A Scientific Approach*. New York: McGraw-Hill, 1964.

LEATHER, J. “Second-language speech research: an introduction”. In: J. Leather (ed.). *Phonological Issues in Language Learning*, 1-58. Oxford: Basil Blackwell, 1999.

LEE, W. *Thoughts of contrastive linguistics in the context of foreign language teaching*. In: J. Alatis (ed.), *Contrastive Linguistics and its Pedagogical Implications*. Washington: Georgetown University Press, 1968.

LEHISTE, I. *Lectures on Language Contact*. Cambridge, MA: The MIT Press, 1988.

LINDBLOM, B. “Economy of speech gestures”. In: MACNEILAGE, P. (ed.). *The Production of Speech*. p. 217-245, Nova York: Springer, 1983.

LLISTERRI, J. “Relationships between speech production and speech perception in a second language”. In: ELENIOUS, K. & BRANDERUD, P. (eds.), *Proceedings of the 13th International Congress of Phonetic Sciences*, Vol. 4, 92-99. Stockholm: KTH / Stockholm University, 1995.

LOCKE, J. L. “Mechanisms of phonological development in children: Maintenance, learning, and loss,” in *Papers from the Sixteenth Regional Meeting of the Chicago Linguistic Society* (Chicago: Chicago Linguistic Society), pp. 220–238, 1980.

LOPES, C. e ABIB, J. Teoria da percepção no behaviorismo radical. *Psic.: Teor. e Pesq.*, vol.18, no.2, p.129-137, 2002.

MACKEN, M., e BARTON, D. A longitudinal study of the acquisition of the voicing contrast in American English word-initial stops, as measured by voice onset time. *Journal of Child Language*, 7, 41-74, 1980.

MANUEL, S. Y. "Speakers nasalize /o/ after /n/, but still hear /o/". *Journal of Phonetics*, 23, p. 453–476, 1995.

McCARTHY, J., PRINCE, A. "Faithfulness and reduplicative identity". *UMOP* 18, p. 249–384, 1995.

McCLELLAND, J.L. e ELMAN, J.L. "The TRACE model of speech perception". *Cognitive Psychology* 18 (1): 1–86, 1986.

McGURK, H., e McDONALD, J. "Hearing lips and seeing voices". *Nature*, 264, p. 746–747, 1976.

MAJOR, Roy C., IOUP, Georgette e WEINBERGER, Steven H. (eds.) *A model for interlanguage phonology. Interlanguage phonology: the acquisition of a second language sound system..* New York: Newbury House/Harper & Row, 1987,p. 101-125 .

MINAGAWA-KAWAI, Y., MORI, K., NAOI, N., KOJIMA, S. "Neural Attunement Processes in Infants during the Acquisition of a Language-Specific Phonemic Contrast". *The Journal of Neuroscience* 27 (2): 315–321, 2006.

MOULTON, W.G. *The Sounds of English and German*. Chicago: University of Chicago Press, 1962.

MUNSON, Benjamin, EDWARDS, Jan, SCHELLINGER, Sarah K., BECKMAN, Mary E. e MEYER, Marie K. *Deconstructing phonetic transcription: Covert contrast, perceptual bias, and na extraterrestrial view of Vox Humana* *Clinical Linguistics and Phonetics*, April-May 2010, 24:245-260

MYERS, Scott. "Final devoicing: Production and perception studies". In: BOROWSKY, T., KAWAHARA, S., SHINYA, T., e SUGAHARA, M. (eds.) *Prosody Matters: Essays in Honor of Elisabeth Selkirk*. Londres: Equinox Press, 2011.

NISHIDA, Gustavo. "Em busca de um primitivo de análise: o problema da representação do tap em grupos e em coda". *Letras de Hoje*, v. 42, p. 129-150, 2007.

_____. *A natureza intervocálica do tap em PB*. 2009. Tese de mestrado, UFPR/Curitiba, 2009.

O'DELL, M., e PORT, R. F. "Discrimination of word-final voicing in German". *Journal of the Acoustical Society of America*, 73(S1), S31(A), 1983.

OLIVEIRA, F. R. M. Análise acústica de fricativas e africadas produzidas por japoneses aprendizes de português brasileiro. Tese de mestrado, UFPR/Curitiba, 2011.

OOSTENDORP, Marc van. "Incomplete devoicing in formal phonology". *Lingua* 118, p. 1362–1374, 2007.

OLLER, D., L.; WIEMAN, W. D. e ROSS, C. "Infant babbling and speech". *Journal of Child Language*, 3, p. 1-11, 1976.

PADGETT, J. *Russian voicing assimilation, final devoicing, and the problem of [v] (or, The mouse that squeaked)*. Ms., University of California, Santa Cruz, ROA #528, 2002.

PIKE, K. L. *Tone languages*. Michigan: Univ. of Michigan Press. Ann Arbor, MI, p. 187, 1948.

PIROTH, Hans Georg. JANKER, Peter. "Speaker-dependent differences in voicing and devoicing of German obstruents". *Journal of Phonetics*, v. 32, p. 81-109, 2004.

POERSCH, J. M. “Simulações conexionistas: a inteligência artificial moderna”. *Revista Linguagem em (Dis)curso*, v.4, n.2, jan./jun. 2004:451.

PORT, R. F., e O'DELL, M. L. “Neutralization of syllable-final voicing in German”. *Journal of Phonetics*, 13, p.455–471, 1985.

_____. “The discreteness of phonetic elements and formal linguistics: response to A. Manaster Ramer”, *Journal of Phonetics*, 24(4), p. 491-511, 1996.

_____; LEARY, A. “Against formal phonology”. *Language* 81, p. 927-964, 2005.

PRICE, P. “How can speech technology replicate and complement good language teachers to help people learn language?” *Proceedings of ESCA Workshop on Speech Technology in Language Learning*. Stockholm: Marholmen Conference Center, 1998.

PRINCE, A.; SMOLENSKY, P. “Optimality theory: Constraint interaction in generative grammar”. *Manuscript, Rutgers University and University of Colorado at Boulder*. Available at ROA and published 2004 at Blackwell, 1993.

PYE, S. “Word-final devoicing of obstruents in Russian”. *Cambridge Papers in Phonetics and Experimental Linguistics*, 5, p. 1–10, 1986.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2012. Disponível em: <<http://www.R-project.org>>. Acesso em: 19.04.2012.

REIS, M. Percepção de sons de língua estrangeira pelo Modelo de Assimilação Perceptual. *Domínios de Lingu@gem. Revista Eletrônica de Linguística*, vol. 5, 1, 2011. Disponível em: <<http://www.seer.ufu.br/index.php/dominiosdelinguagem>>. Acesso em: 20.03.2012.

RICE, K; AVERY, P. "On the Interaction between Sonorancy and Voicing". *Toronto Working Papers in Linguistics*, 10, p. 65-82, 1989.

RICHARDS, J. C. e ROGERS, T. S. *Approaches and Methods in Language Teaching: a description and analysis*. p.173. Cambridge: Cambridge University Press, 1985.

ROCCA, P. D. A. *A tecnologia da fala aplicada ao ensino de entoação da língua inglesa para falantes nativos de língua Portuguesa*. Tese de doutorado: São Paulo, Programa de Linguística Aplicada e Estudos da Linguagem, PUC-SP, 2003.

RUBACH, J. *Cyclic and Lexical Phonology: The Structure of Polish*. Foris: Dordrecht, 1984.

RUMELHART, D. E., HINTON, G. E., and WILLIAMS, R. J. "Learning internal representations by error propagation". In: RUMELHART, D. E., MCCLELLAND, J. L., and the PDP Research Group, editors, *Paralleled Distributed Processing. Explorations in the Microstructure of Cognition. Volume 1: Foundations*. The MIT Press, Cambridge, MA, p. 318-362, 1986.

SANCIER, Michele L.; FOWLER, Carol A. "Gestural drift in a bilingual speaker of Brazilian Portuguese and English". In: *Journal of Phonetics*, v. 25, p. 421-436, 1997.

SCHWARTZ, B. "On explicit and negative data effecting and affecting competence and linguistic behavior". *Studies in Second Language Acquisition*, 15, p. 147-163, 1993.

SELINKER, L.: "Interlanguage". *International Review of Applied Linguistics*, v.10, p. 209-231, 1972.

SELKIRK, E. The syllable. In: HULST, F.V. & SMITH, N. (orgs.) *The structure of phonological representations*. Dordrecht: Foris, p.337-383, 1982.

SHRAGER, M. "Neutralization of Word-Final Voicing in Russian". *Journal of the Acoustical Society of America*, Vol. 112, No. 5, 2002.

SIEVERS, E. *Grundzüge der Phonetik*. 5a. ed. Leipzig: Breitkopf und Härtel, 1901.

SILVA, A. H. P. “Proposta de representação dinâmica para os róticos iniciais do PB”. *Língua Viva* (Guajará-Mirim), EDUFRO - Guajará-Mirim, v. 02, p. 25-30, 2003.

_____. “Primitivos fonológicos de tempo extrínseco vs. primitivos de tempo intrínseco”. *Fórum Linguístico*, v. 5, p. 1-12, 2009.

SLOWIACZEK, L., DINNSEN, D. “On the neutralizing status of polish wordfinal devoicing”. *Journal of Phonetics*, 13, p. 325–341, 1985.

_____; SZYMANSKA, H. “Perception of word-final devoicing in polish”. *Journal of Phonetics*, 17, p. 205–212, 1989.

SOKOLIK, M. E. “Learning without rules: PDP and a resolution of the adult language learning paradox”. *TESOL Quarterly*, v. 24, n. 4, p. 685-696, 1990.

_____; SMITH, M. Assignment of Gender to French Nouns in Primary and Secondary Panguage: A Connectionist Model. *Second Language Research* 8, 1:39-58, 1992.

SMITH, N. V. *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press, 1973.

_____, B.L. A phonetic analysis of consonantal devoicing in children’s speech. *Journal of Child Language* 6, 19 –28, 1979.

STEVIK, E.W. *Memory, Meaning and Method*. Rowley, M.A.: Newbury House, 1976.

TRINDADE, M. J. G. A “Magnetoencefalografia, Aplicações Clínicas”. *Acta Médica Portuguesa*, Lisboa, v. 17, n. 3, p. 231-240, Mai/Jun. 2004. Disponível em: <http://www.actamedicaportuguesa.com/pdf/2004-17/3/231-240.pdf>. Acesso em: Jan, 2012.

VELTEN, H. V. The growth of phonemic and lexical patterns in infant language. *Language* 19, 281–292, 1943.

VIANNA, B. *Nós primatas em linguagem: relações linguísticas como um processo biológico*. UFMG/Belo Horizonte, 2006. Tese (Doutorado em Estudos Linguísticos) - Faculdade de Letras, Universidade Federal de Minas Gerais, 2006.

WARNER, N., JONGMAN, A., SERENO, J., & KEMPS, R. J. J. K. “Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch”. *Journal of Phonetics*, 32(2), p. 251-276, 2004.

WAYLAND, R. P. The relationship between identification and discrimination in cross-language perception: The case of Korean and Thai. In: BOHN, O.-S.; MUNRO, M.J. (Eds.). *Language Experience in Second Language Speech Learning*. Amsterdam/Philadelphia: John Benjamins, p. 201–218, 2007.

WEINREICH, U.: *Languages in contact: findings and problems*. New York: Linguistic Circle of New York, 1953.

WHITE.L. *Universal Grammar and Second Language Acquisition*. Amsterdam/Philadelphia, Pa.: John Benjamins, 1989.

YANTIS, S. *Visual perception: essential readings*. Psychology Press. p. 7, 2011.

ZIEGLER, J.C., FERRAND, L., MONTANT, M. Visual phonology: the effects of orthographic consistency on different auditory word recognition tasks. *Mem. Cogn.* 32, 732–741, 2004.

ZIMMER, M. C. “A desaspiração de plosivas surdas iniciais na transferência do conhecimento fonético do português brasileiro (L1) para o inglês (L2)”. In: *VI CELSUL, 2004, Florianópolis. VI Encontro do CELSUL - RESUMOS*, 2004. p. 141-141.

_____; ALVES, U. K. . A produção de aspectos fonético-fonológicos da segunda língua: instrução explícita e conexãoismo. *Linguagem & Ensino* (UCPel), v. 9, p. 101-143, 2006.

_____; ALVES, U.K. “A dessonorização terminal na aprendizagem da L2: evidências do *continuum* fonética-fonologia”. *Letras de Hoje*. Porto Alegre, v. 42, n. 3, p. 56-68, setembro 2007.

_____. “A leitura em língua estrangeira e os efeitos da frequência e da consistência do insumo lexical em L2”. *Ling. (dis)curso (Impr.)* [online]. vol.10, n.1, p. 111-131, 2010.

_____. On the status of terminal devoicing as an interlanguage process among Brazilian learners of English. *Ilha do Desterro*, n. 55, p. 41-62, 2008.

10 ANEXOS

10.1 QUESTIONÁRIO SOBRE O PERFIL DOS INFORMANTES BRASILEIROS.

Universidade Federal do Paraná
Curso de Pós-Graduação em Letras -
Aluna: Jeniffer Imaregna Alcantara de Albuquerque
Orientadora: Prof^a Dr^a Adelaide Hercília Pescatori Silva

Por favor, responda às perguntas abaixo. Este questionário tem como objetivo coletar informações que serão utilizadas exclusivamente para a pesquisa conduzida. Nenhum dado de nenhum informante será divulgado.

1. Nome: _____

2. Data: _____

3. Idade: _____

4. Fez inglês em outro curso de línguas além do que está fazendo no Celin.
SIM / NÃO

Caso a resposta seja positiva, mencione por quanto ininterrupto você estudou neste local: _____

5. Qual nível você está frequentando nesse momento no Celin?

6. Há quanto tempo você estuda no Celin?

7. Caso você não seja aluno do nível básico, mencione em que nível você entrou no Celin? _____

8. Quantas horas por semana, além do curso, você se dedica ao estudo da língua inglesa? _____

9. Já viajou para um país de língua inglesa? SIM / NÃO

10. Por quanto tempo? _____

11. Quantos anos você tinha na época?

12. Frequentou escola naquele país? SIM / NÃO

13. Que tipo de escola/ curso?

14. Estuda, estudou ou tem contato com outra língua estrangeira? SIM / NÃO

15. Qual língua?

16. Em que cidade foi criado/a?

17. Você é de Curitiba? SIM / NÃO

18. Caso você não seja de Curitiba, mencione onde nasceu e quanto tempo morou lá. Mencione, também, se você morou por mais de 6 meses em outra cidade que não seja Curitiba.

10.2 QUESTIONÁRIO SOBRE O PERFIL DO GRUPO CONTROLE

Universidade Federal do Paraná
Curso de Pós-Graduação em Letras -
Researcher: Jeniffer Imaregna Alcantara de Albuquerque
Adviser: Prof^a Dr^a Adelaide Hercília Pescatori Silva

Please, answer the questions below.

This questionnaire aims only at getting information to help in the analysis of the data of the present study. None of the collected data will be revealed.

1. Name: _____

2. Date: _____

3. Age: _____

4. State/city: _____

5. Have you ever lived in another state/city?

6. Do you know any other language apart from English? Which?

7. Do understand/speak Portuguese?

10.3 PERFIL DOS INFORMANTES

10.3.1 Perfil dos informantes do GE Básico

Informante	Idade	Sexo	Nível no Celin	Horas de estudo no Celin	Naturalidade	Grau de escolaridade	Outra língua estrangeira	Viajou para país de língua inglesa
B1	27	M	B2	90	Curitiba	ESI	Não.	Não.
B2	25	M	B2	90	Curitiba	ESI	Não.	Não.
B3	26	M	B2	120	Curitiba	ESI	Não.	Não.
B4	26	M	B2	90	Curitiba	ESI	Não.	Não.
B5	30	M	B2	90	Curitiba	ESC	Não.	Não.
B6	33	M	B2	120	Curitiba	ESC	Não.	Não.
B7	28	M	B2	90	Curitiba	ESI	Não.	Não.
B8	28	M	B2	120	Curitiba	ESI	Não.	Não.
B9	25	M	B2	90	Curitiba	ESI	Não.	Não.
B10	26	F	B2	120	Curitiba	ESI	Não.	Não.
B11	26	F	B2	120	Curitiba	ESI	Não.	Não.
B12	37	F	B2	90	Curitiba	ESC	Não.	Não.
B13	36	F	B2	120	Curitiba	ESC	Não.	Não.

B14	32	F	B2	90	Curitiba	ESC	Não.	Não.
B15	38	F	B2	120	Curitiba	ESC	Não.	Não.
B16	32	F	B2	90	Curitiba	ESC	Não.	Não.
B17	31	F	B2	90	Curitiba	ESC	Não.	Não.
B18	36	F	B2	120	Curitiba	ESC	Não.	Não.
B19	25	M	B2	90	Curitiba	ESI	Não.	Não.
B20	28	M	B2	120	Curitiba	ESI	Não.	Não.

Notação: B2 =Nível Básico 2; ESC= Ensino Superior Completo; ESI=Ensino Superior Incompleto.

10.3.2 Perfil dos informantes do GE Pré- Intermediário

Informantes	Idade	Sexo	Nível no Celin	Horas de estudo no Celin	Naturalidade	Grau de escolaridade	Outra língua estrangeira	Viajou para país de língua inglesa
P1	26	M	P2	300	Curitiba	ESI	Não.	Não.
P2	31	F	P2	340	Curitiba	ESI	Não.	Não.
P3	34	F	P2	340	Curitiba	ESI	Não.	Não.
P4	27	F	P2	340	Curitiba	ESI	Não.	Não.
P5	25	F	P1	300	Curitiba	ESI	Não.	Sim.
P6	38	M	P1	300	Curitiba	ESI	Não.	Não.
P7	32	M	P1	300	Curitiba	ESI	Não.	Não.

P8	26	M	P2	340	Curitiba	ESI	Não.	Não.
P9	25	M	P1	300	Curitiba	ESI	Não.	Não.
P10	26	M	P2	340	Curitiba	ESI	Não.	Não.
P11	26	M	P2	340	Curitiba	ESI	Não.	Não.
P12	37	M	P1	300	Curitiba	ESC	Não.	Não.
P 13	29	F	P2	340	Curitiba	ESC	Não.	Não.
P 14	33	F	P1	300	Curitiba	ESC	Não.	Não.
P 15	38	F	P2	340	Curitiba	ESC	Não.	Não.
P 16	25	F	P1	300	Curitiba	ESC	Não.	Não.
P17	31	F	P1	300	Curitiba	ESC	Não.	Não.
P18	37	F	P2	340	Curitiba	ESI	Não.	Não.
P19	28	F	P1	300	Curitiba	ESI	Não.	Não.
P20	28	F	P1	300	Curitiba	ESI	Não.	Não.

Notação: P1/P2=Nível Pré-Intermediário 1 (P1) e 2 (P2); ESC= Ensino Superior Completo; ESI=Ensino Superior Incompleto.

10.3.3 Perfil dos informantes do GE Intermediário

Informantes	Idade	Sexo	Nível no Celin	Horas de estudo no Celin	Naturalidade	Grau de escolaridade	Outra língua estrangeira	Viajou para país de língua inglesa
I1	27	F	I4	540	Curitiba	ESC	Não.	Não.
I2	25	M	I4	540	Curitiba	ESC	Não.	Não.

I3	26	F	I4	540	Curitiba	ESI	Não.	Não.
I4	26	F	I4	540	Curitiba	ESI	Não.	Não.
I5	30	M	I3	480	Curitiba	ESI	Não.	Não.
I6	33	M	I3	480	Curitiba	ESI	Não.	Não.
I7	28	M	I3	480	Curitiba	ESI	Sim (Espanhol)	Não.
I8	28	M	I4	540	Curitiba	ESI	Não.	Sim.
I9	25	M	P1	480	Curitiba	ESI	Não.	Não.
I10	26	M	I4	540	Curitiba	ESI	Não.	Não.
I11	26	M	I4	540	Curitiba	ESI	Não.	Não.
I12	37	M	I3	480	Curitiba	ESI	Não.	Não.
I13	36	F	I4	540	Curitiba	ESC	Não.	Não.
I14	32	F	I3	480	Curitiba	ESI	Não.	Não.
I15	38	F	I4	540	Curitiba	ESC	Não.	Não.
I16	32	F	I3	480	Curitiba	ESC	Não.	Não.
I17	31	F	I3	480	Curitiba	ESC	Não.	Não.
I18	36	F	I4	540	Curitiba	ESI	Não.	Não.
I19	25	F	I3	480	Curitiba	ESC	Não.	Não.
I20	28	F	I3	480	Curitiba	ESC	Não.	Não.

Notação: I3/I4=Nível Intermediário 3 (I3) e 4 (I4); ESC= Ensino Superior Completo; ESI=Ensino Superior Incompleto.

10.4 RESULTADOS DO TESTE DE DISCRIMINAÇÃO POR INDIVÍDUO

Nível	BILABIAIS			ALVEOLARES			VELARES		
	PDC	PDV	PV	TDC	TDV	TV	KDC	KDV	KV
Básico	3	1	2	3	2	0	1	2	0
Básico	3	0	0	2	1	1	3	3	1
Básico	2	1	3	3	0	1	3	3	3
Básico	3	0	3	3	1	0	2	1	1
Básico	2	2	1	3	2	0	3	2	1
Básico	3	1	3	1	2	0	3	3	2
Básico	3	1	0	1	3	2	3	3	2
Básico	2	1	3	2	2	0	3	3	1
Básico	3	3	2	1	1	1	0	1	2
Básico	3	0	3	3	1	2	2	0	2
Básico	2	1	0	3	2	0	3	1	2
Básico	0	0	0	3	2	0	3	2	0
Básico	2	2	0	1	2	2	2	2	2
Básico	2	0	0	3	2	0	3	2	1
Básico	3	1	1	2	3	2	0	0	0
Básico	2	0	3	3	3	0	3	2	0
Básico	2	3	2	2	3	1	3	2	0
Básico	1	1	1	0	0	1	0	2	0
Básico	1	0	1	3	3	1	3	3	0
Básico	2	0	1	1	3	2	2	1	1
Pré-Intermediário	3	2	3	2	2	1	3	2	3
Pré-Intermediário	3	1	0	3	2	0	1	1	1
Pré-Intermediário	3	1	2	3	3	1	3	2	2
Pré-Intermediário	3	1	3	3	2	1	3	2	1
Pré-Intermediário	1	3	3	3	3	2	2	0	0
Pré-Intermediário	3	1	3	3	2	0	3	2	0
Pré-Intermediário	1	0	0	0	0	1	2	1	1
Pré-Intermediário	3	1	2	2	2	1	1	1	0
Pré-Intermediário	3	1	2	3	3	1	3	2	2
Pré-	3	1	3	2	3	1	3	3	2

Intermediário									
Pré-Intermediário	3	1	2	3	3	1	3	2	2
Pré-Intermediário	3	3	3	3	2	3	3	2	2
Pré-Intermediário	2	2	0	3	3	1	3	3	2
Pré-Intermediário	1	0	0	0	0	0	0	0	1
Pré-Intermediário	3	3	2	3	3	1	3	3	3
Pré-Intermediário	3	2	1	3	3	1	2	3	1
Pré-Intermediário	3	0	2	3	2	1	3	3	0
Pré-Intermediário	3	1	2	3	3	1	3	2	2
Pré-Intermediário	3	2	3	3	2	1	3	2	2
Pré-Intermediário	3	2	3	2	2	0	2	3	0
Intermediário	3	1	2	2	3	3	2	3	3
Intermediário	3	3	3	3	2	2	3	3	2
Intermediário	2	1	0	1	0	0	3	0	0
Intermediário	3	2	3	3	3	2	3	2	1
Intermediário	3	1	1	3	3	0	1	2	2
Intermediário	3	3	3	3	2	1	3	3	2
Intermediário	3	1	2	2	3	3	2	3	3
Intermediário	2	2	2	2	1	0	3	2	3
Intermediário	3	1	2	2	3	3	2	3	3
Intermediário	1	1	1	1	1	1	0	1	1
Intermediário	3	1	2	3	3	0	3	3	0
Intermediário	3	2	3	3	1	0	3	3	1
Intermediário	3	1	1	3	3	0	3	3	0
Intermediário	3	1	1	3	3	0	3	3	0
Intermediário	3	1	1	2	2	1	2	2	2
Intermediário	2	0	2	2	2	0	3	2	0
Intermediário	2	2	1	3	3	0	3	2	0
Intermediário	1	1	0	3	3	1	3	3	0
Intermediário	3	0	3	2	0	3	3	2	0
Intermediário	3	0	3	2	0	3	3	2	0

Notação: PDC= consoante surda e consoante vozeada sem manipulação (Dado Completo); DV= consoante surda e consoante vozeada com manipulação de Duração de Vogal; V= consoante surda e consoante vozeada com manipulação de vozeamento.

Pré-Intermediário	3	3	3	3	3	1	3	3	0
Pré-Intermediário	3	2	3	3	3	1	3	2	0
Pré-Intermediário	3	1	3	3	3	3	3	3	3
Pré-Intermediário	3	0	3	3	2	2	2	2	2
Pré-Intermediário	3	3	3	3	2	3	3	3	2
Pré-Intermediário	3	1	2	3	1	1	3	3	0
Pré-Intermediário	3	2	1	3	3	1	3	3	0
Pré-Intermediário	3	0	3	3	3	3	3	3	3
Pré-Intermediário	3	1	3	3	1	2	3	3	3
Pré-Intermediário	3	2	3	2	1	2	3	3	2
Intermediário	3	1	3	3	2	3	3	3	2
Intermediário	3	3	3	3	3	3	3	3	3
Intermediário	3	1	3	3	3	3	3	3	3
Intermediário	3	1	3	3	2	3	3	3	3
Intermediário	3	0	3	3	3	2	3	3	0
Intermediário	3	3	3	3	3	3	3	3	2
Intermediário	3	2	2	3	3	3	3	1	0
Intermediário	3	1	3	3	3	3	3	3	3
Intermediário	3	0	3	3	1	3	3	3	3
Intermediário	3	1	3	3	3	3	3	3	3
Intermediário	3	1	3	3	3	3	3	3	2
Intermediário	3	3	3	3	3	3	3	3	3
Intermediário	3	3	3	3	3	3	3	3	3
Intermediário	3	3	3	2	3	3	3	3	3
Intermediário	3	2	0	3	2	1	3	3	1
Intermediário	3	3	3	3	3	2	3	3	1
Intermediário	3	1	3	3	3	3	3	3	3
Intermediário	3	3	3	3	2	1	3	3	2
Intermediário	3	1	3	3	3	3	3	3	3
Intermediário	0	1	3	2	2	2	2	3	2

Notação: DC= Consoante vozeada sem manipulação (Dado Completo); DV= Consoante vozeada com manipulação de Duração de Vogal; V= Consoante vozeada com manipulação de vozeamento.