

UNIVERSIDADE FEDERAL DO PARANÁ

SUSANA PINHEIRO DA CRUZ PRESTES

**PRODUÇÃO DE CONSOANTES OCLUSIVAS INICIAIS DO INGLÊS
POR FALANTES NATIVOS DE PB**

CURITIBA

2013

UNIVERSIDADE FEDERAL DO PARANÁ

SUSANA PINHEIRO DA CRUZ PRESTES

**PRODUÇÃO DE CONSOANTES OCLUSIVAS INICIAIS DO INGLÊS
POR FALANTES NATIVOS DE PB**

Dissertação em Estudos Linguísticos
apresentada ao Curso de Pós-Graduação em
Letras, Setor de Ciências Humanas, Letras e
Artes da Universidade Federal do Paraná, como
requisito parcial à obtenção do grau de Mestre
em Linguística.

Orientadora: Prof.^a Dr.^a Adelaide Hercília
Pescatori Silva

CURITIBA

2013



Ata sexcentésima vigésima quarta, referente à sessão pública de defesa de dissertação para a obtenção de título de mestre a que se submeteu a mestranda **SUSANA PINHEIRO DA CRUZ PRESTES**. No dia seis de dezembro de dois mil e treze, às quatorze horas, na sala 1005B, 10.º andar, no Edifício Dom Pedro I, do Setor de Ciências Humanas da Universidade Federal do Paraná, foram instalados os trabalhos da Banca Examinadora, constituída pelos seguintes Professores Doutores: **ADELAIDE HERCÍLIA PESCATORI SILVA**, Presidente, **DENISE CRISTINA KLUGE** e **UBIRATÃ KICKHOFEL ALVES**, designados pelo Colegiado do Curso de Pós-Graduação em Letras, para a sessão pública de defesa de dissertação intitulada: “PRODUÇÃO DE CONSOANTES OCLUSIVAS INICIAIS DO INGLÊS POR FALANTES NATIVOS DO PB”, apresentada por **SUSANA PINHEIRO DA CRUZ PRESTES**. A sessão teve início com a apresentação oral da mestranda sobre o estudo desenvolvido. Logo após a senhora presidente dos trabalhos concedeu a palavra a cada um dos Examinadores para as suas arguições. Em seguida, a candidata apresentou sua defesa. Na sequência, a Professora **ADELAIDE HERCÍLIA PESCATORI SILVA** retomou a palavra para as considerações finais. Na continuação, a Banca Examinadora, reunida sigilosamente, decidiu pela aprovação da candidata. Em seguida, a senhora Presidente declarou **APROVADA** a candidata, que recebeu o título de **Mestre em Letras**, área de concentração **Estudos Linguísticos**, devendo encaminhar à Coordenação em até 60 dias a versão final da dissertação. Encerrada a sessão, lavrou-se a presente ata, que vai assinada pela Banca Examinadora e pela candidata. Feita em Curitiba, no dia seis de dezembro de dois mil e treze. xxxxxxxxxxxxxxxxxxxx

Dr.^a Adelaide Hercília Pescatoti Silva

Dr.^a Denise Cristina Kluge

Dr. Ubiratã Kickhofel Alves

Susana Pinheiro da Cruz Prestes



PARECER

Defesa de dissertação da mestranda SUSANA PINHEIRO DA CRUZ PRESTES para obtenção do título de **Mestre em Letras**.

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

“PRODUÇÃO DE CONSOANTES OCLUSIVAS INICIAIS DO INGLÊS POR FALANTES NATIVOS DO PB”

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
DENISE CRISTINA KLUGE		Aprovado
UBIRATÃ KICKHOFEL ALVES		Aprovado

Curitiba, 06 de dezembro de 2013.

Prof. Dr. Rodrigo Tadeu Gonçalves
Vice-Coordenador

DEDICATÓRIA

Ao meu precioso Vinicius

AGRADECIMENTOS

A Deus, que me concedeu a saúde e a perseverança para concluir este trabalho.

Aos meus pais, João e Raimunda, a quem devo minha vida, assim como a integridade física e equilíbrio emocional. Obrigada pelo cuidado, pelos ensinamentos e pelo amor.

A meu esposo, pelo apoio, pelo amor e por me fazer sentir que estou no caminho certo e que minha conquista é motivo de seu orgulho. Obrigada por me manter bem alimentada e feliz!

Aos estimados colegas do Celin e da UFPR, especialmente à Inês, Álvaro, Iara, Anderson, Gislaine, Luís Gustavo e Ana Paula, por terem, de uma forma ou de outra, contribuído para o andamento desta pesquisa.

Às pessoas que disponibilizaram seu tempo para serem meus sujeitos de pesquisa.

À Jeniffer, por estar disposta a me ajudar sempre que necessário e a oferecer palavras de apoio e encorajamento nos momentos certos.

À Prof.^a Dr.^a Maria Lúcia de Castro Gomes, que gentilmente disponibilizou seu tempo para me ajudar nas gravações finais. Agradeço ainda pela leitura do meu trabalho e as sugestões feitas.

Aos professores doutores Denise Cristina Kluge e Ubiratã Kickhofel Alves, pela leitura criteriosa deste trabalho e valiosas contribuições nas bancas de qualificação e defesa. Agradeço, em especial, à Prof.^a Dr.^a Denise, por ter sido tão solícita em socorrer-me com a parte estatística desta pesquisa e por colaborar com importantes sugestões.

À Prof.^a Dr.^a Clarissa Jordão, pela compreensão e encorajamento para seguir minhas inclinações acadêmicas. Agradeço também pelas discussões feitas durante as aulas, que me fizeram ver várias questões sob outro prisma.

E, finalmente, agradeço de forma especial a minha orientadora, a Prof.^a Dr.^a Adelaide Hercília Pescatori Silva, por ter-me aceitado quando resolvi mudar de projeto e abraçar a área da fonética e fonologia. Obrigada pelo conhecimento compartilhado em suas aulas, pela paciência, competência e gentileza.

RESUMO

Na presente pesquisa, investigamos a produção de consoantes oclusivas surdas e sonoras do inglês em início de palavra por falantes de português brasileiro (PB). Para tanto, analisamos dados de sujeitos de três grupos diferentes: falantes nativos de inglês norte-americano, brasileiros falantes proficientes de inglês e brasileiros aprendizes de inglês em nível intermediário. Utilizou-se como parâmetro acústico o VOT (*Voice Onset Time*) em sua duração relativa a fim de testar as seguintes variáveis: nível de proficiência, ponto de articulação, altura da vogal seguinte e posição ântero-posterior da vogal seguinte.

Para as oclusivas surdas, foi atestada uma associação positiva entre o nível de proficiência do aprendiz e sua capacidade crescente em aproximar-se dos sons da L2, corroborando pesquisas anteriores (ZIMMER, 2004; FRANÇA, 2011). Já a premissa de que quanto mais posterior o ponto de articulação, mais longa é a duração de VOT nas oclusivas surdas (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1999; YAVAS, 2008; FRANÇA, 2011; SCHWARTZHAUPT, 2012) foi confirmada apenas para os grupos de informantes brasileiros. A duração de VOT das oclusivas surdas diante de vogal alta foi, na maior parte dos casos, mais longa do que diante de vogal baixa na produção dos grupos, corroborando em parte estudos anteriores (PORT; ROTUNNO, 1979; OHALA, 1981; YAVAS, 2007). Quanto à influência da posição ântero-posterior da vogal seguinte, nossos resultados coincidem em parte com aqueles verificados por Schwartzhaupt (2012), pois foi apurado VOT mais longo para oclusivas diante de vogal posterior nos pontos bilabial e alveolar, mas diante de vogal anterior no ponto velar.

As oclusivas sonoras tiveram sua realização variando basicamente de duas maneiras: ora o início do vozeamento antecedia o *burst*, resultando em VOT de valores negativos; ora o início do vozeamento ocorria concomitantemente ao momento da soltura ou alguns segundos depois dela, resultando em VOT zero ou de valores positivos próximos a zero, o que é considerado VOT de intervalo curto. Observamos haver uma associação entre o nível de proficiência e a proporção de dados produzidos com VOT de intervalo curto. Além disso, encontramos evidências de que as oclusivas com VOT de intervalo curto comportam-se de maneira diversa daquelas com VOT de intervalo longo no que diz respeito à influência do ponto de articulação e altura da vogal seguinte. Em outras palavras, embora ambas exibam valores positivos de VOT, estratégias articulatórias diferentes parecem estar em jogo em sua produção.

Palavras-chave: consoantes oclusivas, VOT, produção.

ABSTRACT

In this study, we investigated the production of voiced and unvoiced stops of English in word-initial position by speakers of Brazilian Portuguese. In order to do so, we analyzed data from subjects belonging to three different groups: native speakers of American English, Brazilian proficient speakers of English and Brazilian learners of English at intermediate level. The acoustic parameter used in the analyses was the VOT (Voice Onset Time) in its relative duration to test the following variables: level of proficiency, place of articulation, height and frontness of the following vowel.

For voiceless stops, we found positive relationship between the level of proficiency of learners and their increasing ability to approach the sounds of L2, which corroborates previous studies (ZIMMER, 2004; FRANÇA, 2011). The assertion that the farther back the closure, the longer the VOT (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1999; YAVAS, 2008; FRANÇA, 2011; SCHWARTZHAUPT, 2012) was confirmed only for the Brazilian subjects.

The VOT for voiceless stops before a high vowel was, in most cases, longer than before a low vowel, which partially corroborates previous studies (PORT; ROTUNNO, 1979; OHALA, 1981; YAVAS 2007). Regarding the influence of the frontness of the following vowel, our results partially coincide with those found by Schwartzhaupt (2012), as we found the longest VOTs before a back vowel for bilabial and alveolar stops, but before a front vowel for velar stops.

The voiced plosives occurred basically in two ways: either the onset of voicing preceded the burst, resulting in negative VOT values or the start of voicing occurred concurrently with the time of release or a few seconds after it, resulting in short lag VOT. We observed a positive relationship between the level of proficiency and the proportion of data produced with short lag VOT. Moreover, we found evidence that short-lag plosives behave differently from long-lag ones with respect to the place of articulation and the height of the following vowel. In other words, even though they both exhibit positive VOT, different articulatory strategies seem to be at stake in their production.

Keywords: stops, VOT, production.

LISTA DE FIGURAS

Figura 1: Representação da configuração do trato vocal na articulação das oclusivas bilabial, alveolar e velar. Fonte: Edwards, 1992.....	15
Figura 2: Oscilograma e espectrograma da palavra <i>gap</i> . Entre as linhas pontilhadas destaca-se o momento da oclusão. Fonte: <i>corpus</i> da autora	17
Figura 3: Oscilograma e espectrograma da palavra <i>kit</i> . No intervalo entre as linhas pontilhadas temos a consoante [k]. O burst é a barra vertical mais escura no espectrograma. Fonte: <i>corpus</i> da autora	19
Figura 4: Oscilograma e espectrograma da palavra <i>bop</i> . Note a trajetória levemente ascendente de F2. Fonte: <i>corpus</i> da autora	20
Figura 5: Oscilograma e espectrograma da palavra <i>dop</i> . Note a trajetória marcadamente descendente de F2. Fonte: <i>corpus</i> da autora	21
Figura 6: Oscilograma e espectrograma da palavra <i>gosh</i> . Note a trajetória ligeiramente descendente de F2. Fonte: <i>corpus</i> da autora	21
Figura 7: Oscilograma e espectrograma da palavra <i>back</i> . Entre as linhas pontilhadas, destaca-se a duração da palavra-alvo. Logo em seguida, é possível divisar a oclusão proveniente do [t] na expressão <i>to me</i> que completa a sentença-veículo. Fonte: <i>corpus</i> da autora.	65
Figura 8: Oscilograma e espectrograma da palavra <i>back</i> . Note que não há como saber o limite entre a duração da palavra-alvo e a oclusão proveniente do [t] em <i>to me</i> . Fonte: <i>corpus</i> da autora.	65
Figura 9: Oscilograma e espectrograma da palavra <i>keek</i> . A medição de VOT começa a partir do último burst da sequência. Fonte: <i>corpus</i> da autora. .	66
Figura 10 Oscilograma e espectrograma da palavra <i>get</i> . A medição de VOT começa a partir do início da barra de sonoridade até o registro do burst (-66ms). Fonte: <i>corpus</i> da autora.	66
Figura 11: Oscilograma e espectrograma da palavra <i>gook</i> . A medição de VOT começa a partir do início da barra de sonoridade até o registro do burst (50ms). Fonte: <i>corpus</i> da autora.	67
Figura 12: Oscilograma e espectrograma da palavra <i>buff</i> . Não há barra de sonoridade imediatamente anterior ao burst (VOT=zero). Fonte: <i>corpus</i> da autora.	67

LISTA DE TABELAS

Tabela 1: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do inglês norte-americano no estudo de Lisker e Abramson (1964).....	33
Tabela 2: Valores médios em milissegundos apurados para as oclusivas surdas do inglês norte-americano em diferentes estudos.	35
Tabela 3: Médias de VOT em milissegundos encontradas nas oclusivas [p t k] diante de vogal alta posterior [ʊ] e anterior [ɪ], segundo Alves e Dias, 2010.	38
Tabela 4: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do português brasileiro em diferentes estudos.....	39
Tabela 5: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do português brasileiro diante de diferentes vogais. (KLEIN, 1999).....	41
Tabela 6: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do inglês norte-americano em estudo de Lisker e Abramson (1964).....	43
Tabela 7: Valores médios em milissegundos apurados para as oclusivas sonoras do inglês norte-americano em três diferentes estudos.	43
Tabela 8: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro em três estudos diferentes.....	45
Tabela 9: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro em estudo de Klein (1999).....	45
Tabela 10: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro diante de diferentes vogais (KLEIN, 1999).....	46
Tabela 11: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por falantes nativos de PB (FRANÇA, 2011).	48
Tabela 12: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por falantes nativos norte-americanos e falantes de PB com proficiência em inglês (ALVES, 2011).	50
Tabela 13: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por brasileiros falantes proficientes de inglês em diferentes estudos.	51
Tabela 14: <i>Corpus</i> utilizado no presente estudo	59
Tabela 15: Média da duração relativa de VOT de consoantes oclusivas surdas para nativos, falantes proficientes e aprendizes.	71
Tabela 16: Média da duração relativa de VOT de consoantes oclusivas sonoras para nativos, falantes proficientes e aprendizes.	73
Tabela 17: Média e desvio padrão da duração relativa de VOT das oclusivas surdas por ponto de articulação.	80
Tabela 18: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com pré-vozeadas por ponto de articulação.	84
Tabela 19: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.	84
Tabela 20: Média e desvio padrão da duração relativa de VOT das oclusivas surdas por ponto de articulação.	86
Tabela 21: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação.....	88

Tabela 22: Valores médios encontrados para o VOT absoluto das oclusivas sonoras pré-vozeadas do inglês norte-americano em Lisker e Abramson (1964) e neste estudo.	90
Tabela 23: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.	91
Tabela 24: Valores médios encontrados para o VOT absoluto das oclusivas sonoras de intervalo curto do inglês norte-americano neste estudo, em Lisker e Abramson (1964), Klatt (1975) e Macken e Barton (1978)	92
Tabela 25: Média da duração relativa de VOT e desvio padrão para as oclusivas surdas com relação à altura da vogal seguinte.	93
Tabela 26: Resultados do teste de Wilcoxon quanto à influência da altura da vogal seguinte.	95
Tabela 27: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras pré-vozeadas com relação à altura da vogal seguinte.....	96
Tabela 28: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à altura da vogal seguinte ..	97
Tabela 29: Média e desvio padrão da duração relativa de VOT para as oclusivas surdas com relação à posição ântero-posterior da vogal seguinte. ...	99
Tabela 30: Resultados do teste de Wilcoxon para oclusivas surdas quanto à posição ântero-posterior da vogal seguinte.....	100
Tabela 31: Média e desvio padrão das oclusivas sonoras pré-vozeadas com relação à posição ântero-posterior da vogal seguinte.	101
Tabela 32: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à posição ântero-posterior da vogal seguinte.	102

LISTA DE GRÁFICOS

Gráfico 1: Duração relativa de VOT das consoantes oclusivas surdas por informante e por grupo.....	71
Gráfico 2: Diferentes realizações das oclusivas sonoras por informante no estudo-piloto.....	72
Gráfico 3: Duração relativa de VOT de consoantes oclusivas sonoras por informante e por grupo.....	73
Gráfico 4: Duração relativa de VOT de consoantes oclusivas surdas por informante e por grupo diante das vogais alta [i] e baixa [æ].	76
Gráfico 5: Duração relativa de VOT de consoantes oclusivas surdas por informante e por grupo diante das vogais anterior [ɪ] e posteriores [ʊ].	77
Gráfico 6: Média da duração relativa de VOT das consoantes oclusivas surdas por ponto de articulação.	81
Gráfico 7: Variação das oclusivas sonoras por informante no experimento.....	83
Gráfico 8: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação	84
Gráfico 9: Média da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.....	85
Gráfico 10: Média da duração relativa de VOT das consoantes oclusivas surdas por ponto de articulação.	86
Gráfico 11: Duração relativa das oclusivas surdas por informante no grupo N.	87
Gráfico 12: Duração das oclusivas surdas por informante nos grupos P e A.....	88
Gráfico 13: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação	89
Gráfico 14: Média da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.....	91
Gráfico 15: Média da duração relativa de VOT das oclusivas surdas com relação à altura da vogal seguinte.....	93
Gráfico 16: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas com relação à altura da vogal seguinte	96
Gráfico 17: Média da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à altura da vogal seguinte	97
Gráfico 18: Média da duração relativa de VOT para as oclusivas surdas com relação à posição ântero-posterior da vogal seguinte.	99
Gráfico 19: Média da duração relativa de VOT para oclusivas sonoras pré-vozeadas com relação à posição ântero-posterior da vogal seguinte.....	101
Gráfico 20: Média da duração relativa de VOT para oclusivas sonoras de intervalo curto com relação à posição ântero-posterior da vogal seguinte.....	102

Sumário

1. INTRODUÇÃO	1
2. FUNDAMENTAÇÃO TEÓRICA	6
2.1 MODELOS E ABORDAGENS DE AQUISIÇÃO DE SEGUNDA LÍNGUA 6	
2.1.1 A PERSPECTIVA BEHAVIORISTA	6
2.1.2 A PERSPECTIVA INATISTA	7
2.1.3 A PERSPECTIVA COGNITIVISTA	10
2.2 CARACTERIZAÇÃO DAS CONSOANTES OCLUSIVAS	14
2.3 TEMPO DE INÍCIO DO VOZEAMENTO (VOT)	22
2.3.1 A RELAÇÃO ENTRE VOT E O PONTO DE ARTICULAÇÃO	24
2.3.2 OUTROS FATORES QUE INFLUENCIAM O VOT	26
2.3.3 ESTUDOS SOBRE A DURAÇÃO DO VOT EM DIFERENTES LÍNGUAS	28
2.4 A PRODUÇÃO DE CONSOANTES OCLUSIVAS NO INGLÊS E NO PB	32
2.4.1 OCLUSIVAS SURDAS NO INGLÊS	32
2.4.2 OCLUSIVAS SURDAS NO PORTUGUÊS BRASILEIRO	36
2.4.3 OCLUSIVAS SONORAS NO INGLÊS	42
2.4.4 OCLUSIVAS SONORAS NO PORTUGUÊS BRASILEIRO	44
2.5 PRODUÇÃO DE OCLUSIVAS NA AQUISIÇÃO DE INGLÊS POR FALANTES NATIVOS DE PB	46
2.6 CONCLUSÃO	52
3. METODOLOGIA	54
3.1 PERGUNTAS DA PESQUISA E HIPÓTESES	54
3.2 O CORPUS	58
3.3 OS SUJEITOS	59
3.4 OS PROCEDIMENTOS DE COLETA	63
3.5 A ANÁLISE DOS DADOS	64
3.6 A ANÁLISE ESTATÍSTICA	67
3.7 CONCLUSÃO	68
4. O ESTUDO-PILOTO	70
4.1 VOT E NÍVEL DE PROFICIÊNCIA.....	70
4.2 VOT E O PONTO DE ARTICULAÇÃO.....	74
4.3 VOT E ALTURA DA VOGAL SEGUINTE.....	75
4.4 VOT E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL SEGUINTE... 76	
4.5 CONCLUSÃO.....	78
5. O EXPERIMENTO	79
5.1 VOT E O NÍVEL DE PROFICIÊNCIA	80
5.1.1 OCLUSIVAS SURDAS E O NÍVEL DE PROFICIÊNCIA	80

5.1.2	OCLUSIVAS SONORAS E O NÍVEL DE PROFICIÊNCIA.....	82
5.2.	VOT E O PONTO DE ARTICULAÇÃO	85
5.2.1	OCLUSIVAS SURDAS E O PONTO.....	86
5.2.2	OCLUSIVAS SONORAS E O PONTO	88
5.3	VOT E A ALTURA DA VOGAL SEGUINTE	92
5.3.1	OCLUSIVAS SURDAS E A ALTURA DA VOGAL	93
5.3.2	OCLUSIVAS SONORAS E A ALTURA DA VOGAL	95
5.4	VOT E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL SEGUINTE... ..	98
5.4.1	OCLUSIVAS SURDAS E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL	99
5.4.2	OCLUSIVAS SONORAS E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL	100
5.5	CONCLUSÃO.....	103
6	CONSIDERAÇÕES FINAIS	105
7	REFERÊNCIAS BIBLIOGRÁFICAS.....	111
8	APÊNDICES.....	120

1. INTRODUÇÃO

O ensino de pronúncia do inglês tem sido, nas últimas décadas, objeto de inúmeras discussões que envolvem o uso da língua em contextos internacionais e a crescente interação entre falantes que não têm o inglês como primeira língua. (CRYSTAL, 2003; JENKINS, 1988; JENKINS, 2000) Distinguir, portanto, o certo do errado em tal área tem-se tornado uma tarefa complexa, uma vez que adotar o paradigma do falante nativo é algo cada vez menos popular em trabalhos recentes. A própria noção tradicional de falante nativo é discutível e, segundo Kramsch (1999, apud JENKINS, 2000) por ter-se tornado tão diversificada, ela acabou por perder sentido, especialmente quando falamos de uma língua como o inglês, que faz parte do repertório de populações multilíngues como as encontradas na Índia e em vários países da África. Além disso, Crystal (apud JENKINS, 2000) aponta que menos de 3% da população britânica apresenta o sotaque padrão britânico (*Received Pronunciation*) em sua forma pura. Neste cenário surge a perspectiva de ensino de inglês como língua internacional (*English as an International Language – EIL*) ou inglês como língua franca (*English as a Lingua Franca – ELF*).

Na área do ensino de pronúncia, uma alternativa recentemente proposta é a adoção do LFC ou *Lingua Franca Core*¹(JENKINS, 2000), um apanhado de aspectos de pronúncia, que pretende selecionar aquilo que é essencial para a inteligibilidade entre falantes não nativos em contexto de comunicação internacional e aquilo que é dispensável em tal contexto. Baseado em dados empíricos, o *Lingua Franca Core* sugere que se priorize, por exemplo, a instrução da diferença entre os sons das vogais [ɪ] e [i], ao passo que considera desnecessária a ênfase na produção das fricativas dentais [θ] e [ð], cujos sons poderiam ser substituídos por [f] e [v] ou [t] e [d] sem prejuízo da inteligibilidade no contexto EIL.

Apesar de ser tentador abraçar a ideia do inglês como língua internacional, é preciso lembrar que os alunos têm diferentes objetivos e isso deve ser levado em consideração. Dauer (2005), em um artigo sobre o uso do *Lingua Franca Core*, identifica a facilidade com que se passa a aceitar as prioridades do inglês como língua internacional, muito embora não se perceba a fluidez da fronteira entre o ensino de inglês como segunda língua² e seu ensino como língua internacional.

¹ Segundo Jenkins (2000), a constituição do *Lingua Franca Core*, foi realizada após muitos anos de observação de problemas de comunicação em salas de aula e outros ambientes sociais multilíngues, além de inúmeras gravações de alunos de diferentes L1s engajados em situações comunicativas.

² A autora usa o termo 'segunda língua' para referir-se à situação na qual o inglês é aprendido em país onde também é usado como língua nativa. No contexto brasileiro, o inglês é aprendido como 'língua estrangeira'. Entretanto, as observações feitas pela autora cabem, em certa medida, ao nosso contexto.

Para professores de inglês como língua internacional (EIL) que tenham sido treinados em um ambiente de inglês como segunda língua (ESL) ou usem materiais que tenham sido escritos sob a perspectiva do falante nativo como ouvinte é muito fácil aceitar tais prioridades, mesmo que não sejam aplicáveis a sua situação. Professores de inglês como segunda língua geralmente possuem turmas mistas, compostas não somente por alunos estrangeiros que planejam permanecer no país pelo resto da vida, mas também por alunos e profissionais que irão retornar a seus países de origem dentro de alguns meses ou anos. (p. 549) (tradução minha)³

A heterogeneidade de interesses entre aprendizes de inglês que estudam no exterior leva-nos a considerar o quanto isso repercute no esforço que os aprendizes farão para se aproximar em menor ou maior grau do sotaque de um falante nativo. É possível que alguns almejem inserir-se completamente na cultura do país que elegeram para viver permanentemente, enquanto outros busquem uma maneira rápida e prática de comunicar-se em inglês ao voltarem a seus países de origem. O mesmo deve valer para o contexto de ensino brasileiro, uma vez que nossos professores encontram alunos com os mais diversos objetivos quanto à utilização da língua.

Além disso, é preciso cautela ao se conduzir uma pesquisa sobre pronúncia do inglês que leve em conta seu caráter de língua internacional. Quando se afirma que a pronúncia considerada desviante de um dado segmento não interfere na inteligibilidade do inglês para falantes de outras línguas, acabamos por desconsiderar que cada uma dessas línguas possui seu próprio padrão fonético-fonológico e que este irá, inevitavelmente, interferir na forma como tal segmento é percebido. Portanto, seria necessário testar a inteligibilidade de determinada pronúncia para virtualmente todas as línguas que utilizassem o chamado EIL.

É natural, portanto, que diante de um cenário de paradigmas em mudança, torne-se difícil para o professor tomar decisões sobre o que deve ou não ser priorizado em sala de aula em termos de pronúncia. Se por um lado, temos diversas variedades a serem exploradas; por outro, falantes nativos de português brasileiro (doravante PB) ou quaisquer outros pertencentes ao chamado Círculo em Expansão⁴, precisam ter uma variedade na qual se ancorar.

³ It is very easy for teachers of EIL who have been trained in an ESL setting or who use textbooks that have been written primarily from a NS-as-listener perspective to accept those priorities even though they might not be applicable to their situation. ESL teachers often have mixed classes, composed of not only foreign-born students who plan to stay in the host country for the rest of their lives, but also students and professionals who will return to their home country in a few months or a few years. (p. 549)

⁴ Em *Standards, codification and sociolinguistic realism: The English language in the outer circle* (1985), Kachru divide os falantes de inglês em três círculos. O Círculo Interno engloba falantes de inglês em países nos quais ele é primeira língua (Estados Unidos, Reino Unido, Irlanda, etc), no Círculo Externo, o inglês é falado

Enquanto as observações de Jenkins (2000) em seu *Lingua Franca Core* não forem submetidas a testes com maior rigor metodológico, adotar o paradigma do falante nativo, embora não seja ideal, pode acabar se configurando como a alternativa mais viável. Mesmo que saibamos que os chamados sotaques britânico e norte-americano não passam de construtos linguísticos, ainda assim eles fornecem um norte para professores que, de outra forma, estariam em um território em que virtualmente qualquer pronúncia seria inteligível.

Foi justamente com a motivação de analisar com mais detalhe uma particularidade de pronúncia presente no LFC que surgiu a ideia para esta pesquisa. Estamos falando de um aspecto que, segundo Jenkins (2000) não é normalmente incluído, muito menos priorizado nas aulas de pronúncia, embora seja indispensável para a inteligibilidade entre não nativos. A autora (op. cit., p. 144) se refere à “aspiração [h] que antecede as plosivas /p/, /t/ e /k/ quando ocorrem em posição inicial de sílaba tônica. Sem a ajuda desse sopro de ar, um ouvinte terá dificuldade em identificar o som como surdo. Assim, um /p/ não aspirado pode ser confundido com um /b/, um /t/ com /d/ e um /k/ com /g/”. Tradução minha⁵

Assim, para que se tenham maiores condições de avaliar a relevância da instrução explícita desse aspecto da pronúncia do inglês, o primeiro passo é observar de que maneira a produção das oclusivas surdas do inglês por falantes de PB efetivamente difere da produção de falantes nativos da língua. Para tanto, recorreremos à análise acústica da produção de oclusivas por falantes de PB e falantes nativos de inglês norte-americano. Todavia, estenderemos nossa análise às oclusivas sonoras, com a expectativa de contribuir para novos vislumbres e uma compreensão mais aprofundada do fenômeno em foco. Em última análise, desejamos possibilitar que estudos futuros estejam melhor amparados para encontrar formas de levar os falantes brasileiros a refinarem sua pronúncia da língua-alvo.

De todo modo, salientamos que, independentemente da visão de ensino de pronúncia que se tenha, este trabalho é pertinente tanto àqueles que conservam o paradigma do falante nativo como aos que se alinham à perspectiva do inglês como língua internacional, já que o *Lingua Franca Core* (JENKINS, 2000) contempla a aspiração de consoantes oclusivas surdas

como segunda língua em um ambiente multilíngue (Singapura, Índia, Nigéria, etc) e no Círculo em Expansão, o inglês é ensinado como língua estrangeira, sem gozar de status especial no país (Brasil, Japão, China, etc).

⁵ (...) the aspiration [h] following the fortis plosives /p/, /t/, and /k/ when they occur in initial position in a stressed syllable. Without the help of this puff of air, a listener will find it more difficult to identify the sound as voiceless. Thus, an unaspirated /p/ may be mistaken for /b/, a /t/ for /d/, and a /k/ for /g/.

em início de palavra como um aspecto imprescindível para a inteligibilidade entre falantes não nativos.

Feitas essas considerações, o objetivo principal desta pesquisa é analisar acusticamente a produção de consoantes oclusivas surdas e sonoras do inglês produzidas por falantes de PB, comparando-a com a produção de falantes nativos de inglês. Para tanto, dividiremos nossos sujeitos em três grupos: falantes nativos de inglês norte-americano (grupo controle), falantes nativos de PB proficientes em inglês e falantes nativos de PB aprendizes de inglês.

Estamos especificamente interessados em observar a interferência dos seguintes aspectos na produção das oclusivas: nível de proficiência do sujeito, ponto de articulação da consoante, altura da vogal seguinte e posição ântero-posterior da vogal seguinte.

O capítulo 1 apresenta um panorama de diferentes abordagens teóricas sobre aquisição de língua, de modo que seja possível justificar a perspectiva que elegemos, especialmente no que tange à concepção de aquisição do sistema fonético/fonológico e o primitivo de análise aqui adotado.

O capítulo 2 fornece uma visão geral do que define acusticamente as consoantes oclusivas surdas e sonoras no PB e na língua inglesa. Introduce ainda o conceito de *Voice Onset Time* (VOT), ou tempo de início de vozeamento, bem como esclarece a opção pela utilização desse parâmetro acústico na análise de nossos dados. Por fim, são elencados diferentes estudos sobre a produção de consoantes oclusivas surdas e sonoras na língua inglesa e no PB, e de inglês por falantes de PB.

No capítulo 3 delineamos as perguntas e hipóteses desta pesquisa e justificamos as reformulações feitas para o experimento em si, com base nos achados do piloto. Em especial, enfatizamos a necessidade de aprimorar a análise das oclusivas sonoras dividindo com mais precisão os dados de acordo com sua realização em duas categorias: oclusivas com pré-vozeamento e oclusivas com VOT de intervalo curto. Além disso, justificamos a introdução do critério ponto de articulação para refinar a análise da duração de VOT no que diz respeito à altura e posição ântero-posterior da vogal seguinte.

Passamos em seguida aos critérios metodológicos utilizados para a constituição do *corpus*, seleção dos sujeitos e análise dos dados. Nesses quesitos, não foram necessárias modificações após a execução do piloto. Reforçamos nossa opção por apresentar os dados em termos de duração relativa de VOT como uma estratégia de normalização dos dados, a fim de neutralizar quaisquer desvios que venham a existir por conta da taxa de elocução dos informantes e aumentar a confiabilidade dos dados.

O capítulo 4 tem como foco os resultados do estudo-piloto, ao passo que o capítulo 5 detalha a execução e os resultados obtidos com o experimento completo.

As considerações finais são tema do capítulo 6, no qual voltamos às perguntas e hipóteses lançadas pela pesquisa, procurando esclarecê-las à luz dos achados deste estudo. Encerramos com nossas ponderações sobre os resultados obtidos, seus possíveis efeitos para o ensino da língua inglesa e as possibilidades que se abrem para futuros trabalhos.

2. FUNDAMENTAÇÃO TEÓRICA

Há toda sorte de teorias que procuram explicar a maneira como aprendemos uma língua, seja ela materna (L1) ou estrangeira (L2)⁶. Neste capítulo apresentaremos um cenário das principais abordagens teóricas que interessam especialmente à aquisição de L2, bem como modelos que contemplam o contraste entre os sistemas fonológicos de L1 e L2. Tal percurso tem o objetivo de esclarecer a concepção de aquisição de segunda língua e o aporte teórico que servirão de base para o presente estudo.

Este capítulo objetiva ainda fornecer uma visão geral do que define acusticamente as consoantes oclusivas surdas e sonoras no PB e na língua inglesa, bem como apresentar o parâmetro mais comumente utilizado para estudá-las: o VOT, esclarecendo nossa opção pela utilização desse parâmetro acústico na análise de nossos dados. Apresentamos ainda um levantamento de trabalhos cujo foco é a produção das consoantes oclusivas surdas e sonoras na língua inglesa e no PB, bem como a produção das oclusivas de inglês como língua estrangeira por brasileiros. Finalmente, conduzimos uma discussão desses trabalhos, tendo como contraponto a pesquisa aqui realizada.

2.1 MODELOS E ABORDAGENS DE AQUISIÇÃO DE SEGUNDA LÍNGUA

Esta seção apresentará brevemente os principais pressupostos das perspectivas behaviorista, inatista e cognitivista. Quanto à última, analisaremos mais detidamente os modelos conexionista e emergentista, que possuem base dinâmica, e alinham-se à perspectiva adotada neste trabalho.

2.1.1 A PERSPECTIVA BEHAVIORISTA

Nas décadas de 50 e 60 do século XX, os modelos de aquisição de segunda língua foram fortemente influenciados pelas teorias psicológicas em voga na época, em especial a

⁶ Embora saibamos que seja possível a distinção entre os conceitos de segunda língua e língua estrangeira, trataremos os termos como intercambiáveis neste estudo.

teoria behaviorista. Para os adeptos do behaviorismo, a aprendizagem⁷ de uma língua, tanto quanto qualquer outra aprendizagem, institui-se a partir da formação de hábitos, que são moldados por meio de estímulo, resposta e o posterior reforço positivo ou negativo, conforme a situação (SKINNER, 1957). Aqui, rejeita-se qualquer especulação sobre o processo mental envolvido na aprendizagem e erros cometidos pelos aprendizes são vistos negativamente, como consequência da formação de maus hábitos. Especificamente quanto à pronúncia de uma L2, quaisquer desvios da forma-alvo têm origem na percepção equivocada do aprendiz, que pode ter recebido um estímulo incorreto ou simplesmente não ter sido capaz de perceber determinado som a fim de imitá-lo corretamente. Soluciona-se o problema com a repetição intensiva seguida de reforço positivo quando do domínio da forma-alvo.

À abordagem behaviorista está ligada a chamada Hipótese da Análise Contrastiva ou *Contrastive Analysis Hypothesis* (CAH), cujos principais defensores foram Charles Fries (1945), Lado (1957) e Lee (1978). Parte-se do pressuposto de que as diferenças entre duas línguas podem ser usadas para prever dificuldades na aprendizagem da língua-alvo. Assim, quando duas línguas se assemelham, ocorre transferência positiva; quando elas diferem, há transferência negativa, ou seja, interferência (LARSEN-FREEMAN; LONG, 1991). Pesquisas nessa área buscaram prever uma hierarquia de dificuldade no processo de aprendizagem baseada na correspondência entre L1 e L2. Sabendo que as práticas pedagógicas da época eram dominadas pela memorização de diálogos, imitação e automatização, a Hipótese da Análise Contrastiva tinha o papel de antecipar e minimizar os erros do aprendiz a fim de que a formação de maus hábitos pudesse ser evitada ao máximo.

2.1.2 A PERSPECTIVA INATISTA

Críticas ao behaviorismo ganham força em 1959, quando da resenha de Noam Chomsky ao livro *Verbal Behavior* (SKINNER, 1957). Chomsky argumenta que a teoria de Skinner falha em explicar por que a produção dos aprendizes foge à simples reprodução do *input* recebido e por que seus erros trazem marcas sistemáticas de que algo além de mera repetição está ocorrendo, ou seja, há evidências de que os aprendizes estão envolvidos de forma mais complexa na construção de “regras” para guiar sua produção e, portanto, a aprendizagem de uma língua não é apenas resposta a um estímulo fornecido (ELLIS, 1997).

⁷ Consideraremos aqui os termos aquisição e aprendizagem como sinônimos, por não estarmos alinhados à distinção proposta por Krashen (1982).

Além disso, quando a Hipótese da Análise Contrastiva é sujeita a testes empíricos, sérias falhas são reveladas, uma vez que erros que foram previstos acabaram por não se materializar e outros para os quais não havia previsão foram detectados. Por exemplo, Whitman e Jackson (1972, apud LARSEN-FREEMAN; LONG, 1991) testaram 4 diferentes previsões de interferência inglês/japonês segundo a CAH em 2500 estudantes japoneses do ensino secundário por meio de questões de múltipla escolha e do tipo ‘cloze’⁸. Os autores concluíram que a Análise Contrastiva mostrou-se inadequada para prever os problemas que tais aprendizes enfrentaram e que a interferência provou desempenhar um papel muito pequeno na aprendizagem da língua inglesa, ao menos em nível sintático.

Assim, conforme a CAH e o behaviorismo passam a ser desacreditados, entra em cena uma visão que incorpora a predisposição biológica do ser humano em aprender línguas. O behaviorismo é, portanto, suplantado pela teoria mentalista ou inatista de Chomsky (1957, 1965), e a aquisição de uma língua passa a ser vista como um processo que se dá de dentro para fora. Chomsky postula que os indivíduos são dotados de um dispositivo de aquisição da linguagem (*Language Acquisition Device*, LAD) e que a aquisição de uma língua se dá quando somos expostos a *input* linguístico. Em outras palavras, a língua é inata ao indivíduo, devido à existência de um módulo mental que contempla a chamada Gramática Universal. Por meio dela, somos capazes de apreender os princípios universais que nos permitem fazer uma representação mental da língua a que somos expostos. Embora a teorização de Chomsky estivesse voltada à língua materna, ela serviu de base para diversos modelos e abordagens teóricas relativas à aquisição de L2. Para os inatistas, as falhas de produção apresentam uma janela para o que acontece na mente do falante, pois permitem investigar como o aprendiz elabora hipóteses com base no *input* recebido e é assim capaz de criar sentenças nunca antes ouvidas, que são regidas por uma Gramática Universal.

No que tange ao tratamento dos erros na aquisição de L2, o inatismo serve de base ao chamado Modelo de Análise de Erros, cujo principal teórico é Corder (1967). Tal modelo destaca o papel do aprendiz como sujeito ativo no processo de aprendizagem, por meio da análise do *input* e a formulação de hipóteses com relação à L2. Diferentemente da Hipótese da Análise Contrastiva, que buscava prever os erros realizados pelo aprendiz contrastando L1 e L2, o Modelo de Análise de Erros parte dos erros dos aprendizes para explicar semelhanças e diferenças entre as línguas envolvidas no processo. Corder faz a distinção entre *mistake* e

⁸ As questões do tipo ‘cloze’ são normalmente aplicadas para se testar a compreensão de texto ou as habilidades semânticas e sintáticas do aprendiz, que é solicitado a fornecer palavras previamente excluídas de um texto de forma sistemática.

error. Ele aponta que o primeiro pode ser cometido tanto por aprendizes quanto por falantes nativos e é apenas um equívoco causado por fadiga ou falta de atenção, sendo passível de autocorreção. Já o que ele chama de *error* é um desvio sistemático produto do estágio de aquisição da L2 e reflexo da competência do aprendiz que, portanto, não consegue corrigi-lo.

Também ligado à perspectiva inatista é o conceito de Interlíngua do linguista americano Larry Selinker. Refere-se a um sistema linguístico construído pelo aprendiz, que guarda semelhanças com a língua-alvo, mas é em parte dependente da L1. De acordo com Ellis (1997) a interlíngua é um sistema único, regido por regras que dependem de uma gramática mental e recebem influência de ambas as línguas. Possui um caráter transitório e pode ser entendida como um *continuum* em direção à L2. Embora alguns pesquisadores da área afirmem que os sistemas construídos pelos aprendizes são compostos de regras variáveis, outros defendem que os sistemas da interlíngua são homogêneos e que a variabilidade encontra-se no desempenho, não na competência do falante. Postula-se ainda que diferentes erros sejam o reflexo de diferentes estratégias empregadas pelo aprendiz. Erros de omissão, por exemplo, decorrem da estratégia de simplificação de estruturas gramaticais as quais o aprendiz não está pronto para assimilar.

Além do Modelo da Análise de Erros e da Interlíngua, vale destacar ainda a chamada Hipótese de Marcação Diferencial, que se utiliza do conceito de marcação, segundo o qual algumas estruturas são mais básicas ou naturais na língua e, portanto não marcadas, ao passo que outras são mais complexas e menos frequentes na língua, sendo consideradas marcadas. Por exemplo, a estrutura de plural, tanto no inglês quanto no português, é marcada em relação ao singular, pois pressupõe a adição do morfema correspondente. Com relação à aquisição de uma língua estrangeira, Eckman (1977, apud LARSEN-FREEMAN; LONG, 1991) prevê que as áreas de uma L2 que diferem da L1 e são mais marcadas do que na L1 causarão dificuldade de aquisição, porém se forem menos marcadas do que na L1, não haverá dificuldade. A Hipótese de Marcação Diferencial prevê, portanto, que falantes nativos de alemão terão dificuldade em produzir o contraste surda/sonora das obstruintes em final de palavra do inglês (*wet/wed*). Isso porque das três posições nas quais tal contraste pode existir nas línguas, a posição em início de palavra é a mais comum, o que significa ser menos marcada, seguida da posição medial e, finalmente, a posição final é a menos frequente e mais marcada nas línguas. Consequentemente, como as duas línguas diferem nessa particularidade, já que o alemão não conta com o contraste surda/sonora em final de palavra, a Hipótese de Marcação Diferencial prevê corretamente que falantes de alemão terão dificuldade em adquirir o contraste, mas

falantes de inglês não sofrerão para abrir mão desse contraste na aquisição do alemão, já que se trata de um elemento mais marcado em sua L1 do que na L2.

2.1.3 A PERSPECTIVA COGNITIVISTA

Vimos que a perspectiva behaviorista concebe a aprendizagem como dependente essencialmente do *input* recebido, sendo determinada de fora para dentro e moldada via estímulo, resposta e reforço. Por outro lado, as teorias de base inatista fazem o caminho oposto, enfatizando a relevância do indivíduo e seu aparato mental e minimizando a importância do *input*, cujo papel é apenas acionar a Gramática Universal da qual todo falante é dotado.

Há, no entanto, uma terceira perspectiva, representada por teorias que entendem a aprendizagem de uma língua como resultado da relação complexa entre os mecanismos internos do aprendiz e uma série de fatores sociais e ambientais. A linguística cognitivista assume como centrais a percepção e categorização do que ocorre no mundo pelo indivíduo, bem como sua capacidade de organizar e relacionar o conhecimento que progressivamente acumula. As experiências do falante com a língua têm papel decisivo e fazem com que sua produção e compreensão linguística mudem sempre que um novo dado é processado (ELLIS, 2003). Sob essa perspectiva, não se faz distinção entre competência e desempenho, muito menos conta-se com uma gramática inata que será desencadeada por *input* linguístico. O aprendiz opera um sistema de processamento complexo para o qual a informação linguística é semelhante a qualquer outro tipo de informação. Modelos teóricos como o conexionismo e o emergentismo são representantes da perspectiva cognitivista.

Conexionistas postulam que os mecanismos neurais são responsáveis pela geração do conhecimento de forma geral, inclusive o linguístico. Ao contrário dos inatistas, não entendem a linguagem como uma faculdade inata, mas assumem que sua aquisição ocorre por meio de conexões sinápticas, que se fazem e desfazem ao longo da vida de um indivíduo e permitem que o conhecimento linguístico se desenvolva. Ao observar um dado linguístico, o aprendiz é capaz de fazer generalizações, que ocasionam, por sua vez, conexões neurais. O aprendizado se dá quando tais conexões são fortalecidas conforme haja a recorrência dos dados linguísticos ou repetição de experiências já vivenciadas pelo aprendiz. (MITCHELL; MYLES, 2004). Ellis (2003) nos informa sobre existência de experimentos conexionistas com redes neurais artificiais nos quais é possível observar como regularidades são extraídas

de informações nos diversos níveis linguísticos (morfológico, fonológico, semântico, sintático, prosódico). Um mecanismo simples de processamento de linguagem é capaz de encontrar generalizações que governam esses sistemas porque, à medida que novo *input* é inserido, o *input* antigo é revisto e as informações são ajustadas a cada novo estágio, ou seja, o mecanismo de redes neurais artificiais “aprende” como funciona o sistema linguístico, assim como um indivíduo aprende uma língua com base na recorrência dos dados aos quais é exposto.

Entretanto, Ellis (2003) afirma que, na aprendizagem de uma língua, também devemos levar em conta a perspectiva do indivíduo, seu sistema perceptual e de atenção, que fazem com que certas informações sejam retidas e outras não.

Emergentistas compartilham a visão dos conexionistas de que nascemos sem quaisquer mecanismos de aquisição da linguagem. Eles argumentam que a hipótese do inatismo não é plausível, tendo em vista o fato de que a neurobiologia tem recentemente constatado que a arquitetura do córtex cerebral é plástica e altamente dependente da experiência e, portanto, sua especialização é consequência do aprendizado e não um mecanismo inato. Estudos com transplantes heterotópicos têm demonstrado que o córtex auditivo, por exemplo, pode vir a se especializar para a visão. Logo, o mapeamento da área cortical é aparentemente resultado da aprendizagem e não de informação genética que estabelece a representação fixa da gramática universal no córtex. Segundo o emergentismo, a aprendizagem de uma língua se dá a partir de associações, um conjunto de mapeamentos entre as formas fonológicas e conceitos ou intenções comunicativas (ELLIS, 2003). Posto de outra forma, a aprendizagem envolve a análise de padrões e regularidades que emergem quando da exposição a uma língua, cujo desenvolvimento depende de bilhões de associações feitas a partir de seu uso (MITCHELL; MYLES, 2004).

Segundo Zimmer et al (2009) os emergentistas concebem os processos de aquisição de L1 e L2 como dependentes do mesmo mecanismo cognitivo, uma vez que ambos os processos resultam da capacidade de observar simultaneamente diversas limitações probabilísticas, de modo que aspectos que isoladamente não seriam relevantes, adquirem relevância quando processados em conjunto com outros aspectos probabilísticos presentes no *input*.

Tanto o conexionismo quanto o emergentismo estão interessados em compreender de que forma o cérebro humano processa e aprende novas informações e que estratégias são utilizadas para acessar o conhecimento linguístico. Além disso, ambos os modelos apostam em uma relação dinâmica entre os componentes do sistema linguístico, de maneira que uma mudança implementada no sistema fará com que seus componentes interajam de maneira

integrada. Logo, para melhor compreender tais modelos, faz-se necessário conhecer a teoria dos sistemas dinâmicos, que teve sua origem na matemática e se propõe a descrever sistemas com estados numéricos que se desenvolvem através do tempo.

Um sistema dinâmico é aquele cujo comportamento em um dado momento depende de seu estado em um momento anterior, ou seja, a modificação de uma variável tende a afetar todo o sistema. Segundo Blank e Zimmer (2011), a maneira como os elementos de um sistema se encontram em um dado tempo é o que determina seu estado total nesse tempo. Podemos dizer que o sistema configura-se como uma totalidade de estados (espaço de estados) que interagem continuamente e não podem ser separados de forma estanque.

Compreender a aquisição de uma língua como um sistema dinâmico implica concebê-la como um processo de desenvolvimento não linear, cujo comportamento é imprevisível, pois quando uma informação é recebida, pode ser processada de diferentes maneiras, a depender das variáveis atuantes no processo de aprendizagem, que podem ser tanto de natureza cognitiva quanto social. (ELMAN et al., 1996).

Dessa forma, passamos a considerar a aquisição como um processo que se desenvolve e muda com o decorrer do tempo, além de depender da interação de uma série de variáveis. No que tange à aquisição de uma L2 ou L3, existe a associação dos subsistemas linguísticos do aprendiz interagindo com fatores como *input*, *feedback* recebido, ambiente de aprendizagem, idade do aprendiz, sua aptidão, motivação, estilo cognitivo e estratégias de aprendizagem (BLANK; ZIMMER, 2011).

Associa-se à perspectiva dinâmica o modelo da Fonologia Gestual de Browman e Goldstein (1986, 1989, 1990, 1992). Seu primitivo de análise, o gesto articulatório, é dotado de tempo intrínseco, diferentemente do primitivo de análise da fonologia de base gerativa, o chamado traço distintivo, que é dotado de tempo extrínseco. Decorre daí que cada som presente na cadeia da fala pode resultar da ação de um ou mais gestos articulatórios, que podem se sobrepor temporalmente. Para realizar um som, os indivíduos organizam diferentes manobras articulatórias de maneira coordenada, tarefa que executam automaticamente, devido a um conhecimento já internalizado. Os gestos pertencem aos subsistemas oral, laríngeo ou vélico e, em cada subsistema, são especificados o grau e local de constricção de cada gesto, o que permite a representação de variações relativas à sua magnitude gestual e seu tempo de ativação (SILVA, 2003).

Uma diferença essencial entre o tratamento via traços da fonologia gerativa e o tratamento via gestos da fonologia gestual ou articulatória é que no primeiro modelo há a necessidade de uma interface computacional entre a representação fonológica e a realização

fonética, o que é desnecessário no segundo modelo, uma vez que a representação gestual já é a especificação direta dos gestos produzidos (BEST; HALLÉ, 2010).

A adoção do gesto articulatório como primitivo de análise permite que se contemplem detalhes de produção que fogem a categorizações discretas, como é o caso da produção de oclusivas em início de palavra, objeto de investigação deste trabalho. Uma vez que tal fenômeno está ligado à duração do VOT, expressa por meio de um contínuo numérico, um tratamento via traços distintivos não seria adequado, pois somente permitiria a análise dos dados de forma categórica. Logo, o foco no detalhe fonético exige que busquemos um modelo que tenha por primitivo o gesto, pois ele vai muito além da simples caracterização dos tempos de soltura, contemplando também os movimentos de todos os articuladores envolvidos.

Assim, ao tratamos da produção de consoantes oclusivas, a Fonologia Gestual nos permite afirmar que a produção de um falante de PB difere da produção de um falante nativo de inglês quanto à coordenação dos gestos que estão envolvidos na produção do segmento, ao invés de termos de opor as produções como essencialmente distintas.

A Fonologia Gestual contribuiu para a elaboração de um modelo de análise de percepção da fala chamado Modelo de Assimilação Perceptual (*Perceptual Assimilation Model*, PAM), proposto por Catherine Best (1994, 1995). Quanto à aquisição de segunda língua, o PAM-L2 (BEST; TYLER, 2007) busca explicar como se dá a transferência fonético-fonológica entre L1 e L2 no nível segmental. O PAM-L2 postula que na infância, o indivíduo aprende a ignorar detalhes fonéticos que não são importantes na percepção dos sons de sua L1, portanto ele é “treinado” a perceber apenas os sons que são relevantes em sua língua materna. Posteriormente, no aprendizado da L2, ao entrar em contato com um som inexistente em sua L1, o aprendiz tende a assimilar tal som de acordo com as pistas acústicas da L1, em vez de inseri-lo em uma nova categoria. Isso ocorre porque os elementos fônicos da L1 e da L2 interagem em um mesmo espaço fonológico, o que faz com que o aprendiz não se dê conta de quais categorias acústicas pertencem a um ou ao outro sistema. Assim, é necessário que o aprendiz volte a sintonizar o sistema de percepção a fim de perceber detalhes fonéticos relevantes para a aquisição da L2 e ignorar detalhes irrelevantes.

É esse movimento que flagramos na aquisição de oclusivas iniciais do inglês, especialmente as surdas, por aprendizes brasileiros. Sabendo-se que a aspiração de oclusivas surdas em início de palavra não constitui uma pista acústica relevante no PB, nossos aprendizes assimilam esse padrão com base na L1, deixando de produzi-las com o nível de aspiração que as caracteriza no inglês.

Associada à transferência dos padrões do sistema fonológico da L1 para a L2, está a transferência dos padrões de correspondência entre letra e som da L1 para a L2, o conhecimento grafo-fônico-fonológico. Embora o português e o inglês façam uso do mesmo sistema alfabético, a relação entre a forma gráfica e a realização de alguns sons se dá de maneira diferente. Assim, ao entrar em contato com os grafemas da L2, o aprendiz atribui a eles a mesma ativação fonético-fonológica da L1, mesmo que a realização seja diferente (ZIMMER; ALVES, 2006). É o que ocorre com relação às oclusivas, que são foco deste estudo. Embora sua realização possa se dar de maneira diferente nos dois sistemas, a grafia desses segmentos coincide em ambas as línguas, o que passa a reforçar a transferência desse padrão da L1 para a L2.

Diante das perspectivas e modelos teóricos apresentados neste capítulo, a concepção de aquisição de língua que nos parece mais compatível com as expectativas do presente estudo é a de base cognitivista, por conceber a aquisição de uma língua como um processo complexo e dinâmico, dependente de mecanismos mentais e biológicos associados a aspectos sociais e que envolvem o conjunto de experiências do indivíduo ao longo da vida. Além disso, em vista da natureza gradiente de nossos dados, acreditamos que a Fonologia Gestual, cujo primitivo de análise é o gesto articulatório, seja a opção mais adequada a fim de investigar a produção de oclusivas em início de palavra por falantes de PB. Por fim, muito embora o PAM seja um modelo voltado à percepção, consideramos que seus pressupostos são os que mais se alinham aos interesses deste estudo, por oferecerem maior poder explicativo com relação à produção de um detalhe fonético fino que se baseia em faseamento gestual, como é o caso do VOT das oclusivas do inglês.

2.2 CARACTERIZAÇÃO DAS CONSOANTES OCLUSIVAS

De forma geral, as consoantes caracterizam-se pela resistência à passagem do ar egresso dos pulmões em algum ponto do trato vocal. O que as diferencia, no entanto, é o ponto em que o ar é obstruído e a maneira como isso ocorre. Além disso, é possível classificar as consoantes mediante a existência ou não de vibração das pregas vocais.

As consoantes oclusivas, também chamadas plosivas, diferenciam-se das demais por apresentarem um bloqueio momentâneo total do fluxo de ar no interior do trato vocal, ou seja, uma oclusão articulatória total em algum ponto do trato. Embora as consoantes nasais também apresentem tal oclusão, sua classificação leva em conta um fator mais saliente de sua

produção: o abaixamento do véu palatino e o conseqüente escape do ar pela cavidade nasal. Assim sendo, as chamadas consoantes oclusivas são primordialmente aquelas em cuja produção o véu palatino encontra-se levantado, inexistindo, portanto, murmúrio nasal.

As oclusivas podem ser divididas entre surdas e sonoras. Para o PB, tal divisão depende primordialmente da existência ou não de vibração das pregas vocais durante sua produção; caso haja tal vibração, a oclusiva é considerada sonora ou vozeada e, em sua ausência, é classificada como surda ou desvozeada. Na língua inglesa, todavia, a distinção entre uma oclusiva surda e uma sonora pode depender de outras pistas acústicas, já que nem sempre há a vibração das pregas durante sua produção. Devemos atentar, portanto, para pistas como a duração da consoante, que tende a ser maior para oclusivas surdas (VELOSO, 1995), bem como a amplitude do *burst*, que é levemente mais intensa para os segmentos surdos (LADEFOGED; MADDIESON, 1996; BONATTO, 2007). No caso das oclusivas em início de palavra, que são foco deste trabalho, observamos nas surdas a presença da aspiração, que é um evento acústico no qual o ar passa pelas pregas vocais e através da laringe gerando um ruído que se assemelha a um sopro.

Na língua inglesa, bem como no português, a oclusão ou obstrução total ocorre basicamente em três pontos do trato: quando ela se dá pelo encontro dos lábios, indo o lábio inferior de encontro ao superior, temos as oclusivas bilabiais [p,b] (Fig.1 A), quando a oclusão se dá entre o ápice da língua e a região dos alvéolos (logo atrás da arcada superior), falamos em oclusivas alveolares [t,d] (Fig. 1 B) e, finalmente, quando a oclusão ocorre envolvendo o dorso da língua e o véu palatino, dizemos que as oclusivas são velares [k,g] (Fig. 1 C). Em cada par já mencionado, a primeira consoante é surda e a segunda é sonora.

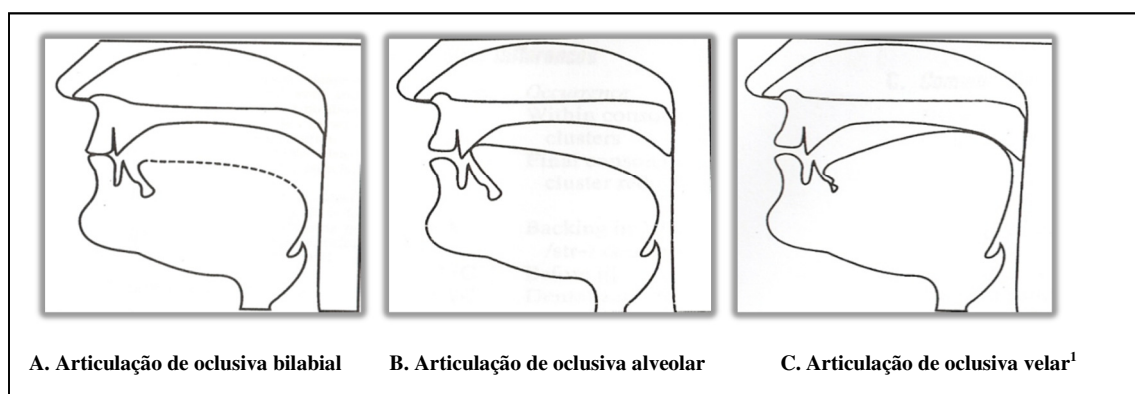


Figura 1: Representação da configuração do trato vocal na articulação das oclusivas bilabial, alveolar e velar. Fonte: Edwards, 1992.

Embora a língua inglesa conte ainda com a oclusiva glotal [ʔ], sua realização se dá em contextos muito específicos. Ela aparece, por exemplo, na expressão de negação normalmente

transcrita ortograficamente como *uh-uh* e foneticamente transcrita como [ʔʌʔʌ]. Surge ainda, em alguns dialetos, como alofone do [t] seguida de uma nasal silábica em palavras como *beaten* [biʔn] e *kitten* [kiʔn], ou como alofone de oclusivas surdas finais em palavras como *rap*, *rat* e *rack* (LADEFOGED, 2001). Por ter sua realização de tal forma restrita, tomamos a decisão de não contemplar a oclusiva glotal nesta pesquisa.

Para estudar as consoantes oclusivas e quaisquer outros sons presentes na fala, a Fonética Acústica atualmente lança mão de programas que permitem analisar o sinal sonoro com grande detalhe. Neles, podemos visualizar um dado acústico através de duas janelas: o oscilograma, que dá conta da forma de onda do sinal sonoro, e o espectrograma, que fornece as frequências de cada som de uma cadeia sonora qualquer em função do tempo. Atualmente, o *software* de processamento de áudio mais difundido é o *Praat* (BOERSMA; WEENINK, 2012).

Os sons da fala têm origem no processo de fonação. Esta acontece no momento da expiração, quando o ar egresso dos pulmões se propaga e encontra a glote, onde estão as pregas vocais. A pressão do ar nesse ponto faz com que as pregas vibrem resultando no chamado tom laríngeo, que dá origem aos sons da fala através das diversas configurações que o trato assume. A frequência de vibração das pregas vocais é medida em Hertz e a chamamos de *frequência fundamental* ou F0. Seu correlato perceptual é o *pitch*, ou seja, a altura em que um som é realizado. Assim, quanto mais grave é um som, menor é a sua frequência fundamental e, inversamente, quanto mais agudo, maior a sua frequência.

Quando o tom laríngeo sofre a ação dos articuladores do trato, que são todas as estruturas envolvidas na produção da fala, desde a laringe até os lábios, obtemos os sons característicos da fala humana. Ao produzimos consoantes, a constrição no trato é mais significativa do que na produção das vogais, embora possa ocorrer em maior ou menor grau, a depender do segmento produzido. Além disso, tal constrição ocorre em diversos pontos localizados por toda a extensão do trato (pontos de articulação). Por outro lado, as vogais são produzidas com constrição mínima do trato, e sua ocorrência limita-se à região que vai do palato até o véu palatino.

No que tange à análise acústica, quando estamos diante de um som vocálico, encontramos várias faixas de frequência ressaltada, as quais denominamos *formantes*⁹. Em um espectrograma, é possível vê-los ao longo de um segmento vocálico, em uma tonalidade mais escura. As vogais são, por excelência, os segmentos nos quais divisamos formantes, mas

⁹ Segundo Maia (1985) um formante é “uma faixa de frequência onde ocorre uma elevação da amplitude dos componentes espectrais da voz em função da resposta do ressoador (isto é, do próprio trato)”.

também podemos encontrá-los nas consoantes nasais, líquidas e aproximantes. A relação entre os dois primeiros formantes (F1 e F2) é normalmente suficiente para a identificação e análise de um segmento vocálico. Na transição entre uma vogal e qualquer outro segmento, os formantes apresentam mudanças características em sua trajetória.

As consoantes oclusivas, por sua vez, podem ser identificadas por meio de três eventos acústicos: a oclusão, a soltura e a transição formântica entre a oclusiva e a vogal seguinte. Durante a fase de oclusão, que, na língua inglesa, tem duração média de 50 a 100 milissegundos (KENT; READ, 1992), e no PB dura entre 73 e 142ms (MELO et al., 2011), há pouca ou nenhuma energia acústica, já que há bloqueio total do fluxo de ar no interior do trato vocal. Nos segmentos sonoros, pode haver, especialmente na produção de falantes de PB, a presença de uma barra de sonoridade horizontal ao pé do espectrograma durante a oclusão, o que indica a existência de vibração das pregas naquele intervalo.

Na Figura 2, temos a seleção da palavra *gap* [gæp] produzida por um informante brasileiro aprendiz de inglês. Entre as linhas pontilhadas destaca-se o momento da oclusão e, na porção inferior do espectrograma, nota-se uma barra cinza escura que caracteriza o vozeamento presente nas oclusivas sonoras. O vozeamento significa que, no momento em que o ar egresso dos pulmões passa pela glote, existe vibração das pregas vocais.

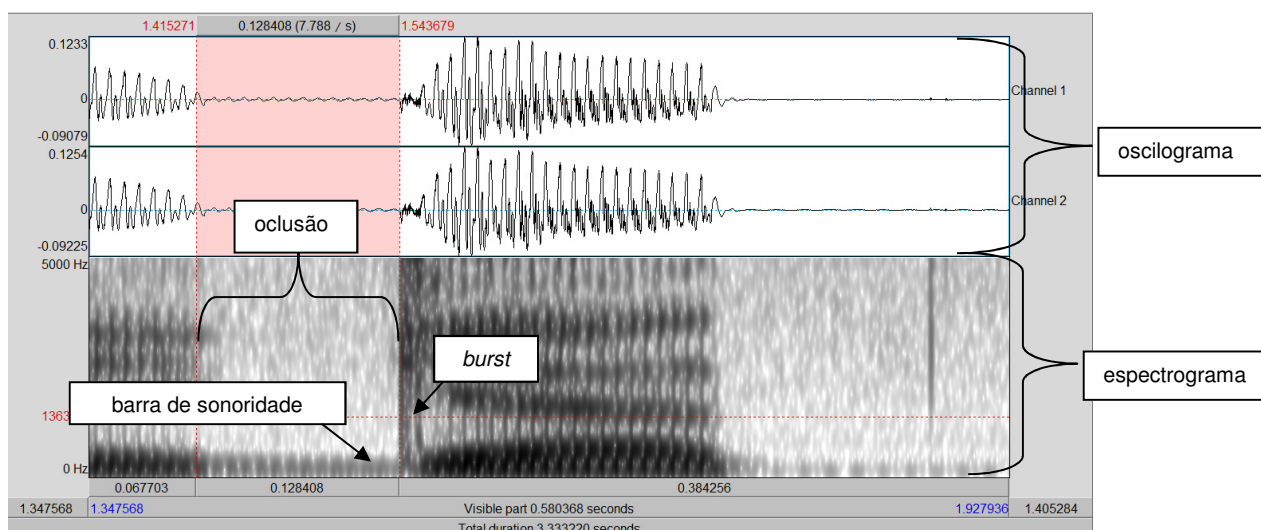


Figura 2: Oscilograma e espectrograma da palavra *gap*. Entre as linhas pontilhadas destaca-se o momento da oclusão. Fonte: *corpus* da autora

Na fase de soltura, é identificado o chamado *burst*, que dura entre 5ms e 40ms e está entre os mais breves eventos articulatórios da fala (KENT; READ, 1992). Caracteriza-se pelo momento em que os articuladores se afastam, liberando o fluxo de ar cuja propagação pelo trato estava impedida. É identificado no oscilograma como uma forma de onda aperiódica e,

no espectrograma, por meio de uma coluna vertical mais escura. Ocasionalmente, é possível divisar dois ou mais *bursts* seguidos, os chamados *bursts* múltiplos. Segundo Bonatto (2007, p. 54) a presença de *bursts* múltiplos pode ser devida “a características físicas de superfície de contato do palato ou uma alta pressão supraglótica”.

Na língua inglesa, o *burst* pode ou não ser seguido de aspiração, que, por sua vez, é um evento acústico no qual o ar passa pelas pregas vocais e através da laringe gerando um ruído semelhante a um sopro. Para tanto, as pregas devem estar maximamente próximas e tensionadas, embora não cheguem a se encostar em momento algum. O IPA (*International Phonetic Alphabet*) prevê um sinal diacrítico próprio para representar a aspiração, com vemos na transcrição da palavra *peak* [p^hik]. As oclusivas surdas prevocálicas do inglês são normalmente aspiradas em posição tônica. Entretanto, a aspiração não ocorre em alguns contextos: quando a oclusiva é antecedida de [s] na mesma sílaba, como na palavra *school* [skul], quando é sucedida por uma lateral, como em *plan* [plæn], ou um rótico, como em *cream* [krim]¹⁰. As oclusivas sonoras, por outro lado, são normalmente não aspiradas, tanto no inglês quanto no PB, independentemente do contexto em que ocorram.

À exceção dos contextos acima listados, portanto, é possível afirmar que, além do vozeamento, a aspiração também é uma pista para a distinção surda/sonora nas consoantes oclusivas prevocálicas do inglês. No PB, entretanto, a aspiração ocorre em níveis consideravelmente mais baixos, o que faz com que ela não seja normalmente considerada uma pista acústica na distinção entre oclusivas surdas e sonoras nessa língua.

A Figura 3 mostra a realização da palavra *kit* [kɪt] por uma informante norte-americana. Entre as linhas pontilhadas é possível divisar a duração total da consoante oclusiva [k]. Inicialmente, divisamos o momento da oclusão; em seguida, há a barra vertical de tonalidade mais escura a que chamamos *burst*; mais adiante, há um intervalo entre o *burst* e o início dos pulsos periódicos que caracterizam a produção da vogal. Tal intervalo é a aspiração. Vale notar também que, por estarmos observando a produção de uma oclusiva surda, não encontramos na porção inferior do espectrograma a barra de sonoridade durante a oclusão.

¹⁰ Fazemos aqui uma transcrição fonética larga para os dados arrolados como exemplo.

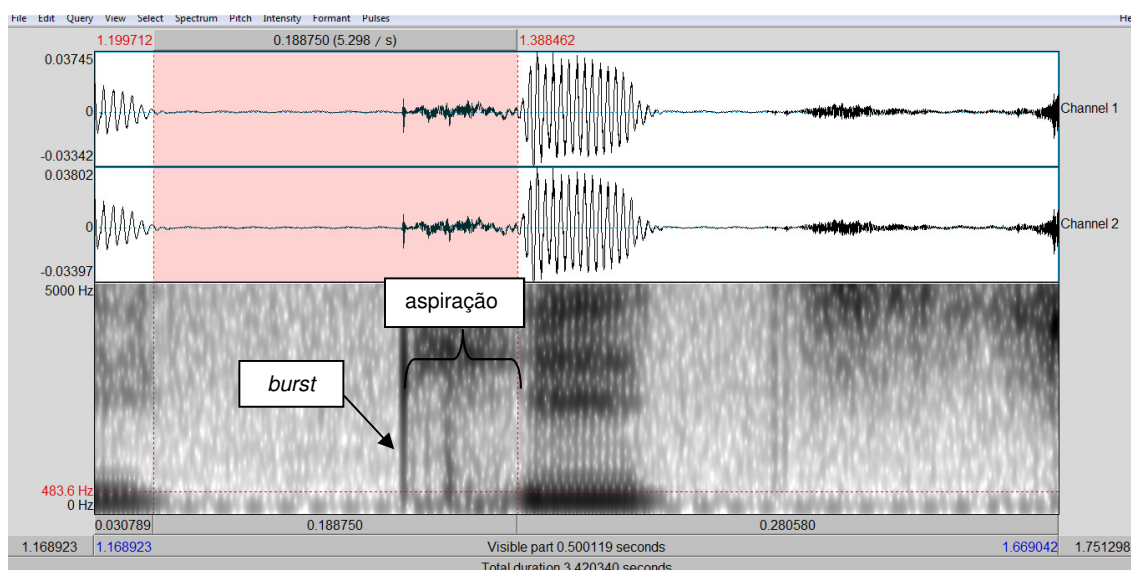


Figura 3: Oscilograma e espectrograma da palavra *kit*. No intervalo entre as linhas pontilhadas temos a consoante [k]. O *burst* é a barra vertical mais escura no espectrograma. Fonte: *corpus* da autora

Por fim, o ponto de articulação de uma oclusiva pode ser identificado pela trajetória que os formantes exibem em sua transição para a vogal seguinte. Kent e Read (1992) apontam que a mudança mais inequívoca na transição formântica entre uma oclusiva e a vogal seguinte se dá com relação ao primeiro formante (F1). Isso porque oclusivas apresentam o mais alto nível de constricção entre as consoantes, o que se reflete em um F1 de frequência muito baixa. Logo, a observação de um aumento no F1 na transição da oclusiva para a vogal seguinte gera uma pista sobre o modo de articulação da consoante.

Já a mudança da trajetória de F2 está relacionada ao ponto de articulação. No contexto em que a vogal seguinte é, por exemplo, posterior baixa como [ɑ], temos que os formantes traçam uma trajetória diferente para cada ponto de articulação: haverá um aumento sutil na frequência de F2 quando a oclusiva em questão for bilabial, uma diminuição acentuada quando for alveolar e uma diminuição sutil quando for velar. Entretanto, esse comportamento na trajetória dos formantes não se mantém da mesma forma para as todas as vogais do inglês¹¹ e, portanto, não é pista suficiente para determinar a identidade da oclusiva (KENT; READ, 1992).

Observemos a mudança na trajetória dos formantes nos dados desta pesquisa quando o segmento seguinte é [ɔ], uma vogal posterior meio-baixa, portanto, semelhante à posterior baixa [ɑ]. As palavras selecionadas nas Figuras 4, 5 e 6 são *bop* [bɒp], *dop* [dɒp] e *gosh* [gɔʃ], produzidas por um informante brasileiro aprendiz de inglês. Nos espectrogramas,

¹¹ O inglês britânico padrão (*Received Pronunciation*) conta com até 20 vogais, ao passo que a variedade norte-americana padrão (*General English*) possui 16 vogais (JENKINS, 2000).

identificamos os formantes pelas linhas horizontais mais escuras que se estendem por toda duração da vogal e estão destacadas por uma série de pequenos pontos. O primeiro formante (F1) está na porção mais inferior do espectrograma e nota-se que a sua trajetória é levemente ascendente nos três casos (Figs. 4, 5 e 6). Quanto ao F2, cujo início da trajetória está no cruzamento entre as linhas pontilhadas, vemos que para a bilabial sua trajetória é levemente ascendente; para a alveolar, é marcadamente descendente; e para a velar, levemente descendente. Logo, a mudança na trajetória dos formantes é semelhante àquela encontrada quando a vogal em questão é [ɑ]. Ressaltamos, entretanto, que esse padrão de trajetória formântica não será o mesmo diante de uma vogal anterior alta, por exemplo.

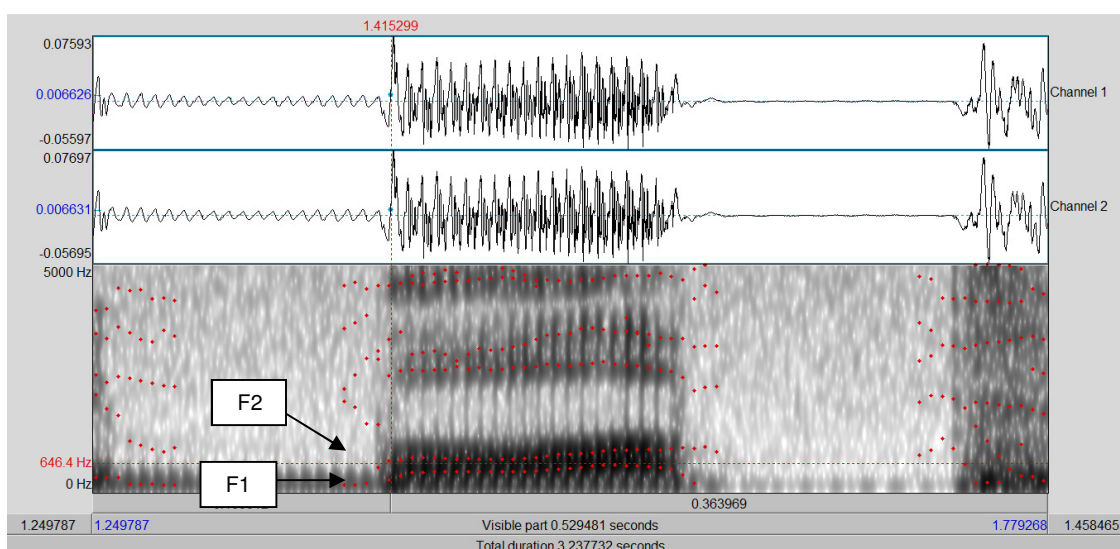


Figura 4: Oscilograma e espectrograma da palavra *bop*. Note a trajetória levemente ascendente de F2. Fonte: *corpus* da autora

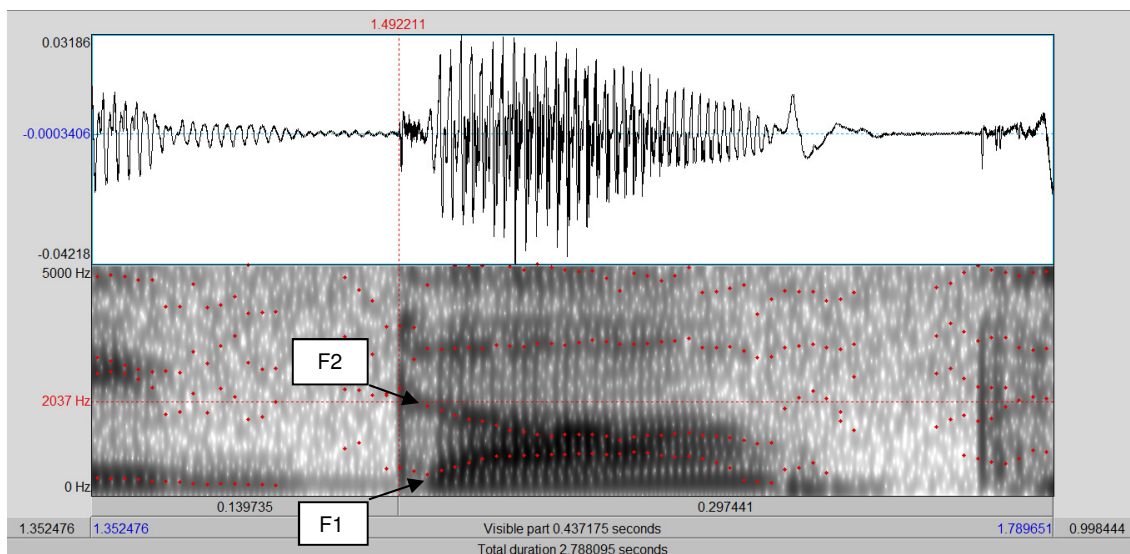


Figura 5: Oscilograma e espectrograma da palavra *dop*. Note a trajetória marcadamente descendente de F2. Fonte: *corpus* da autora

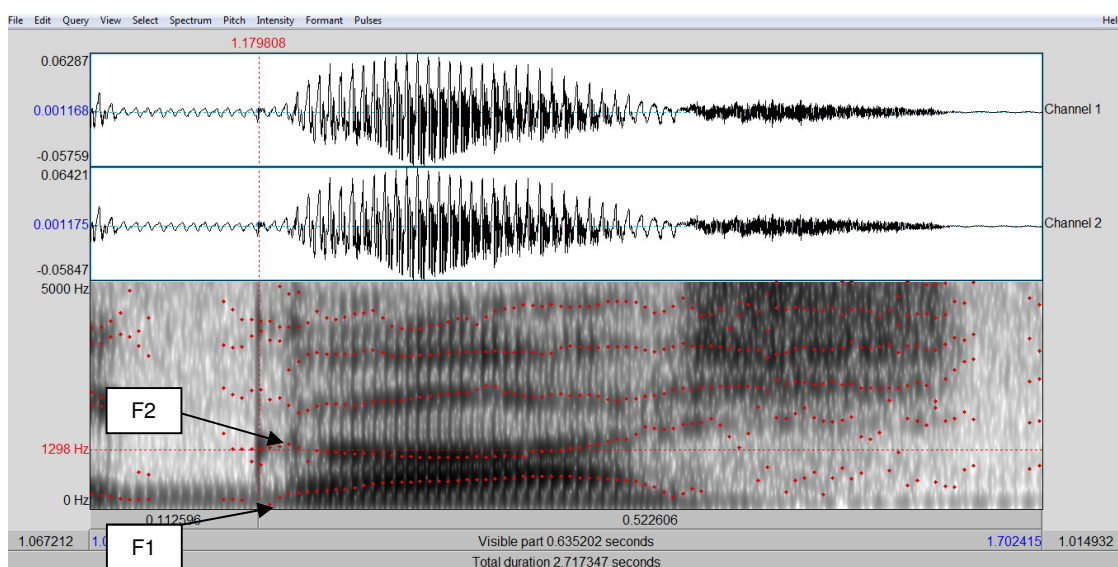


Figura 6: Oscilograma e espectrograma da *gosh*. Note a trajetória ligeiramente descendente de F2. Fonte: *corpus* da autora

Apesar disso, a literatura aponta que é possível encontrar um comportamento regular no que diz respeito à frequência inicial de F2 na transição da oclusiva para a vogal seguinte (KENT; READ, 1992). Daí resulta a teoria de que há um *locus* acústico para cada ponto de articulação, ou seja, a frequência inicial de F2 será mais ou menos a mesma para cada ponto de articulação. O valor inicial de F2 para as bilabiais é de aproximadamente 800 Hz, para alveolares é de 1800 Hz e as velares podem ter F2 inicial com frequências de 3000 Hz ou 1300 Hz. Revendo as Figuras 4, 5 e 6, encontraremos, em cada uma delas, o início da trajetória de F2 indicado pela seta. À esquerda, em vermelho, observa-se o valor da frequência

inicial de F2. Para a bilabial (Fig. 4), temos 646 Hz; para alveolar (Fig. 5), 2037 Hz; e para a velar (Fig. 6), 1298 Hz.¹² Destacamos que, embora tais valores sejam condizentes com a teoria do *locus* acústico, nem sempre encontramos correspondências tão exatas em nossos dados, o que faz com que seja aconselhável a associação de mais de um parâmetro para identificar o ponto de articulação das oclusivas.

2.3 TEMPO DE INÍCIO DO VOZEAMENTO (VOT)

A análise acústica das oclusivas pode ser feita por meio de diferentes pistas acústicas. Podemos, por exemplo, distinguir uma oclusiva surda de uma sonora atentando para a barra de sonoridade, que é identificada na porção inferior do espectrograma durante a fase da oclusão (KENT; READ, 1992). É possível ainda observar pistas como a duração da consoante, que tende a ser maior para oclusivas surdas (VELOSO, 1995), bem como a amplitude do *burst*, que é levemente mais intensa para os segmentos surdos (LADEFOGED; MADDIESON, 1996; BONATTO, 2007). Além disso, sabe-se que, quando uma vogal antecede uma oclusiva sonora, sua duração é normalmente mais longa do que se estivesse diante de uma oclusiva surda (CHEN, 1970¹³, apud BONATTO, 2007). É comum que várias dessas pistas sejam associadas para contrastar oclusivas surdas e sonoras em uma dada língua.

É possível ainda caracterizar as oclusivas por meio de um parâmetro acústico chamado VOT (*Voice Onset Time*), que é o intervalo existente entre a soltura dos articuladores e o início do vozeamento do segmento seguinte. Sua primeira descrição encontra-se em um trabalho de Lisker e Abramson (1964) no qual os autores buscaram identificar um parâmetro único para distinguir oclusivas surdas e sonoras. Para fins de medição do VOT, o momento da soltura é determinado como ponto zero. Se o vozeamento é anterior à soltura, o VOT tem valores negativos e estamos diante de uma típica oclusiva sonora; se o vozeamento é posterior, o VOT é positivo e a oclusiva é tipicamente surda. Há casos em que o vozeamento coincide com a soltura, então dizemos que o VOT tem valor zero. No PB, é comum que o vozeamento das sonoras se inicie antes da soltura e, portanto, tenha valores negativos (ISTRE,

¹² Saliente-se que estamos aqui recorrendo a apenas um dado para cada ponto de articulação, e haverá diferentes realizações para tais segmentos por outros sujeitos e mesmo por um único sujeito.

¹³ CHEN, M. *Vowel length variations as a function of the voicing of the consonant environment*. In: *Phonetica*, v. 22, p. 129-159, 1970.

1983¹⁴ apud KLEIN, 1999; KLEIN, 1999; MELO et al., 2011); ao passo que no inglês, o vozeamento de uma sonora em início de palavra, muitas vezes, ocorre no mesmo momento da soltura (LISKER; ABRAMSON, 1964; MACKEN; BARTON, 1978) o que caracteriza VOT = zero. Nesses casos, não é possível divisar a barra de sonoridade imediatamente anterior ao momento da soltura. Logo, a presença da barra de sonoridade durante a oclusão não é, isoladamente, um parâmetro aconselhável para distinguir surdas e sonoras, visto que, em certos casos, ela só irá ocorrer concomitantemente ou após a soltura.

Esse é um dos motivos pelos quais o VOT torna-se um parâmetro mais apropriado para esta pesquisa, inclusive no que tange à análise das oclusivas surdas. Isso porque, no contexto aqui pesquisado, o contraste entre o inglês de um nativo norte-americano e um falante de PB se dá basicamente pelos níveis de aspiração, que costumam ser consideravelmente mais longos no inglês de um nativo (ALVES, 2011; FRANÇA, 2011). Ora, sabendo-se que a aspiração é o correlato acústico do VOT nas oclusivas surdas, temos que a medição do VOT de uma oclusiva surda culmina na própria medição de seu nível de aspiração. Portanto, ao optarmos pela análise dos dados por meio da medição do VOT, utilizamos um parâmetro que nos permite, a um só tempo, analisar os índices de aspiração presentes nas oclusivas surdas e os níveis de sonoridade que caracterizam as sonoras.

Uma vez esclarecida a nossa opção, é necessário que o leitor se familiarize um pouco mais com o conceito de VOT. Lisker e Abramson (1964) propuseram o foco na relação de tempo entre o início do vozeamento e a soltura dos articuladores da oclusiva a fim de obter um parâmetro único para a análise do contraste surda/sonora nas oclusivas em início de palavra.

No estudo em questão, os pesquisadores investigaram 11 línguas (inglês norte-americano, cantonês, holandês, húngaro, espanhol porto-riquenho, tâmil, coreano, armênio oriental, tailandês, hindi e marati) por meio da análise acústica de dados de um total de 17 informantes, que produziram um grupo de palavras nas quais havia todas as oclusivas licenciadas para ocorrerem em posição inicial de palavra em sua língua materna. Além disso, os sujeitos foram interpelados a construir duas sentenças: uma na qual a palavra-alvo se encontrasse em posição inicial na frase e a outra na qual não se encontrasse. Após a análise dos dados, os autores observaram que as 11 línguas por eles investigadas apresentavam oclusivas iniciais prevocálicas em três situações: (a) consoantes cujo vozeamento inicia-se antes da soltura dos articuladores (pré-vozeado ou VOT negativo); (b) oclusivas cujo

¹⁴ ISTRE, G. L. (s/d) *Um estudo do VOT em monolíngues brasileiros*. (trabalho não publicado).

vozeamento ocorre concomitantemente à soltura ou logo após (VOT zero ou intervalo curto); e (c) oclusivas cujo vozeamento se inicia com certo retardo em relação à soltura (VOT positivo ou intervalo longo). A presença do intervalo longo coincide com o evento da aspiração, uma vez o retardo no vozeamento da vogal seguinte gera a produção do ruído em questão, por conta da passagem do ar através das pregas vocais parcialmente abertas. Assim sendo, Lisker e Abramson (1964) concluíram que as 11 línguas analisadas se encaixavam em três grupos a depender do número de categorias de oclusivas: as línguas de duas categorias (inglês norte-americano, cantonês, holandês, húngaro, espanhol porto-riquenho e tâmil), as línguas de três categorias (coreano, armênio oriental e tailandês) e as de quatro categorias (hindi e marati). Vale lembrar que embora inglês e espanhol sejam ambas línguas de duas categorias, a forma como o VOT se apresenta em cada uma delas é diferente. No inglês norte-americano, as oclusivas sonoras são caracterizadas com VOT de intervalo curto e as surdas apresentam intervalo longo, ao passo que no espanhol porto-riquenho há a presença de VOT negativo para as sonoras e intervalo curto para as surdas. O português é também uma língua de duas categorias, cujo comportamento se assemelha ao do espanhol.

2.3.1 A RELAÇÃO ENTRE VOT E O PONTO DE ARTICULAÇÃO

Outro trabalho que é referência para o estudo de VOT foi conduzido por Cho e Ladefoged (1999). O objetivo principal dos pesquisadores foi analisar a relação entre duração de VOT e o ponto de articulação, sendo seu foco apenas as oclusivas surdas aspiradas e não aspiradas, em posição inicial e antes de vogais não altas. Eles optaram por utilizar palavras isoladas, uma vez que contextos diferentes poderiam afetar o VOT.

O estudo de Cho e Ladefoged (1999) envolveu a análise de 18 línguas em vias de extinção. Eles notaram que, com exceção da língua Dahalo, falada no Quênia, o ponto velar é o que detém o VOT mais longo em todas as 13 línguas que não contrastam velares e uvulares. Logo, o padrão observado em estudos anteriores que abordavam apenas uma língua, repetiu-se na análise de número maior de línguas: pontos de articulação mais posteriores no trato tendem a exibir VOTs mais longos. O mesmo foi observado em estudos feitos no Brasil por Klein (1999), Bonatto (2007) Melo et al. (2011). É preciso notar, entretanto, que no estudo de Cho e Ladefoged (1999), a diferença de VOT entre bilabiais e alveolares não se provou significativa nas 13 línguas que contavam com ambos os pontos de articulação, resultado

também encontrado em outras pesquisas que demonstraram diferença não significativa ou, até mesmo, sobreposição dos valores apresentados nos dois pontos (LISKER; ABRAMSON, 1964; ABRAMSON; LISKER, 1971).

A partir dos dados analisados, Cho e Ladefoged (1999) acharam plausível dividir as oclusivas surdas em 4 categorias: (a) oclusivas não aspiradas, com VOT por volta de 30ms, (b) levemente aspiradas, com VOT aproximado de 50ms, (c) aspiradas, com VOT em torno de 90ms e (d) altamente aspiradas, para línguas como o Tlingit e Navajo. Além disso, observaram que línguas com mais de dois tipos de oclusivas tendem a enfatizar esse contraste dispersando os valores de VOT através de um *continuum*. Entretanto, essa tendência não é necessariamente seguida por todas as línguas. Para esses autores, para serem consideradas aspiradas, as oclusivas devem possuir os seguintes valores como limiares: 55ms para [p], 70ms para [t] e 80ms para [k], o que condiz com os valores médios encontrados por Lisker e Abramson (1964) em palavras isoladas.

Cho e Ladefoged (1999) elencam uma série de características fisiológicas/aerodinâmicas que causam variação de VOT devido ao ponto de articulação. A primeira delas diz respeito ao volume da cavidade supraglotal atrás do ponto de constrição. Quanto mais posterior o ponto de constrição, menor será esse volume e, portanto, o ar egresso dos pulmões no instante da soltura exercerá maior pressão sobre o ponto de constrição. Como resultado, será necessário mais tempo para que tal pressão atinja o ponto adequado ao início da vibração das pregas. Considerando que o VOT é o intervalo de tempo entre a soltura dos articuladores e o início da vibração das pregas, temos que o VOT será mais longo para uma consoante velar, cujo ponto de constrição é mais posterior, do que para uma alveolar ou bilabial. Observando o mesmo fenômeno sob outro prisma, temos que quanto maior a massa de ar à frente do ponto de constrição, mais tempo será necessário para que tal massa seja movida e o ar egresso dos pulmões seja finalmente liberado.

Em seguida, Cho e Ladefoged (1999) apontam a interferência do movimento dos articuladores na variação do VOT. Para tal, observam que tanto os lábios quanto a ponta da língua, utilizados para a realização das oclusivas bilabial e alveolar, respectivamente, movem-se mais rapidamente do que o dorso da língua, recrutado para a realização da velar (HARDCASTLE, 1973¹⁵, apud CHO E LADEFOGED, 1999). Diante disso, a velar tomará mais tempo para ser articulada e, portanto, terá o início do vozeamento retardado em relação às demais, o que resultará em um VOT de maior duração. Além disso, o movimento da

¹⁵ HARDCASTLE, W.J. *Some observations on the tense-lax distinction in initial stops in Korean*. Journal of Phonetics, v.1, 1973.

mandíbula afeta de maneiras diferentes a velocidade com que os demais articuladores se movimentam: o lábio inferior tem seu movimento acelerado, ao passo que o dorso da língua é afetado de maneira menos intensa pelo movimento da mandíbula.

Outro fator que interfere na duração do VOT é a extensão da área de contato entre articuladores. De maneira geral, quanto maior a área de contato, mais longo o VOT. As velares, por exemplo, que notadamente possuem VOT de duração mais longa, são produzidas com constrição entre o dorso da língua e o palato mole, e possuem uma área de contato consideravelmente maior do que aquela existente para os demais pontos.

No que diz respeito às oclusivas aspiradas, a mudança da área de abertura glotal é mais um fator que causa variação de VOT. Isso porque a glote, anteriormente aberta para permitir a aspiração, precisa ir progressivamente se fechando depois da soltura para permitir a vibração das pregas. A área de abertura glotal tem redução mais lenta para as oclusivas velares porque a pressão intraoral é reduzida mais vagarosamente. Dessa forma, o início do vozeamento se dará posteriormente para velares em relação a alveolares e bilabiais.

Finalmente, há que se considerar o ajuste temporal entre a duração da oclusão e o VOT. Em pontos mais posteriores, a oclusão tende a ser mais curta justamente porque a aspiração será mais longa. Resulta daí uma relação temporal inversa entre oclusão e VOT, ou seja, quanto mais longa a oclusão, mais curta a duração da aspiração e mais rapidamente se dá o início do tempo de vozeamento.

2.3.2 OUTROS FATORES QUE INFLUENCIAM O VOT

Além do ponto de articulação, há outros fatores que influenciam a duração do VOT. Em fala contínua, por exemplo, a duração do VOT é mais curta para oclusivas surdas e mais longa para sonoras, se comparada à sua produção em palavras isoladas. Baran et al.¹⁶ (1977, apud DOCHERTY, 1992) apontam que a diferença entre os valores médios de VOT em consoantes oclusivas homorgânicas chega a 80ms em palavras isoladas, ao passo que em fala contínua esse valor cai para 30ms. Isso nos leva a constatar que é preciso cautela ao analisar os resultados das pesquisas que tratam de índices de VOT e, ao tratarmos de valores absolutos, é sempre prudente atentarmos para as condições nas quais os dados foram coletados. Docherty

¹⁶ BARAN et al. *Phonological contrastivity in conversation: a comparative study of voice onset time*, Journal of Phonetics, v. 5, 1977.

(1992) nos alerta para esse fato ao apresentar resultados de diferentes pesquisas realizadas com oclusivas. Um estudo realizado por Zlatin (1974) aponta valores médios de VOT para consoantes iniciais do inglês que variam de 28ms a 90ms, sendo que os valores absolutos variam entre 10ms e 220ms. Docherty (1992) alega que essa variação se deve a uma série de fatores, entre eles, o fato de que a maioria dos estudos realizados não controla a taxa de elocução dos sujeitos, apesar de essa variável interferir na duração do VOT. Outro fator que tem influência na grande variação de resultados é a própria forma de medição do VOT que, segundo o autor, pode se dar de diferentes maneiras. Ele cita Fischer-Jorgensen e Hutters (1981¹⁷, apud Docherty, 1999), que identificaram três possíveis pontos onde é possível marcar o início do vozeamento da vogal que precede a oclusiva: (a) o próprio início do vozeamento indicado pela periodicidade de baixa frequência na forma de onda, (b) o início do F1 e (c) o início de formantes mais altos.

A qualidade da vogal também parece ter influência na duração do VOT. Port e Rotunno (1979) observaram ser o VOT 7% mais longo diante das vogais altas [i] e [u] do que diante da vogal baixa [a]. Resultados semelhantes foram encontrados por Ohala (1981).

É preciso levar em conta ainda a tonicidade da sílaba na qual se encontra a oclusiva, uma vez que tem sido observado que a duração de VOT é maior em oclusivas que estão em sílabas tônicas (LISKER AND ABRAMSON, 1967; GIMSON, 1980). Klein (1999) e Alves et al. (2008) também verificaram que há interdependência entre tonicidade e VOT nas oclusivas do PB. Tal fato parece estar ligado aos correlatos acústicos de acento, uma vez que nessa língua, a duração da sílaba tônica supera a das demais sílabas, sendo o principal correlato físico do acento no PB. (MASSINI-CAGLIARI, 1992)

O número de sílabas também faz diferença. Klatt (1975, apud DOCHERTY, 1999) estima que o VOT é 8% menor em palavras dissílabas quando comparadas a monossílabas.

Por fim, a sonoridade da consoante final na sequência CVC também parece ter influência na duração do VOT, que, na língua inglesa, é mais longo quando a consoante final é sonora (PORT; ROTUNNO, 1979). Na presente pesquisa, não nos foi possível controlar essa variável, já que optamos por um *corpus* constituído exclusivamente por palavras reais da

¹⁷ FISCHER-JORGENSEN, E; HUTTERS, B. *Aspirated stop consonants before low vowels: a problem of delimitation*. Annual Report Inst. of Phonetics of Copenhagen, 1981.

língua inglesa ao invés de logatomas¹⁸, o que nos impediu de obter um equilíbrio entre palavras com consoante final surda ou sonora, ou a exclusiva utilização de uma ou outra.

2.3.3 ESTUDOS SOBRE A DURAÇÃO DO VOT EM DIFERENTES LÍNGUAS

Há estudos sobre a duração do VOT em consoantes oclusivas nas mais diferentes línguas. Na maior parte, eles traçam comparação com o inglês. Alguns deles estudam o VOT em início de palavra, outros em fim. Muitos deles apresentam tanto aspectos de produção quanto de percepção.

Caramazza et al. (1973) investigaram como se dá a produção e a percepção de oclusivas para falantes canadenses bilíngues de inglês e francês. Para fins de comparação, os pesquisadores testaram também falantes monolíngues de ambas as línguas. A produção de oclusivas iniciais foi avaliada por meio de análises acústicas e o *corpus* foi composto de 9 palavras para cada uma das 6 consoantes oclusivas. A percepção foi testada fazendo com que os sujeitos identificassem palavras que tiveram o VOT modificado em três pontos do *continuum*. Os sujeitos tinham idades entre 17 e 25 anos.

Os pesquisadores observaram que o VOT parece ser uma variável de distinção de vozeamento importante para o inglês, mas não necessariamente para o francês. Isso porque os dados dos falantes monolíngues de francês apontam sobreposição nos valores de VOT para os pares homorgânicos [pa/ba], [ta/da] e [ka/ga], o que não ocorre nos dados dos falantes monolíngues de inglês, nos quais há clara separação para cada contraste. Os falantes bilíngues, por sua vez, apresentaram disparidade na maneira como produziram as distinções de vozeamento a depender da língua. Quando se tratava dos dados produzidos em francês, os sujeitos apresentaram sobreposição de valores de VOT tal qual aquela apresentada pelos monolíngues de francês. Entretanto, quando se tratava do inglês, muito embora os valores de VOT tenham mudado para se aproximar daqueles produzidos pelos monolíngues de inglês, não chegaram a se igualar a eles. No que tange aos testes estatísticos, sua realização só foi possível para as oclusivas surdas, uma vez que os dados provenientes das sonoras apresentaram distribuição descontínua. Dessa forma, observou-se que, para as oclusivas surdas, as produções em francês dos bilíngues não apresentaram diferença estatisticamente significativa em relação às produções dos monolíngues [$F(1,84) = 3,56; P > 0,05$], ao passo que suas produções em inglês diferiam significativamente daquelas dos monolíngues [$F(1,84) =$

¹⁸ Logatomas são sequências que, embora não constituam itens lexicais de uma língua, são fonologicamente possíveis para o sistema linguístico em questão; por exemplo, “panata” no PB.

44,10; $P < 0,001$]. Os pesquisadores apontam que tal diferença possivelmente se deve ao fato de que todos os sujeitos bilíngues têm o inglês como segunda língua e que, portanto, a influência entre as línguas seja unilateral, apenas da primeira para a segunda.

Caramazza et al (1973) apontam ainda que os resultados obtidos para os monolíngues de inglês são consistentes com aqueles em Lisker e Abramson (1964) e Abramson e Lisker (1965, 1970) para o inglês americano, e sugerem que o VOT é fonêmico, tanto no inglês americano quanto no canadense.

No entanto, os autores sugerem que o mesmo não ocorre para o francês, uma vez que os resultados obtidos dos sujeitos monolíngues dessa língua apontam para o fato de que, além de vozeamento e aspiração, outras pistas podem ser mais relevantes para os falantes de francês canadense. Entre elas podem estar a força articulatória (fortis-lenis)¹⁹ e a taxa de transição formântica, que diz respeito às frequências que os formantes assumem na transição entre as oclusivas e as vogais que as sucedem. Os autores concluem, portanto, que as distribuições de VOT realizadas pelos falantes bilíngues em francês indicam a ausência de interferência fonológica da segunda língua (inglês) na primeira (francês). Entretanto, o contrário não é verdadeiro, já que a produção dos bilíngues mostrou diferenças significativas em relação à dos monolíngues de inglês. Os pesquisadores apontam ainda que a maior diferença na produção do francês e do inglês entre os sujeitos bilíngues concentrou-se no VOT positivo, ou seja, a capacidade de controlar o VOT para as oclusivas surdas mostrou-se mais eficiente, ao passo que o VOT negativo manteve-se praticamente inalterado para as sonoras nas duas línguas.

Ainda no campo do bilinguismo, Sancier e Fowler (1997) conduzem três experimentos para analisar a flutuação no sotaque de uma informante brasileira de 27 anos que estuda nos Estados Unidos e passa as temporadas de férias no Brasil, permanecendo, portanto, alguns meses em cada país. Essa estudante aprendeu a língua inglesa, por volta dos 15 anos, ou seja, após o chamado Período Crítico²⁰. A motivação para o estudo surgiu com o relato da estudante de que, quando retornava ao Brasil, seu pai estranhava seu sotaque e se referia a ele como “explosivo”. Tendo em mente as pesquisas de Flege (1987) e Major (1992) e considerando que o sotaque da estudante poderia se dever, entre outros fatores, aos níveis de aspiração das oclusivas surdas por ela realizados, as pesquisadoras resolveram medir os níveis de VOT em sua produção. Assim, ela foi gravada em três momentos diferentes: após 4.5

¹⁹ Força articulatória diz respeito à quão energeticamente um segmento é produzido com relação à energia muscular, à tensão dos órgãos fonadores ou à energia total mobilizada na sua produção.

²⁰ A Hipótese do Período Crítico preconiza que há um declínio na capacidade de aprender línguas que está relacionado à maturação do indivíduo.

meses nos Estados Unidos, ou seja, um pouco antes de sua partida para o Brasil; depois de 2.5 meses no Brasil; e, novamente, um pouco antes de retornar ao Brasil, depois de 4 meses nos Estados Unidos. O *corpus* desse estudo contou com 6 sentenças que a informante deveria traduzir do português para o inglês e vice-versa, em 5 repetições aleatorizadas. O objetivo era observar de que maneira a produção da estudante era afetada pela língua falada a seu redor. Para isso, foram medidos os níveis de VOT nas três diferentes ocasiões acima citadas e sua produção foi avaliada por falantes nativos de ambas as línguas com relação ao grau de sotaque percebido.

No primeiro experimento, 13 falantes nativos de PB julgaram que as sentenças que a estudante produziu logo após sua estada nos Estados Unidos eram mais marcadas por sotaque do que aquelas produzidas no fim de sua estada no Brasil. No segundo experimento, 33 falantes nativos de inglês americano julgaram que as sentenças produzidas logo após sua estada no Brasil apresentavam mais sotaque do que as produzidas depois de meses nos Estados Unidos. Entretanto, os acertos aqui não ocorreram de maneira tão acurada e os pesquisadores acreditam que isso se deve ao fato de que os americanos tinham de decidir entre dois sotaques estrangeiros em maior ou menor grau, ao passo que os brasileiros tinham a tarefa facilitada ao julgar entre um sotaque nativo e um estrangeiro.

O terceiro experimento concentrou-se na medição de VOT das oclusivas [p] e [t] produzidas pela informante, a fim de detectar se a língua ambiente ditava mudanças na produção de ambas as línguas. Para tanto, foram selecionadas sentenças que continham as consoantes [p] e [t] no português, bem como [p^h] e [t^h] no inglês. Foram analisados 30 *tokens* de cada um dos segmentos e, na maior parte das vezes, as oclusivas encontravam-se em posição inicial ou tônica.

Os resultados demonstraram que a língua falada e a experiência recente em um dos ambientes (Brasil ou Estados Unidos) tiveram efeito significativo nos índices de VOT produzidos pela informante. Quanto à língua falada, o VOT para [p^h] no inglês mostrou-se, em média, 22 milissegundos mais longo do que o VOT para [p] nas produções em português, ao passo que [t^h] foi 33ms mais longo que [t]. Quanto à experiência recente, as pesquisadoras constataram um aumento de, em média, 6ms no VOT de [p] e 5ms no VOT de [t] após meses nos Estados Unidos. Isso levou Sancier e Fowler (1997) a concluir que sua informante apresentou a chamada flutuação gestual, ou seja, uma mudança na coordenação dos gestos de desvozeamento laríngeo e constrição oral, responsáveis pela mudança de VOT encontrada nos dados. De acordo com as autoras, essa flutuação pode se dever a uma predisposição do falante/ouvinte em imitar aquilo que percebe na fala do outro, o que reforçaria uma teoria

gestual como a do realismo direto²¹ ou a teoria motora²². Além disso, as autoras acreditam que é preciso lançar mão da correspondência entre os fones da L1 e L2 defendida por Flege (1987), para que se possa explicar o paralelismo existente na flutuação de VOT na produção das duas línguas. Assim, seria possível compreender que a falante nativa de PB formou em sua língua uma categoria de /t/ que a permite uma realização autêntica de [t] no PB e uma realização com sotaque de [t^h] no inglês, sendo que possuem uma ligação que faz com que ambas se modifiquem de acordo com a língua ambiente.

Mas se a produção da informante de Sancier e Fowler (1997) está sujeita à flutuação gestual, o que dizer de indivíduos bilíngues que adquiriram as duas línguas desde a infância? Haverá o mesmo tipo de interação entre L1 e L2 observada em um sujeito que adquire a língua após o chamado Período Crítico? Rocca (2002) se propõe a investigar essa questão e, para isso, analisa o grau de precisão com que falantes bilíngues de infância reproduzem o VOT de oclusivas surdas em sílaba tônica, no início de palavras no PB e no inglês. Ela analisou a produção de três sujeitos bilíngues com idades entre 28 e 48 anos, que adquiriram PB e inglês antes dos 3 anos de idade. O *corpus* continha 6 frases-veículo cujas palavras-alvo foram *papa*, *tapa* e *capa* para o PB e *pop*, *top* e *cop* para o inglês. Foram realizadas 12 repetições. Os resultados foram comparados com as medidas de VOT obtidas por Lisker e Abramson (1964) para o inglês e a referência para o PB veio de Behlau (1986²³, apud Rocca, 2002).

Os resultados apontam para níveis de controle de VOT que se distinguem daqueles apresentados por falantes monolíngues tanto de PB quanto de inglês. Embora os sujeitos façam distinção clara entre o inglês e o PB com relação aos índices de VOT das oclusivas, não é possível afirmar que tenham adquirido as duas línguas com as mesmas características dos falantes monolíngues. Rocca (2002) conclui que é muito difícil precisar o que está em jogo quando tentamos compreender a aquisição de bilíngues de infância. Para tanto, ela evoca a argumentação de Flege (1998), segundo a qual podemos levar em conta dois possíveis fatores: a maturação dos centros de linguagem no cérebro quando da exposição às duas línguas ou a interação entre as duas línguas e o estágio de desenvolvimento em que um sistema linguístico se encontra quando da introdução do segundo. Rocca aponta que o

²¹ Para informações mais detalhadas, ver FOWLER, C.A. *An event approach to the study of speech perception from a direct-realist perspective*. In: *Journal of Phonetics*, v.14, p. 3-28, 1986 e NISHIDA, G., *Sobre teorias de percepção da fala* (Tese Doutorado, UFPR, 2012).

²² Para informações mais detalhadas, ver LIBERMAN, A.M.; MATTINGLY, I.G. *The motor theory of speech perception revised*, In: *Cognition*, v. 21, p. 1-36, 1985 e NISHIDA, G., *Sobre teorias de percepção da fala* (Tese Doutorado, UFPR, 2012).

²³ BEHLAU, M. S. *Análise do Tempo de Início de Sonorização na Discriminação dos Sons Plosivos do Português*. Tese de Doutorado. Escola Paulista de Medicina. São Paulo, 1986.

problema, nesses casos, é que a aquisição de duas línguas e o desenvolvimento do sistema neural ocorrem em um mesmo momento e é, portanto, difícil isolar uma dessas variáveis para estudo. De toda forma, esses resultados corroboram as pesquisas de Sancier e Fowler (1992) e a hipótese de que a interferência fonética entre as línguas é bi-direcional.

Além dos estudos acima apresentados, há diversos outros que comparam o inglês a outras línguas. Podemos citar no espanhol, Flege e Eefting (1987, 1988), Bohn e Flege (1993) e Toribio et al (2005); no francês, Saerens et al (1987) e Flege e Hillenbrand (1984); no russo, Ringen e Kulikov (2010); no polonês, Keating et al. (1981); no árabe, Flege e Port (1981) e Khattab (2002); no turco, Ögüt et al (2006); no húngaro, Gósy (2001); e no mandarim, Chen et al (2007).

2. 4 A PRODUÇÃO DE CONSOANTES OCLUSIVAS NO INGLÊS E NO PB

Já munidos do conceito de VOT e cientes da existência de trabalhos que se utilizam desse parâmetro para analisar consoantes oclusivas de diversas línguas, passaremos agora a observar mais detidamente trabalhos que focam a produção dessas consoantes nas línguas que interessam mais especificamente à presente pesquisa: o inglês, por ser a língua aqui investigada, e o português, por ser a língua materna dos sujeitos cuja produção comparamos a dos nativos de inglês. Inicialmente, faremos um levantamento de trabalhos que se concentram nas consoantes oclusivas surdas no inglês, apresentando os valores médios de VOT encontrados em cada uma delas. Em seguida, listaremos os trabalhos que se ocupam das oclusivas surdas no PB. O mesmo procedimento se repetirá para as oclusivas sonoras e, por fim, trataremos de elencar trabalhos que focam a produção de oclusivas na aquisição de inglês por falantes nativos de PB.

2.4.1 OCLUSIVAS SURDAS NO INGLÊS

Na seção 2.2, vimos que Lisker e Abramson (1964) conduziram um trabalho pioneiro a fim de testar a utilização do VOT como parâmetro acústico único na distinção entre

consoantes oclusivas surdas e sonoras. Dentre as 11 línguas pesquisadas, encontrava-se o inglês norte-americano que, segundo os pesquisadores, foi caracterizado como uma língua de duas categorias, ou seja, observou-se a existência de oclusivas iniciais prevocálicas cuja duração de VOT (a) aproximava-se a zero e (b) possuía intervalo longo. Os 4 informantes produziram palavras em que constavam todas as oclusivas iniciais prevocálicas do inglês licenciadas para ocorrerem nessa posição [p, b, t, d, k, g]. De início, as palavras foram produzidas isoladamente e, em um segundo momento, os informantes foram interpelados a construir duas sentenças: uma na qual a palavra-alvo se encontrasse em posição inicial de frase e a outra em posição medial. Lisker e Abramson (1964) apuraram os seguintes resultados com relação às oclusivas surdas: os valores médios de VOT para [p], [t] e [k] foram, respectivamente: 58ms, 70ms e 80ms em palavras isoladas; 28ms, 39ms e 43ms em posição inicial na sentença; e 34ms, 37ms e 49ms em posição medial. Segue um resumo dos dados na Tabela 1.

Inglês norte-americano

	[p]	[t]	[k]
Palavras isoladas	58	70	80
Posição inicial	28	39	43
Posição medial	34	37	49

Tabela 1: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do inglês norte-americano no estudo de Lisker e Abramson (1964)

Observa-se, portanto, que os valores de VOT em palavras isoladas superam consideravelmente aqueles em posição inicial ou medial. Resultados semelhantes foram encontrados em pesquisas posteriores como as já mencionadas de Baran et al. (1977, apud Docherty, 1999) e Zlatin (1974).

Klatt (1975²⁴, apud Docherty, 1999), por sua vez, analisou dados de 3 adultos do sexo masculino falantes nativos de inglês, que produziram oclusivas em posição inicial inseridas na frase-veículo “Say_____ *instead.*” Para as surdas, ele obteve os seguintes valores médios de VOT: 47ms para [p], 65ms para [t] e 70ms para [k]. Nota-se que a duração de VOT aqui é superior àquela que Lisker e Abramson (1964) obtiveram quando solicitaram a seus sujeitos que construíssem sentenças nas quais a palavra-alvo estivesse inclusa. Diante disso, é possível notar que, na realização de frases-veículo, o VOT das oclusivas tende a ser mais longo do que quando essas frases são construídas espontaneamente pelos sujeitos. Isso reforça a

²⁴ KLATT, D. *Voice onset time, friction and aspiration in word initial consonant.* In: Journal of Speech and Hearing Research, v. 18, 1975.

necessidade da utilização de valores de duração relativa a fim de que tenhamos maior acurácia na análise dos dados que envolvam a medição de VOT.

Macken e Barton (1978), ao estudar a aquisição do contraste de vozeamento por crianças monolíngues de inglês, utilizam como referência os valores de VOT encontrados na fala de adultos em três diferentes estudos (LISKER; ABRAMSON, 1964; KLATT, 1975²⁵; ZLATIN, 1974) e reportam os seguintes valores como variações normais nos valores de VOT: de 47 a 81ms para [p], 67 a 87ms para [t] e 70 a 90ms para [k].

Já Major (1992) realizou um estudo para investigar a interferência da L2 na L1. Ele analisou os índices de VOT em dados de 5 sujeitos bilíngues (todos do sexo feminino), falantes nativas de inglês americano, residentes no Brasil por períodos que variavam entre 12 a 35 anos. Como grupo controle, o pesquisador obteve dados de 5 falantes nativos de inglês residentes nos Estados Unidos e 5 falantes nativos de PB residentes no Brasil. O *corpus* de seu estudo contou com a leitura de palavras em lista e, posteriormente, os sujeitos foram interpelados a montar frases com as palavras-alvo em posição medial. Os valores médios de VOT obtidos por Major (1992) para o inglês americano foram de 78ms para [p], 84ms para [t] e 93ms para [k].

Schwartzhaupt (2012), por sua vez, investigou os efeitos de ponto de articulação, qualidade da vogal e número de sílabas da palavra-alvo na duração do VOT de consoantes oclusivas surdas. Ele analisou os dados de 10 brasileiros falantes proficientes de inglês e 5 falantes nativos de inglês norte-americano. As palavras-alvo monossílabas e dissílabas foram inseridas em frase-veículo e os resultados foram apresentados tanto em termos absolutos quanto em duração relativa de VOT sobre a frase-veículo. Os resultados concernentes à produção de monossílabos por falantes nativos de inglês, considerando os valores médios de VOT absoluto foram de 64,34ms para [p], 79,88 para [t] e 91,68 para [k]²⁶. Os valores de duração relativa em relação à frase-veículo foram de 5,26% para [p], 6,59% para [t] e 7,5% para [k]. A aplicação do teste estatístico MANOVA revelou diferenças significativas com relação ao ponto de articulação [$F(2,117) = 26,163; p < 0,05$], que foram confirmadas através da aplicação do teste de comparação pareada de Bonferroni, cujos resultados reportados foram [p]x[t] = $p < 0,05$; [p]x[k] = $p < 0,05$ e [t]x[k] = $p < 0,05$.

²⁵ KLATT, D. *Voice onset time, friction and aspiration in word initial consonant*. In: *Journal of Speech and Hearing Research*, v. 18, 1975.

²⁶ Esses valores correspondem à média dos resultados apurados pelo autor, uma vez ele opta por separá-los com relação à qualidade da vogal.

A apresentação dos valores médios de VOT das pesquisas já mencionadas pode ser visto na Tabela 2. Note que optamos pela seleção dos valores de VOT obtidos por Lisker e Abramson (1965) quando a palavra-alvo está em posição medial na sentença, uma vez que os estudos de Klatt (1975) e Schwartzhaupt (2012) foram feitos em condições semelhantes.

Inglês norte-americano			
	[p]	[t]	[k]
Lisker e Abramson (1964)	34	37	49
Klatt (1975)	47	65	70
Macken e Barton (1978)	47-81	67-87	70-90
Major (1992)	78	84	93
Schwartzhaupt (2012)	64	80	92

Tabela 2: Valores médios em milissegundos apurados para as oclusivas surdas do inglês norte-americano em diferentes estudos.

Nota-se que os maiores valores de VOT foram obtidos por Major (1992). Isso possivelmente se deve ao fato de que em seu estudo foi incluída a realização de palavras isoladas, o que já sabemos, aumenta consideravelmente a duração do VOT. Entretanto, os valores médios encontrados por Schwartzhaupt (2012) são praticamente tão altos quanto os de Major (1992), embora seu experimento tenha sido feito com a utilização de frases-veículo.

Voltemo-nos agora à questão da influência da vogal seguinte sobre duração do VOT. Segundo Yavaş (2007) oclusivas diante de vogais altas tendem a apresentar VOTs mais longos porque tais vogais ocorrem com a cavidade oral mais fechada, o que faz com que a transição entre a soltura da oclusão e a configuração do trato para a realização de uma vogal alta se dê de maneira menos abrupta, o que resulta em retardamento no início do vozeamento em relação a uma oclusiva que estivesse diante de uma vogal baixa. Yavaş (2007) conduz um experimento sobre aquisição de língua inglesa com 16 falantes nativos de espanhol de diferentes nacionalidades. O *corpus* de seu estudo conta com 12 dissílabas inseridas em frases-veículo. Os resultados apontam que, para todos os sujeitos, o VOT das oclusivas diante de vogal alta mostrou-se mais longo para [p] e [k], não sendo confirmado para [t] com relação a 5 dos sujeitos. Estatisticamente, todas as comparações resultaram em diferenças significativas, exceto quanto aos pares: (alveolar+v. alta) vs. (alveolar +v.baixa) e (alveolar+v.baixa) vs. (velar+v.baixa). Tais resultados se assemelham aos encontrados anteriormente por Port e Rotunno (1979). Neste estudo, o *corpus* contou com 18 palavras isoladas lidas por 5 sujeitos falantes nativos de inglês norte-americano. Seus resultados demonstram que oclusivas diante das vogais altas [i] e [u] possuem VOT 7% mais longo do

que diante da vogal baixa [a], tendo sido a diferença estatisticamente significativa atestada por testes *t* com valor de $p < 0,01$ em cada caso.²⁷

Já no estudo de Schwartzhaupt (2012) não foram encontradas diferenças significativas em relação à altura da vogal na produção de falantes nativos. O teste MANOVA foi aplicado para cada ponto de articulação separadamente e apresentou os seguintes resultados: $[F(3,36) = 0,627; p = 0,602]$ para [p], $[F(3,36) = 0,916; p = 0,443]$ para [t] e $[F(3,36) = 0,895; p = 0,453]$ para [k]. O autor observa, entretanto, que os valores absolutos sugerem uma diferença que poderia mostrar-se significativa, especialmente com relação à bilabial, caso o número de sujeitos fosse ampliado. Ele aponta que para [p] o valor médio de VOT absoluto diante da vogal alta posterior é de 70,82ms, ao passo que diante da vogal baixa anterior é de 56,58ms.

2.4.2 OCLUSIVAS SURDAS NO PORTUGUÊS BRASILEIRO

Quanto ao PB, é sabido que a duração do VOT nas consoantes surdas iniciais é consideravelmente menor do que no inglês. No estudo de Major (1992), a análise dos dados de 5 falantes de PB que constituíam grupo controle do experimento resultou nos seguintes valores médios de VOT: 11ms para [p], 15ms para [t] e 35ms para [k].

Klein (1999) analisou dados de 4 sujeitos falantes nativos de PB e utilizou um *corpus* com logatomas e palavras reais, todos seguindo a estrutura CV (consoante + vogal). A pesquisadora realizou dois estudos-piloto: no primeiro, o *corpus* constituiu-se de logatomas monossilábicos, produzidos tanto isoladamente quanto em frases-veículo do tipo “*Digo _____ pra ela*”; o segundo foi conduzido com palavras dissilábicas, trissilábicas e polissilábicas produzidas tanto isoladamente quanto em diferentes posições na sentença. O contexto silábico em ambos os estudos constituiu-se das vogais [e, a, o].

Os resultados apontam que o PB, assim como o espanhol, o holandês, o húngaro e o tâmil, é uma língua de 2 categorias, ou seja, o VOT localiza-se primordialmente em duas áreas distintas: nas oclusivas sonoras localiza-se na área de pré-sonorização e, nas oclusivas

²⁷ Os autores não incluem os resultados obtidos através do teste T, limitando-se a afirmar que a diferença foi significativa.

surdas, na área de retardo curto. Apesar de o inglês possuir também duas categorias, elas se encontram nas áreas de retardo curto e retardo longo.

Os valores médios de VOT obtidos por Klein (1999) para a área de retardo curto das oclusivas surdas no PB foram os seguintes: 15,58ms para [p], 16,69ms para [t] e 36,36ms para [k]. Se comparados com os resultados médios obtidos por Istre (1985²⁸, apud KLEIN, 1999), a discrepância maior se localiza na bilabial, cujo VOT médio foi de 11,95ms. Para a alveolar o valor médio encontrado por Istre foi de 18,48ms, ao passo que para a velar foi de 38,53ms.

A fim de observar os efeitos do treino perceptual na identificação e produção de oclusivas surdas do inglês, Reis e Nobre-Oliveira (2007) analisaram dados de 11 acadêmicos da Universidade Federal de Santa Catarina que, à época da coleta, estudavam inglês há 4 anos em média. Para fins de comparação, foram medidos os valores de VOT dos sujeitos tanto em dados produzidos em inglês como em português. O *corpus* foi formado por uma lista de 21 palavras, nas quais a oclusiva encontrava-se sempre em posição inicial de palavra. Os valores obtidos para o PB foram os seguintes: 17,27ms para [p], 23,55ms para [t] e 46,55ms para [k].

Alves e Dias (2010) fizeram uma análise acústica da produção de oclusivas surdas do PB diante das vogais altas [ɪ] e [ʊ] por 4 sujeitos adultos (dois do sexo masculino e dois do feminino). O *corpus* foi composto de frases nas quais foram inseridas palavras-alvo terminadas em [pʊ], [tʊ], [kʊ], [pɪ] e [kɪ], como por exemplo, “horóscopo”, “contato”, “buraco”, “gripe” e “craque”. Foi excluída a alveolar diante de [ɪ], pois os sujeitos realizavam a variante africada [tʃɪ]. As pesquisadoras encontraram as seguintes médias de VOT para as oclusivas: 48,17ms para [p], 50,07ms para [t] e 63,89ms para [k], o que confirmou a presença de considerável grau de aspiração na realização desses segmentos. Valores tão altos como esses, no entanto, não foram encontrados em estudos semelhantes, e são ainda mais inquietantes se lembrarmos que as oclusivas desse *corpus* encontram-se em sílaba átona final, que é reduzida e costuma apresentar duração consideravelmente menor que sílabas em outras posições acentuais (MAJOR, 1985). Por outro lado, é preciso notar que os valores apurados pelas pesquisadoras incluem tanto dados nos quais as vogais finais são realizadas (59,9%), quanto dados nos quais ocorre seu apagamento (40,1%). É possível que, devido ao apagamento da vogal final, tenha havido um prolongamento da oclusiva. De fato, quando os dados são analisados de forma separada, as médias de VOT das oclusivas diante de vogais apagadas superam as médias das oclusivas diante de vogais cuja realização se concretiza. O teste estatístico de *Mann-Whitney*, com nível de significância de $p < 0,05$ confirma que essa

²⁸ ISTRE, G. L. (s/d) *Um estudo do VOT em monolíngües brasileiros*. (trabalho não publicado).

diferença é significativa para [p] ($p < 0,007$) e para [k] ($p < 0,006$).²⁹ Na Tabela 3, é possível observar as médias de VOT das oclusivas diante de [ɪ] e diante de [ʊ] nos contextos em que as vogais são apagadas e nos contextos em que não são.

	[p]		[t]	[k]	
	[ʊ]	[ɪ]	[ʊ]	[ʊ]	[ɪ]
não apagadas	37,67	31,38	36,52	47,83	52,95
apagadas	65,88	84,82	56,51	74,72	94,40

Tabela 3: Médias de VOT em milissegundos encontradas nas oclusivas [p t k] diante de vogal alta posterior [ʊ] e anterior [ɪ], segundo Alves e Dias, 2010.

O teste estatístico aplicado demonstra que, quando há apagamento de vogal, o VOT das oclusivas tem duração maior diante de [ɪ], sendo essa diferença significativa para [p] ($p < 0,007$) e para [k] ($p < 0,006$).³⁰ Os dados em que não há apagamento de vogal não foram submetidos ao teste estatístico.

Outro fator analisado em Alves e Dias (2010) é a influência exercida pelo contexto em que as palavras-alvo encontram-se: posição medial ou posição final de sentença. A esse respeito, as autoras continuam optando por contrastar somente os dados nos quais há apagamento de vogal e atestam que, em posição final os valores médios de VOT ultrapassam aqueles das oclusivas em posição medial na sentença. Elas apontam que essa diferença é significativa para [p] ($p < 0,001$), [t] ($p < 0,002$) e [k] ($p < 0,001$).

França (2011), em seu trabalho sobre a aquisição de oclusivas surdas do inglês por falantes de PB, analisa dados de 22 sujeitos, que foram solicitados a produzir sentenças tanto em português quanto em inglês. Para o PB, o *corpus* é constituído de logatomas dissílabos paroxítonos inseridos em início de frase-veículo. Os valores encontrados são os seguintes: [p] = 19,56ms, [t] = 21,66ms e [k] = 47,20ms. O teste de *Friedman* demonstrou diferença significativa entre os três pontos de articulação ($p < 0,001$).³¹

Por fim, o estudo de Schwartzhaupt (2012) apresenta valores médios de 15,13ms para [p], 17,87ms para [t] e 58,05ms para [k]. A duração relativa de VOT tem médias de 1,35%, 1,56% e 4,98%, respectivamente. Na Tabela 4, temos um resumo dos resultados obtidos nas pesquisas supracitadas:

²⁹ Em todo o artigo, as autoras não incluem os valores de Z obtidos pelo teste de *Mann-Whitney*, limitando-se a apresentar os valores de p a fim de caracterizar as diferenças encontradas como estatisticamente significativas.

³⁰ É preciso que se questione se realmente podemos falar em VOT quando as vogais finais são apagadas, já que, por definição, o VOT é o intervalo entre a soltura dos articuladores e o vozeamento do segmento seguinte, que, nesse caso, está ausente.

³¹ França (2011) menciona apenas o valor de p em seu trabalho, não incluindo os resultados do teste de *Friedman*.

Português brasileiro			
	[p]	[t]	[k]
Istre (1985)	11,95	18,48	38,53
Major (1992)	11	15	35
Klein (1999)	15,58	16,69	36,36
Reis e Nobre-Oliveira (2007)	17,27	23,55	46,55
Alves e Dias (2010)	37,67/31,38	36,52	47,83/52,95
França (2011)	19,56	21,66	47,20
Schwartzhaupt (2012)	15,13	17,87	58,05

Tabela 4: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do português brasileiro em diferentes estudos

Primeiramente, deve-se atentar para o alto valor de VOT encontrado para a oclusiva velar no estudo de Schwartzhaupt (2012). O autor sugere que a discrepância entre esse valor e aqueles encontrados para as oclusivas bilabial e alveolar possa se dever ao fato de que os sujeitos de seu estudo, sendo falantes proficientes de inglês, tendam a transferir um padrão de VOT mais longo da L2 para sua língua materna. Embora seja uma justificativa possível, seria interessante investigar por que a transferência desse padrão não se estenderia às demais oclusivas. É razoável supor, por exemplo, que uma vez que o [k] apresentou VOT consideravelmente mais longo nos dados dos nativos de Schwartzhaupt (2012), isso poderia ocasionar maior percepção do VOT nesse ponto e, conseqüentemente, levar a transferência. Há de se considerar ainda se o dialeto dos falantes de PB da região em questão não possui, por si só, oclusivas velares com VOT mais longo, o que explicaria a transferência. Submetendo os dados ao teste MANOVA, apurou-se diferença significativa com relação ao ponto de articulação para o PB [$F(2,87) = 178,486; p < 0,05$]. A posterior aplicação do teste de Bonferroni sugere que a diferença é significativa entre [k] e [p] ($p < 0,05$) e entre [k] e [t] ($p < 0,05$). Entretanto, entre [p] e [k] a diferença não se provou significativa ($p = 0,674$).

Quanto ao estudo de Alves e Dias (2010), notam-se valores consideravelmente altos de VOT para as oclusivas bilabial e alveolar, mesmo quando levamos em conta apenas os valores obtidos para as consoantes diante das vogais que foram efetivamente realizadas. Isso pode ser devido ao uso exclusivo de vogais altas seguindo as oclusivas (YAVAŞ, 2007). Além disso, o *corpus* de seu estudo inclui palavras-alvo tanto em posição medial como em final de sentença. Este último contexto parece propiciar vogais mais longas e, possivelmente, as oclusivas que estão diante de tais vogais possam ter o VOT alongado também.

Klein (1999), inclusive, encontra resultados que reforçam, pelo menos em parte, essa perspectiva. Ela analisou frases nas quais as oclusivas estão em palavras trissílabas, todas em contexto vocálico [a] nas seguintes posições: (a) inicial, (b) inicial seguindo um determinante,

(c) medial e (d) final. Sua conclusão foi a de que os maiores valores médios de VOT para [t] = 13,27ms e [k] = 30,93ms ocorreram em posição final e para [p] 13,12ms, em posição medial. Ao aplicar o teste *t-de-Student* com nível de significância $p < 0,05$, as diferenças significativas ocorreram somente entre os seguintes pontos: para [p] entre a posição inicial e medial ($t = -2,499$; $p = 0,02$) e entre a inicial com determinante e a medial ($t = -3,126$; $p = 0,009$); para [t] e para [k] somente entre a medial e a final, com respectivos ($t = -2,611$; $p = 0,02$) e ($t = -3,258$; $p = 0,007$).

Além da influência do ponto de articulação sobre a duração do VOT, Klein (1999) analisa ainda os dados obtidos de palavras inseridas em sentenças, a fim de testar como a tonicidade da sílaba e a posição na sentença influenciam a duração do VOT. Seus dados sugerem que o VOT sofre influência da tonicidade da sílaba; porém, as alterações não são sistemáticas. As oclusivas bilabial e alveolar apresentaram valores médios de VOT maiores na posição pós-tônica, ao passo que a velar teve os maiores índices na posição tônica.

Destacamos ainda que, dentre todos os estudos, apenas o de Reis e Nobre-Oliveira (2007) utiliza exclusivamente palavras isoladas. Major (1992) e Klein (1999) utilizam tanto palavras isoladas quanto frases e Schwartzhaupt (2012) utiliza exclusivamente frases-veículo. De todo modo, a relação entre ponto de articulação e VOT é atestada em todos os trabalhos, confirmando que quanto mais posterior o ponto, maiores são os valores de VOT.

No que diz respeito às vogais seguintes, Alves e Dias (2010) constataram valores de VOT significativamente maiores para as oclusivas diante de [i] do que diante de [u], o que sugere que a posição anterior propicia a maior duração do VOT. Já os resultados encontrados por Schwartzhaupt (2012) sugerem VOT mais longo em bilabiais diante de vogais altas posteriores do que diante de anteriores, enquanto o inverso se dá para oclusivas velares. Tais resultados podem indicar que o ponto de articulação deve ser considerado na investigação da influência da qualidade vocálica na duração do VOT. Além disso, o pesquisador aponta que, para as oclusivas diante de vogais baixas, os valores de VOT encontrados foram muito semelhantes tanto diante de vogais posteriores quanto de anteriores, o que indica que o contraste relativo à posição ântero-posterior não parece funcionar da mesma maneira com vogais altas e baixas.

Ainda quanto ao contexto vocálico, Klein (1999) analisa a produção de oclusivas surdas em logatomas isolados diante de [a], [e] e [o] e aponta que os valores de VOT das oclusivas diante da vogal baixa [a] foram significativamente menores para os três pontos de articulação. Especificamente para a bilabial [p], houve sensibilidade significativa para todos os contextos vocálicos, sendo os maiores índices de VOT flagrados diante da vogal média

posterior [o]. Já nos logatomas produzidos em frases-veículo, os contextos vocálicos interferiram de maneiras diferentes para os três pontos de articulação. Para a bilabial, a vogal média posterior [o] resultou no VOT de maior valor; para a alveolar, a vogal média anterior [e] causou o VOT mais longo. Quanto à velar, a diferença significativa ficou entre a vogal baixa [a] e as médias, que apresentaram os maiores índices de VOT. De forma geral, entretanto, todas as oclusivas apresentaram VOT com valores médios mais baixos diante de vogal baixa (Tabela 5).

Português brasileiro						
	Isolado			Frase-veículo		
	[a]	[e]	[o]	[a]	[e]	[o]
[p]	11,34	13,98	27,77	12,51	14,11	19,84
[t]	14,03	18,64	17,67	15,07	19,19	17,13
[k]	33,84	40,67	39,04	29,68	36,37	35,83

Tabela 5: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas do português brasileiro diante de diferentes vogais. (KLEIN, 1999)

É interessante confrontar esses resultados com os obtidos por Alves e Dias (2010), cujo trabalho abordou apenas oclusivas diante de vogais altas e obteve valores médios de VOT entre 31 e 52ms, portanto, muito mais altos do que os valores médios obtidos por Klein (1999) para a vogal baixa [a], entre 11 e 33ms. Muito embora seja necessário ponderar que as pesquisas foram feitas sob diferentes condições, não devemos descartar a possibilidade de estarmos diante de uma espécie de compensação segundo a qual a duração da vogal afeta a duração do VOT da oclusiva que a precede. Para isso, é preciso levar em conta um modelo no qual o acento lexical se manifeste no nível da sílaba e não do segmento vocálico. Isso nos leva a, segundo Massini-Cagliari (1992, p.36-37):

aceitar que existe um jogo de compensações entre as durações da consoante inicial e do núcleo dentro da sílaba, mas que não prejudica o valor da duração total da sílaba. Em outras palavras, tanto a consoante inicial como o núcleo podem ter mais duração dentro de uma determinada sílaba, desde que o fato de um durar mais que o outro não afete a duração total da sílaba em questão e a relação de proeminência que esta sílaba estabelece com as outras da mesma palavra.

Dessa forma, sabendo-se que vogais mais baixas são inerentemente mais longas (DELGADO MARTINS, 1986³² apud MASSINI-CAGLIARI, 1992; MORAES, 1986³³, apud MASSINI-CAGLIARI, 1992), isso justificaria um VOT mais curto nas oclusivas que as

³² DELGADO MARTINS, M.R. *Sept études sur la perception*. Lisboa: Instituto Nacional de Investigação Científica, 1986.

³³ MORAES, J.A. de. *Acentuação lexical e acentual frasal em português: um estudo acústico-perceptivo*. In: II Encontro Nacional de Fonética e Fonologia. Brasília, 1986.

precedem. Inversamente, vogais altas são inerentemente mais breves, o que levaria a um VOT mais longo.

Andrade e Mascarenhas (1994³⁴, apud KLEIN 1999) apresentam uma pesquisa realizada por Viana (1984, apud KLEIN, 1999) sobre o português europeu que corrobora a perspectiva acima, pois encontra valores médios de VOT maiores para as oclusivas diante de vogais altas do que para não altas.

Por outro lado, é possível levar em conta o aspecto aerodinâmico da produção das oclusivas para explicar sua relação com a altura da vogal seguinte. Yavas e Wildermuth (2006), em sua análise dos efeitos da altura da vogal seguinte na aquisição de oclusivas aspiradas do inglês por falantes nativos de espanhol, observaram que o VOT tende a aumentar de acordo com a altura da vogal, uma vez que vogais altas têm uma cavidade mais obstruída do que as vogais baixas, pois a língua em posição alta resulta em um declive de pressão menos abrupto. Dessa forma, uma oclusiva assim produzida contará com um intervalo de VOT mais longo do que antes da vogal baixa, que demanda uma posição da língua baixa e, portanto, um declive de pressão mais abrupto.

2.4.3 OCLUSIVAS SONORAS NO INGLÊS

De forma geral, considera-se que, na língua inglesa, as consoantes oclusivas sonoras possuem VOT próximo a zero, uma vez que a soltura dos articuladores é normalmente simultânea ao vozeamento do segmento seguinte. Há a possibilidade ainda de que o VOT para oclusivas sonoras assumam valores negativos, quando o vozeamento precede a soltura em alguns milissegundos. Assim sendo, valores médios para o VOT de oclusivas sonoras em inglês ficam entre -20ms e +20ms. (KENT; READ, 1992)

Na análise das oclusivas sonoras, Lisker e Abramson (1964) preferiram não fazer uma média única com os dados dos 4 sujeitos norte-americanos pesquisados, pois observaram que isso levaria a conclusões equivocadas. Isso porque apenas um deles foi responsável por 95% dos dados nos quais as oclusivas sonoras apresentaram pré-vozeamento. Na maior parte dos dados dos outros 3 sujeitos, o VOT foi positivo. Foram feitas, portanto, duas médias distintas: uma de dados com VOT positivo (*voicing lag*) e outra de dados com VOT negativo (*voicing*

³⁴ ANDRADE, A.; MASCARENHAS, I. *Para um estudo do vozeamento em início de vogal diante de consoante oclusiva*. Centro de Linguística da Universidade de Lisboa, CIP-Lisboa, 1994.

lead). Os valores médios encontrados em palavras isoladas para [b], [d] e [g] foram, respectivamente: 1ms, 5ms e 21ms nos dados de VOT positivo e -101ms, -102ms e -88ms para VOT negativo. Para as oclusivas em posição inicial de frase, os pesquisadores obtiveram médias de 7ms, 9ms e 17ms para VOT positivo e -65ms, -56ms e -45ms para o VOT negativo. Finalmente, para as oclusivas em posição medial, as médias foram de 4ms, 7ms e 16ms para VOT positivo e -63ms para o [b] em VOT negativo. A ausência de valores de VOT negativo para [d] e [g] é justificada pelo fato de que os pesquisadores relataram a existência do chamado *unbroken voicing*, ou seja, vozeamento proveniente do segmento precedente que segue sem interrupção até a soltura dos articuladores da oclusiva em questão. Esse fenômeno se deu em todas as ocorrências do [d] e do [g] e, embora tenha ocorrido também com o [b], permitiu que fossem realizadas algumas medições. Na Tabela 6, temos um resumo dos dados das oclusivas sonoras.

Inglês norte-americano			
	[b]	[d]	[g]
Palavras isoladas	1/-101	5/-102	21/-88
Posição inicial	7/-65	9/-56	17/-45
Posição medial	4/-63	7	16

Tabela 6: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do inglês norte-americano em estudo de Lisker e Abramson (1964)

No estudo de Klatt (1975), anteriormente mencionado, cuja análise envolveu oclusivas em posição inicial de palavra inseridas na frase veículo “*Say_____ instead.*”, os seguintes valores médios de VOT foram apurados para as oclusivas sonoras: 11ms para [b], 17ms para [d] e 27ms para [g]. Já em Macken e Barton (1978), foram reportados os seguintes valores de VOT para as oclusivas sonoras do inglês: de 1ms a 11ms para [b], 5 a 17ms para [d] e 21 a 27ms para [g].

Na Tabela 7, temos um resumo dos valores de VOT para as oclusivas sonoras, apurados nas diferentes referências já mencionadas. Como anteriormente, optamos pela seleção dos valores de VOT em Lisker e Abramson (1965) quando a palavra-alvo está em posição medial na sentença.

Inglês norte-americano			
	[b]	[d]	[g]
Lisker e Abramson (1964)	4/-63	7	16
Klatt (1975)	11	17	27
Macken e Barton (1978)	1-11	5-17	21-27

Tabela 7: Valores médios em milissegundos apurados para as oclusivas sonoras do inglês norte-americano em três diferentes estudos.

É interessante registrar que, em um estudo posterior, Lisker e Abramson (1967) voltam a observar que as oclusivas sonoras do inglês apresentam valores de VOT que se

enquadram em duas faixas descontínuas, uma de aproximadamente -100ms e a outra de valores próximos a zero. Isso os leva a sugerir que, no inglês, há duas categorias fonéticas para [b,d,g] ao invés de uma só: a primeira se refere a sonoras produzidas com pré-vozeamento e a segunda a sonoras produzidas com VOT zero ou positivo. Observam ainda que, quando os valores são próximos a zero, há a tendência de que a duração de VOT aumente conforme o ponto de articulação seja mais posterior. Tal relação, entretanto, não ocorre quando há pré-vozeamento.

2.4.4 OCLUSIVAS SONORAS NO PORTUGUÊS BRASILEIRO

No português brasileiro, em contrapartida, o vozeamento nas oclusivas sonoras se inicia bem antes da soltura, o que nos leva a valores negativos que superam os do inglês.

Klein (1999), em uma análise geral, obtém os seguintes valores médios para as oclusivas sonoras: [b] = - 92,27ms, [d] = - 92,07ms e [g] = - 78,20ms. Nota-se que a área de pré-sonorização das bilabiais e alveolares é praticamente a mesma. A autora compara seus resultados com aqueles encontrados por Istre (1983, apud Klein, 1999) e observa que, ao contrário deste autor, que classificou as consoantes sonoras do PB como sendo de precedência média de sonorização, seria mais adequado enquadrá-las na categoria de precedência longa. Istre (1983, apud Klein, 1999) encontrou os seguintes valores de VOT para as oclusivas sonoras: [b] -39,45ms, [d] -60,59ms e [g] -53,69ms.

Melo et al. (2011) realizam uma pesquisa para comparar as características acústicas de oclusivas surdas e sonoras na fala de 11 crianças e 17 adultos sem patologia de fala relatada. São utilizadas palavras e pseudopalavras dissílabas paroxítonas ([papa], [baba], [tata], [dada], [kaka] e [gaga]) inseridas na frase-veículo “Fala _____ de novo”. Os resultados, de forma geral, apontaram para VOT mais longo para oclusivas sonoras em comparação às surdas, com os seguintes valores médios de VOT para oclusivas em *onset* inicial produzidas por adultos com idades entre 19 e 44 anos: [b]= -101,9ms, [d] = -95,6ms e [g]= -80,4ms.

A comparação entre os resultados das três pesquisas segue na Tabela 8. Vale notar que, em todas as pesquisas, ao contrário do que ocorre com as oclusivas surdas, o ponto velar não é o que detém os maiores valores de VOT, sendo, inclusive, o ponto com VOT de menor

duração nos estudos de Klein (1999) e Melo et al.(2011). Tal fato também ocorre nos dados de pré-sonorização coletados por Lisker e Abramson (1964).

Português brasileiro			
	[b]	[d]	[g]
Istre (1983)	-39,45	-60,59	-53,69
Klein (1999)	-92,27	-92,07	-78,20
Melo et al. (2011)	-101,9	-95,6	-80,4

Tabela 8: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro em três estudos diferentes

Voltando a Klein (1999), a pesquisadora observou variação considerável dos valores de VOT nos diferentes contextos analisados. A Tabela 9 mostra os valores médios de VOT em milissegundos encontrados para as oclusivas em logatomas produzidos isoladamente e em frases-veículo.

Português brasileiro			
	[b]	[d]	[g]
Logatoma	-86,57	-98,84	-91,16
Logatoma F/V	-104,98	-93,31	-76,93

Tabela 9: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro em estudo de Klein (1999)

A autora aponta que os resultados para [d] e [g] corroboram os resultados encontrados por Lisker e Abramson (1964), já que no estudo destes pesquisadores, todas as oclusivas inseridas em frases produzidas pelos sujeitos apresentaram valores de VOT reduzidos em comparação aos valores de segmentos produzidos em palavras isoladas. Por outro lado, a oclusiva bilabial não seguiu a mesma tendência.

Com relação ao contexto vocálico, as diferenças se mostraram estatisticamente significativas apenas em parte dos dados. Por exemplo, a produção da oclusiva bilabial diante de [a], [e] e [o] só variou significativamente nos logatomas produzidos em frases-veículo, situação em que a média de VOT da bilabial diante de [o] foi significativamente menor. Para a dental [d] observaram-se diferenças tanto na produção em logatomas isolados quanto em frases-veículo. Isoladamente, no contexto da vogal média [e], o VOT foi significativamente maior, ao passo que nas frases-veículo com a vogal baixa [a], o VOT foi significativamente menor que nas demais. Finalmente, quanto à oclusiva velar nos logatomas isolados, a diferença ficou por conta do contexto de [e] em relação à [o], ao passo que nas frases-veículo, a diferença significativa ficou por conta de [a] com relação às outras vogais.

Diante desses dados, Klein (1999) sugere que existe influência do contexto vocálico para as oclusivas sonoras do PB; embora essas alterações sejam assimétricas através dos diferentes pontos de articulação. Consideração semelhante foi feita por Schwartzhaupt (2012)

em relação a consoantes oclusivas surdas. A Tabela 10 traz o resumo dos resultados obtidos pela pesquisadora.

Português brasileiro						
	Isolado			Frase-veículo		
	[a]	[e]	[o]	[a]	[e]	[o]
[b]	-91,84	-81,43	-87,92	-106,88	-108,80	-99,25
[d]	-89,15	-110,51	-96,85	-85,15	-100,42	-95,02
[g]	-88,46	-99,14	-82,38	-68,11	-87,73	-79,84

Tabela 10: Valores médios em milissegundos encontrados para o VOT das oclusivas sonoras do português brasileiro diante de diferentes vogais (KLEIN, 1999).

Klein (1999) analisa ainda os dados obtidos de palavras inseridas em sentenças, a fim de testar como a tonicidade da sílaba e a posição na sentença influenciam a duração do VOT. A pesquisadora encontrou valores significativamente maiores de VOT para as oclusivas sonoras em posição tônica nos três pontos de articulação, o que nos remete novamente à característica do PB de ter na duração o principal correlato acústico de acento (MASSINI-CAGLIARI, 1992).

Com relação à posição na sentença, a autora analisa frases nas quais as oclusivas estão em palavras trissílabas, todas em contexto vocálico [a] nas seguintes posições: (a) inicial, (b) inicial seguindo um determinante, (c) medial e (d) final. Com a utilização do teste *t-de-Student* e nível de significância de $p < 0,05$, a autora observa que a única oclusiva a apresentar diferença estatisticamente significativa quanto à posição na sentença foi a bilabial, entre as posições medial e final ($t = -2,577 / p = 0,02$). Ela aponta que, ao contrário do esperado, o valor médio de VOT na posição final foi menor na posição medial. Diante disso, temos que, para os dados analisados por Klein (1999), a posição na sentença não parece ter influência na duração de VOT das oclusivas sonoras.

2.5 PRODUÇÃO DE OCLUSIVAS NA AQUISIÇÃO DE INGLÊS POR FALANTES NATIVOS DE PB

Passemos agora a observar mais detidamente o que relata a literatura sobre a produção de oclusivas por aprendizes de inglês falantes nativos de PB.

Em seu trabalho sobre a aquisição de oclusivas surdas do inglês por falantes de PB, França (2011) analisa dados de 22 sujeitos, que foram solicitados a produzir sentenças tanto em português quanto em inglês. Os valores encontrados para consoantes oclusivas surdas da

língua inglesa em dados obtidos de aprendizes falantes de PB são os seguintes: 27,43ms para [p], 45,83ms para [t] e 58,61ms para [k], ou seja, maiores do que produzidos pelos mesmos falantes na L1, a saber, 19,56ms para [p], 21,66ms para [t] e 47,20ms para [k]. Com relação aos dados do PB, a autora comparou a diferença entre os três pontos de articulação por meio de dois testes: o de *Friedman* teve como resultado o valor de $p < 0,001$ na comparação entre os três pontos, o que aponta diferença estatisticamente significativa entre eles; o teste de *Wilcoxon* apresentou resultado de $p < 0,001$ entre [p] e [k] e também entre [t] e [k], ou seja, houve diferença estatisticamente significativa entre esses pontos, embora a diferença entre [p] e [t] tenha sido de $p = 0,003$. Para comparar os dados produzidos em inglês, a autora fez novamente uso do teste de *Friedman*, que apontou valor de $p < 0,001$ na diferença entre os três pontos. O teste de *Wilcoxon* confirmou a diferença estatisticamente significativa, já que os valores obtidos na comparação entre os pontos foram sempre de $p < 0,001$.³⁵

Ressaltamos que a autora se utiliza de dados produzidos em frases-veículo nos quais a palavra-alvo encontra-se em posição inicial. Nos dados em português foram selecionados dissílabos paroxítonos e, nos dados em inglês, tanto monossílabos quanto dissílabos. Aqui se observa o padrão já atestado anteriormente de que a duração do VOT se estende conforme o ponto de articulação se torna mais posterior. Além disso, a diferença na produção de L1 e L2 leva a pesquisadora a concluir que os aprendizes se encontram em uma etapa de desenvolvimento em direção à língua-alvo. Entretanto, é preciso enfatizar que a diferença na constituição do *corpus* pode ter levado a resultados menos confiáveis do que se a pesquisadora mantivesse monossílabos tanto no *corpus* para o PB quanto para o inglês. De toda forma, a conclusão da pesquisadora se torna ainda mais clara quando ela organiza os dados produzidos em inglês conforme o nível de proficiência dos sujeitos. Para todos os pontos de articulação, ocorre aumento na duração do VOT quando os aprendizes passam do nível básico ao intermediário e, mais ainda, quando passam do intermediário ao avançado. Para a oclusiva bilabial, os índices são de 24,63ms para o básico, 26,69ms para o intermediário e 36,07ms para o avançado. Para a dental, temos os valores de 43,16ms, 43,59ms e 56,28, respectivamente. A velar conta com a seguinte escalada: 55,78ms, 60,96ms e 91,71ms. A autora enfatiza que mesmo os valores encontrados no nível básico são superiores àqueles produzidos nos dados em PB, o que, segundo ela, só confirma o crescente desenvolvimento na aquisição deste aspecto fonético-fonológico presente na L2. Na Tabela 11 é possível observar os valores de VOT encontrados por França (2011).

³⁵ Embora a autora sugira um resultado significativo para os dados, ela não fornece os resultados dos testes, limitando-se a apresentar os valores de *p* apurados.

Falantes nativos de PB aprendizes de inglês					
Português brasileiro		Inglês			
		Média	Básico	Intermediário	Avançado
[p]	19,56	27,43	24,63	26,69	36,07
[t]	21,66	45,83	43,16	43,59	56,28
[k]	47,20	58,61	55,78	60,96	91,71

Tabela 11: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por falantes nativos de PB (FRANÇA, 2011).

Além dos resultados relativos a ponto de articulação, França (2011) observa ainda a influência da vogal que sucede a oclusiva, atestando que, nos dados da interlíngua, as oclusivas diante de vogais altas possuem VOT mais longo do que diante de vogais baixas, assim como foi apurado para a interlíngua de falantes nativos espanhol em Yavaş (2007), para o inglês falado por nativos em Port e Rotunno (1979) e para o português falado por nativos em Klein (1999). Com a aplicação do teste de *Friedman*, tal diferença mostrou-se estatisticamente relevante para as três oclusivas, o que reforça a possibilidade de estarmos diante de uma espécie de compensação entre os segmentos que formam a sílaba. Assim, oclusivas tenderiam a apresentar duração mais curta quando associadas a vogais baixas, que são intrinsecamente mais longas e, inversamente, exibiriam maior duração quando diante de vogais altas, cuja duração é intrinsecamente menor. (MASSINI-CAGLIARI, 1992). Por fim, ao comparar a produção de monossílabos e dissílabos na interlíngua, França (2011) observou que os primeiros possuem VOT mais longo que os segundos em todas as oclusivas. Ao aplicar o teste de *Wilcoxon*, ela concluiu que tal diferença foi estatisticamente relevante para a alveolar e velar, ao passo que a bilabial não apresentou diferença estatisticamente significativa entre monossílabos e dissílabos.³⁶

Zimmer (2004) demonstra que conforme os aprendizes adquirem maior proficiência na língua, maiores tendem a ser os índices de VOT na realização de oclusivas surdas. A pesquisadora estuda a possibilidade de transferência de padrões fonético-fonológicos da L1 para a L2. Sabendo que, para as oclusivas surdas, os valores de VOT no PB tendem a ser consideravelmente mais baixos, ela espera que aprendizes brasileiros de inglês venham a transferir essa característica de sua língua materna para a língua estrangeira. A autora confirma essa hipótese em seu trabalho e refere-se ao fenômeno como desaspiração das oclusivas surdas iniciais. Segundo ela, o fator que leva falantes de PB à desaspiração envolve a ativação do conhecimento fonético-fonológico de sua língua materna. Em sua pesquisa, a autora trabalhou com a leitura de palavras e logotomas e analisou dados de 156 estudantes em

³⁶ Novamente a autora sugere um resultado significativo para os dados, mas não fornece os resultados dos testes.

quatro estágios diferentes de aprendizagem: 50 no nível básico, 57 no intermediário, 34 no intermediário-avançado e 15 no avançado. Seu trabalho revela que o processo de desaspiração diminui significativamente conforme aumenta o nível de proficiência dos sujeitos. Ela conclui que a formação categórica de contrastes fonéticos da L2 é dificultada para adultos em razão da experiência linguística do falante e não necessariamente na perda da plasticidade resultante da maturação neuronal. Vale lembrar que a própria questão da perda da neuroplasticidade tem sido questionada, pois estudos têm demonstrado que os neurônios não perdem a plasticidade à medida que os indivíduos envelhecem³⁷. Tal achado acaba sendo um forte contra-argumento para a hipótese do Período Crítico, amplamente utilizada nos estudos sobre aquisição de segunda língua, segundo a qual a capacidade de um indivíduo para aprender línguas diminui conforme a idade, sofrendo um declínio considerável depois da adolescência.

Já a pesquisa de Alves (2011) sobre a produção de oclusivas surdas do português e do inglês por falantes de PB conta apenas com falantes com alto grau de proficiência. Seus sujeitos são todos do sexo feminino e os dados são obtidos por meio de fala semi-espontânea. A proficiência dos sujeitos é justificada pelo fato de que todos estavam matriculados no curso de Linguística Aplicada oferecido pela Universidade Federal de Santa Catarina, e, para tanto, tiveram de passar por um teste escrito e entrevista na qual obtiveram um escore equivalente a, no mínimo, 550 no exame de TOEFL³⁸. Nos dados analisados pela autora, os falantes de PB produzem oclusivas com índices de VOT próximos aos dos nativos que compõem seu grupo-controle e acreditamos que isso se deva ao nível de proficiência de tais sujeitos. Isso nos leva a apostar na hipótese de que, para as consoantes surdas, há uma associação entre nível de proficiência e duração de VOT, tal que, quanto maior a duração de VOT, mais proficiente em língua inglesa é o sujeito. Procuraremos verificar se tal associação é verdadeira para o grupo de sujeitos aqui analisados.

Alves (2011) utiliza dois falantes nativos como grupo controle, muito embora seus dados tenham sido coletados sob condições diferentes, a saber, entrevistas obtidas na internet. A pesquisadora justifica essa escolha pela impossibilidade de encontrar participantes que fossem livres de qualquer influência do PB. Sua solução foi, portanto, selecionar duas entrevistas obtidas em *talk shows*, cujo conteúdo biográfico se assemelhasse às entrevistas

³⁷ Para informações mais detalhadas, ver Ferrari et al. Plasticidade neural: relações com o comportamento e abordagens experimentais. In: Psicologia: Teoria e Pesquisa. v.17, n. 2, 2001

³⁸ TOEFL é a sigla para *Test of English as a Foreign Language* (Teste de inglês como língua estrangeira). Um score de 550 pontos, ou cerca de 80 pontos na versão para a internet, equivale à pontuação mínima necessária para ter o nível C1 do Quadro Comum Europeu de Referência para Línguas ou obter o *Certificate in Advanced English*, emitido pela Universidade de Cambridge.

feitas com os sujeitos da pesquisa. Ambos os entrevistados são norte-americanos. Os resultados encontram-se na Tabela 12.

	Falantes nativos norte-americanos	Falantes de PB proficientes em inglês
[p]	63,86	48,8
[t]	69,13	52,95
[k]	59,07	60,35

Tabela 12: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por falantes nativos norte-americanos e falantes de PB com proficiência em inglês (ALVES, 2011).

É preciso considerar, no entanto, que a autora decidiu descartar as oclusivas produzidas de forma não aspirada pelos informantes de ambos os grupos, levando-se em conta a classificação de Cho e Ladefoged (1999), segundo a qual oclusivas aspiradas possuem VOT de duração igual ou superior a 35 milissegundos.

A diferença entre a produção de nativos e falantes proficientes foi medida por meio do teste *Mann-Whitney* com nível de significância $p < 0,05$. Os resultados mostraram diferença significativa entre os dois grupos para a oclusiva bilabial ($X^2(2) = -2,187$; $p = 0,029$) e alveolar ($X^2(2) = -3,069$; $p = 0,002$), ao passo que não houve diferença estatisticamente significativa na produção da velar ($p = 0,337$).³⁹ Alves (2011) também chama atenção para o fato de que, apesar de diversos trabalhos indicarem o progressivo aumento da duração do VOT conforme o ponto de articulação se torna mais posterior (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1998; LADEFOGED, 2011), o mesmo não se observou com relação aos dados de seus falantes nativos de inglês.

Os falantes de PB, por sua vez, produziram apenas a velar com diferença significativa em relação aos demais pontos ($p = 0,003$) entre [p] e [k]; ($p < 0,001$) entre [t] e [k]. Ela afirma que tal resultado não surpreende, visto que nas velares do PB, a duração do VOT costuma ser consideravelmente mais longa do que nos demais pontos (KLEIN, 1999; ALVES et al., 2008).

Para medir a influência da vogal seguinte na produção dos falantes de PB foi aplicado o teste *Kruskall-Wallis*, que revelou não haver diferença significativa entre oclusivas diante das vogais altas [i], [ɪ], [u] e [ʊ], as médias [ə], [ɛ], [ʌ] e [ɔ] e as baixas [æ] e [ɑ]. Com relação aos falantes nativos de inglês, a pesquisadora limita-se a afirmar que testes estatísticos foram aplicados e também não revelaram diferença significativa entre o VOT das oclusivas diante das vogais altas, médias e baixas. Ela não menciona os resultados obtidos por meio dos testes para nenhum dos casos. A pesquisadora pressupõe que a influência da vogal seguinte possa ter sido mascarada pela taxa de elocução dos sujeitos, uma vez que os dados foram obtidos de

³⁹ A autora não reporta os valores de Z.

fala semi-espontânea. Finalmente, no que tange à influência do número de sílabas, a autora utilizou o teste *Mann-Whitney* com nível de significância $p < 0,05$ para observar se o VOT de oclusivas inseridas em monossílabos exibe realmente maiores valores do que o VOT de oclusivas inseridas em palavras com mais de uma sílaba, como previsto em Lisker e Abramson (1964) e Klatt (1991⁴⁰, apud ALVES, 2011). Para falantes de PB, puderam ser analisadas somente as oclusivas alveolar e velar. Houve diferença significativa somente com relação à oclusiva velar ($Z=2,768$, $p=0,006$), ao passo que para a alveolar o resultado foi de $p=0,552$. Quanto aos nativos, não foi observada diferença significativa no VOT quanto ao número de sílabas para nenhuma das três oclusivas. A diferença para as bilabiais foi de $p=0,823$; para as alveolares, $p=0,337$ e para as velares, $p=0,371$ ⁴¹.

Por fim, o estudo de Schwartzhaupt (2012) aponta os seguintes valores médios de VOT absoluto na produção de monossílabos inseridos em frases-veículo por brasileiros falantes proficientes de inglês: 23,4ms para [p], 44,75ms para [t] e 68,33ms para [k]. Os valores relativos de VOT em relação à frase-veículo são de 2,16%, 4,3% e 6,38%, respectivamente. Se compararmos tais resultados com os obtidos por França (2011) e Alves (2011) para falantes proficientes, veremos que são relativamente mais baixos, exceto para a velar, que supera o valor encontrado em Alves (2011). O autor aponta, no entanto, que a diferença com relação ao estudo de França (2011) pode se dever ao fato de que o número de sujeitos por ela analisados foi menor – 4 em comparação aos 10 sujeitos por ele analisados. Já com relação à pesquisa de Alves (2011), fica difícil justificar os altos valores de VOT absoluto, pois embora ela tenha analisado apenas 5 sujeitos, há que se considerar que seus dados foram obtidos por meio de fala semi-espontânea, o que hipoteticamente levaria à produção de consoantes com VOT mais curto (Baran et al.⁴²1977, apud DOCHERTY, 1992) A Tabela 13 traz a comparação entre os três estudos.

	França (2011)	Alves (2011)	Schwartzhaupt (2012)
[p]	36,07	48,8	23,4
[t]	56,28	52,95	44,75
[k]	91,71	60,35	68,33

Tabela 13: Valores médios em milissegundos encontrados para o VOT das oclusivas surdas produzidas por brasileiros falantes proficientes de inglês em diferentes estudos.

⁴⁰ KLATT, D. *Voice onset time, friction and aspiration in word initial consonant*. In: BAKEN, R.J; DANILOFF, R.G.(Ed.) *Readings in Spectrography of Speech*. San Diego: Singular Publish. Group, Inc. & Kay Elemetrics Corp.

⁴¹ A autora somente reporta o valor de Z que citamos acima, sendo os demais resultados amparados somente pelo valor de p.

⁴² BARAN et al. *Phonological contrastivity in conversation: a comparative study of voice onset time*, *Journal of Phonetics*, v. 5, 1977.

Assim como nos demais casos analisados por Schwartzhaupt (2012), a diferença entre pontos de articulação na interlíngua provou-se significativa através da aplicação do teste MANOVA [$F(2,237) = 157,341; p < 0,05$], o que foi confirmado com o teste de *Bonferroni*, cujos resultados reportados são $[p]x[t] = p < 0,05$; $[p]x[k] = p < 0,05$ e $[t]x[k] = p < 0,05$.

2.6 CONCLUSÃO

Agora cumpre estabelecer os rumos de nossa pesquisa à luz das considerações que pudemos fazer a partir dos estudos apresentados neste capítulo. A primeira observação a ser retomada é a de que diversos fatores podem influenciar a duração do VOT, a saber: o ponto de articulação, a taxa de elocução, o próprio critério para a medição do VOT, a qualidade e a tonicidade da vogal seguinte, o número de sílabas da palavra-alvo e a sonoridade da consoante final (DOCHERTY, 1999). Portanto, a comparação entre os resultados de diferentes estudos deve ser feita com cautela e a constituição do *corpus* deve levar em conta todos os fatores acima citados.

Neste estudo, objetivamos elaborar um *corpus* com monossílabos da língua inglesa, o que está em consonância com os *corpora* utilizados na maior parte das pesquisas aqui apresentadas. Além disso, tencionamos incluir vogais de diferentes qualidades para suceder as oclusivas, o que deverá, a um só tempo, equilibrar a influência que a qualidade da vogal seguinte venha a exercer na duração do VOT e permitir que realizemos nossa própria avaliação dessa variável quanto ao grupo de sujeitos desta pesquisa.

Quanto à taxa de elocução, é importante observar que a maioria dos estudos realizados não a controla adequadamente (DOCHERTY, 1999), apesar de tal variável interferir na duração do VOT. Os estudos apresentados no decorrer deste capítulo utilizam os mais diversos métodos para a obtenção dos dados, desde a produção de fala espontânea e frases-veículo, que resultam em índices de VOT de menor duração, até a utilização de palavras isoladas, o que colabora para que os valores de VOT sejam mais longos. Disso decorre que, por vezes, a comparação que se faz da duração absoluta em diferentes trabalhos não leva em conta o quanto os valores de VOT podem variar a depender da escolha do *corpus*. O trabalho de Rocca (2002), por exemplo, compara a produção de seus sujeitos bilíngues (PB e inglês) com os resultados obtidos por Lisker e Abramson (1964) quando esses pesquisadores utilizavam palavras isoladas, ou seja, 58ms para [p], 70ms para [t] e 80ms para [k]. No entanto, os sujeitos de Rocca (2002) produzem monossílabos em frases-veículos do tipo “*I say ____ again*” e os valores que a pesquisadora toma para comparação poderiam, na

verdade, ter sido aqueles que Lisker e Abramson (1964) apuram para oclusivas inseridas em sentenças em posição medial: 34ms para [p], 37ms para [t] e 49ms para [k] e que são consideravelmente mais baixos. Se Rocca (2002) tivesse optado por utilizar os valores de VOT das oclusivas inseridas em frases, o resultado de sua pesquisa teria sido alterado consideravelmente.

Além disso, as pesquisas que utilizam VOT não costumam fazer uso de sua duração relativa, o que neutralizaria quaisquer desvios que venham a existir por conta da taxa de elocução dos informantes e aumentaria a confiabilidade dos dados. Nesta pesquisa, buscaremos informar o leitor não somente a respeito dos valores de duração absoluta de VOT, como também da duração relativa apurada através dos dados. Entendemos que a insistência na utilização exclusiva de valores absolutos de VOT, em detrimento de valores relativos, constitui empecilho que faz com que analisemos estudos, muitas vezes, enxergando resultados de maneira equivocada, ou, pelo menos, parcial. Por fim, seria interessante que os autores de estudos que envolvessem a utilização do VOT como parâmetro acústico buscassem uniformidade ao menos na extração dos dados, já que fica bastante claro que há subjetividade considerável nesse quesito.

3. METODOLOGIA

O estudo aqui realizado visa analisar a produção de consoantes oclusivas iniciais do inglês por falantes nativos de PB, comparando-a com a produção de falantes nativos de inglês. Este capítulo trata da metodologia utilizada para alcançar os objetivos desta pesquisa. Inicialmente, serão apresentadas as perguntas da pesquisa e as hipóteses levantadas e, em seguida, os critérios utilizados para a constituição do *corpus*, a escolha dos sujeitos e a análise dos dados, tanto do estudo-piloto quanto do experimento completo.

3.1 PERGUNTAS DA PESQUISA E HIPÓTESES

Inicialmente, cabe lembrar que os dados aqui analisados foram coletados junto a três grupos distintos. O primeiro deles, denominado grupo N, é composto por falantes nativos de inglês originários dos Estados Unidos. O segundo e terceiro grupos são formados por falantes de PB e estão divididos da seguinte forma: aprendizes de inglês em estágio intermediário de aprendizagem (grupo A) e falantes proficientes de língua inglesa (grupo P). Detalhamentos sobre os critérios de seleção dos sujeitos em cada um dos grupos encontram-se na seção 3.3. Após a condução do piloto e as devidas alterações, as perguntas de nossa pesquisa bem como as hipóteses a serem testadas culminaram no seguinte:

Pergunta 1: O nível de proficiência na língua inglesa interfere na produção das oclusivas por falantes de PB?

H1: Haverá diferença estatisticamente significativa entre a produção do grupo A e do grupo N com relação à duração do VOT nas oclusivas surdas.

H2: Não haverá diferença significativa entre a produção do grupo P e do grupo N com relação à duração do VOT nas oclusivas surdas.

H3: Haverá diferença significativa entre a produção do grupo A e do grupo P com relação à duração do VOT nas oclusivas surdas.

Embora o PB e o inglês contem com oclusivas surdas e sonoras nos mesmos pontos de articulação, diversos estudos indicam que tais consoantes são produzidas de maneira diferente nas duas línguas (LISKER; ABRAMSON, 1964; KLEIN, 1999; ALVES et al., 2008; SCHWARTZHAUPT, 2012). Com relação às surdas, os estudos apontam que falantes de PB tendem a manter na produção em inglês, características das oclusivas da L1, ou seja, sua produção costuma apresentar níveis de aspiração inferiores aos de um nativo de inglês

(MAJOR, 1992; SANCIER; FOWLER, 1997; ROCCA, 2003; ZIMMER, 2004; COHEN, 2004; REIS; NOBRE-OLIVEIRA, 2007; ALVES, 2011; FRANÇA, 2011). Gostaríamos, portanto, de observar se fato semelhante se aplica aos grupos N e A desta pesquisa.

Além disso, Zimmer (2004) e França (2011) apontam uma relação positiva entre o grau de proficiência do aprendiz e sua capacidade crescente em aproximar-se dos sons da língua-alvo. Sob um ponto de vista dinâmico⁴³, quanto mais alto o nível de proficiência do falante, maior controle ele terá na coordenação dos gestos para realizar os sons da L2 e, no caso da produção das oclusivas surdas por brasileiros, tal fato redundará em segmentos cujos valores médios de VOT tornar-se-ão progressivamente mais altos. Parece haver, portanto, para as consoantes surdas, uma associação entre nível de proficiência e duração de VOT, tal que, quanto maior a duração de VOT, mais proficiente em língua inglesa é o sujeito.

Com relação às oclusivas sonoras, não nos foi possível lançar hipóteses quanto à diferença na produção do grupo N e A, já que mesmo entre falantes nativos, tanto em nosso estudo piloto quanto em experimentos anteriores (LISKER; ABRAMSON, 1964) deflagrou-se falta de uniformidade na realização dessas consoantes, cujo VOT tende a variar entre valores positivos (intervalo curto) e negativos (pré-vozeamento). Isso impediu a condução de testes estatísticos neste estudo, uma vez que contamos com um número pequeno de informantes (6 em cada grupo), que tornou-se ainda mais reduzido ao dividirmos as produções das sonoras levando em conta as duas realizações distintas com as quais nos deparamos. Os dados das oclusivas sonoras serão, portanto, analisados sem o amparo de testes estatísticos.

Pergunta 2: O ponto de articulação interfere nos índices de VOT das oclusivas por nativos e falantes de PB?

H4: Haverá progressivo aumento na duração de VOT de [p], [t] e [k] e a diferença será significativa para todos os pontos, na produção dos grupos N, P e A.

A influência do ponto de articulação sobre os valores de VOT é um dos aspectos mais estudados no que se refere às oclusivas surdas. Tem sido atestado (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1999; KLEIN, 1999; COHEN, 2004; SCHWARTZHAUPT, 2012) que os índices de VOT tendem a aumentar quanto mais posterior é o ponto de articulação da consoante. Logo, uma oclusiva surda bilabial [p]

⁴³ Os modelos dinâmicos de produção da fala têm como primitivos de análise os gestos articulatórios. Assim, cada som corresponde a um ou mais gestos articulatórios, que pertencem a um subsistema oral, laríngeo ou vélico. Para mais informações, ver seção 2.1.3 deste trabalho.

apresentará VOT mais curto do que uma alveolar [t], que terá VOT mais curto que uma velar [k]. Investigaremos se tal tendência se aplicará aos sujeitos desta pesquisa e se observará uniformemente dentro dos três grupos pesquisados.

Embora não tenhamos formulado nenhuma hipótese para as oclusivas sonoras, esperamos encontrar o mesmo aumento progressivo na duração de VOT nos dados de intervalo curto. Quanto aos dados de pré-vozeamento, a expectativa é a de que as durações de nativos de inglês e falantes de PB não difira muito, uma vez que, ao isolarmos esse tipo de dado em diferentes estudos, tem-se por volta de -100ms de duração absoluta de VOT tanto para a produção em inglês (LISKER; ABRAMSON, 1964) quanto em PB (KLEIN, 1999; MELLO et al, 2011). Além disso, esses valores não parecem se modificar muito conforme o ponto de articulação.

Pergunta 3: A altura da vogal seguinte interfere nos índices de VOT das oclusivas por nativos de inglês e falantes de PB?

H5: A duração de VOT das oclusivas surdas diante da vogal alta [i] será significativamente maior do que diante da vogal baixa [æ] na produção dos grupos N, P e A, mesmo ao considerar os pontos de articulação em separado.

A duração do VOT é condicionada por diversos fatores: o ponto de articulação, a qualidade e a tonicidade da vogal seguinte, o número de sílabas da palavra-alvo e a sonoridade da consoante final (DOCHERTY, 1999). Neste estudo, além do ponto de articulação, estamos interessados em observar a influência da vogal seguinte nos índices de VOT. Há indícios de que, na realização de oclusivas surdas, os índices de VOT alcancem maior duração quando a oclusiva encontra-se diante de vogais altas. Segundo Klang et al. (1998), isso ocorre porque vogais altas possuem grau de abertura menor, o que faz com que o ar que escapa da boca sofra maior impedância⁴⁴, retardando assim a pressão transglotal necessária ao início do vozeamento. Entretanto, encontramos tanto pesquisas que corroboram essa visão (KLATT, 1975; YAVAS, 2008; FRANÇA, 2011), quanto pesquisas que a refutam (ALVES, 2011). Por outro lado, também é possível levar em conta um modelo no qual o acento lexical se manifeste no nível da sílaba e não do segmento vocálico. Segundo Massini-Cagliari (1992), isso nos leva a aceitar a existência de um jogo de compensações entre as durações da consoante inicial e do núcleo dentro da sílaba, de modo que as durações possam variar sem, no entanto, afetar a duração total da sílaba e a relação de proeminência que essa

⁴⁴ O dicionário Houaiss define impedância como o “produto entre a densidade de um meio e a velocidade do som neste meio, sendo uma medida da resistência à propagação do som”.

sílaba estabelece com as demais sílabas da palavra. Assim, sabendo-se que vogais mais baixas são inerentemente mais longas (DELGADO MARTINS, 1986⁴⁵ apud MASSINI-CAGLIARI, 1992; MORAES, 1986⁴⁶, apud MASSINI-CAGLIARI, 1992), tal fato justificaria um VOT mais curto nas oclusivas que as precedem. Inversamente, vogais altas são inerentemente mais breves, o que levaria a um VOT mais longo.

Finalmente, é necessário levar em conta a possível interação entre a altura da vogal e o ponto de articulação. Logo, analisaremos como cada consoante em separado se comporta com relação à duração de VOT exibida diante de vogais altas e baixas. Procuraremos observar se a maneira como a altura da vogal afeta o VOT se dá igualmente para as oclusivas sonoras, apesar da inviabilidade da condução de análise estatística para esses dados.

Pergunta 4: A posição ântero-posterior da vogal seguinte interfere nos índices de VOT das oclusivas por nativos de inglês e falantes de PB?

H6: A duração de VOT das oclusivas surdas diante da vogal anterior [ɪ] será significativamente maior do que diante da vogal posterior [ʊ] na produção dos grupos N, P e A, mesmo ao considerar os grupos de articulação em separado.

Alves e Dias (2010) constataram valores de VOT significativamente maiores para as oclusivas diante de [ɪ] do que diante de [ʊ], o que sugere que a posição anterior propicia a maior duração do VOT. Já os resultados encontrados por Schwartzhaupt (2012) sugerem VOT mais longo em bilabiais diante de vogais altas posteriores, enquanto que para oclusivas alveolares e velares é a vogal anterior que parece causar aumento no VOT. Tais resultados podem indicar que o ponto de articulação deve ser considerado na investigação da influência da posição ântero-posterior na duração do VOT.

Diante disso, optaremos pelo mesmo procedimento adotado para a altura da vogal seguinte, ou seja, analisaremos como cada consoante em separado se comporta com relação à duração de VOT exibida diante de vogais anteriores e posteriores.

Procuraremos ainda observar se as oclusivas sonoras são afetadas da mesma maneira que as surdas quanto a posição da vogal.

⁴⁵ DELGADO MARTINS, M.R. *Sept études sur la perception*. Lisboa: Instituto Nacional de Investigação Científica, 1986.

⁴⁶ MORAES, J.A. de. *Acentuação lexical e acentual frasal em português: um estudo acústico-perceptivo*. In: II Encontro Nacional de Fonética e Fonologia. Brasília, 1986.

3.2 O CORPUS

Cabe inicialmente esclarecer que o *corpus* utilizado para o estudo-piloto mostrou-se adequado para a consecução do experimento completo e que, portanto, o que segue é uma descrição única, que vale para ambas as etapas desta pesquisa. Optou-se pela constituição de um *corpus ad hoc*, o que significa dizer que não foi constituído de fala espontânea. Segundo Llisterri (1991, p. 68), os dados dos *corpora* falados podem ser divididos entre:

os que pertencem ao âmbito da fala espontânea e os que são preparados *ad-hoc* para o estudo de um fenômeno determinado. Os primeiros são geralmente obtidos através da realização de gravações de um ou vários informantes falando de maneira mais ou menos livre durante um período de tempo relativamente longo, os segundos são normalmente lidos por informantes selecionados em condições precisas e bem controladas.⁴⁷

A coleta de dados por meio da gravação de fala espontânea ficou fora de questão por consumir muito tempo e não trazer garantias de um número suficiente de dados para análise. Seja frisado, no entanto, que nossa opção pela chamada ‘fala de laboratório’⁴⁸ apresenta a vantagem de permitir maior controle sobre as variáveis envolvidas no fenômeno estudado, o que faz com que o teste de uma hipótese se faça de maneira mais rigorosa (XU, 2010). Para o *corpus* deste estudo, escolhemos vocábulos nos quais as oclusivas surdas [p], [t] e [k] e sonoras [b], [d] e [g] estão em posição inicial de palavra. Selecionamos exclusivamente monossílabos, uma vez que o número de sílabas da palavra-alvo tem-se mostrado um fator que causa variação nos índices de VOT das oclusivas iniciais (LISKER; ABRAMSON, 1967; KLATT, 1975; OHALA, 1981; ALVES, 2011) Cada palavra-alvo foi inserida em posição medial na frase-veículo “Say _____ to me”, a fim de que o ambiente fonético no qual se encontre o segmento possa ser controlado e possamos evitar o chamado efeito de lista causado pela leitura de palavras isoladas (LLISTERRY, 1991, p. 73).

As palavras constituíram pares mínimos do tipo *cap* [kæp] e *gap* [gæp]. Foram selecionadas vogais iguais ou semelhantes para suceder os pares surda/sonora e se buscou contemplar a maior variedade possível com relação à qualidade vocálica. Assim, cada oclusiva foi seguida de três vogais altas [i], [ɪ] e [u ou ʊ], duas médias [ɛ] e [ʌ], e duas baixas

⁴⁷ “los que pertenecen al ámbito del habla espontánea y los que se han preparado ad-hoc para el estudio de un fenómeno determinado. Como veremos más adelante, mientras que los primeros suelen obtenerse realizando grabaciones de uno o varios informantes hablando de manera más o menos libre durante un período de tiempo generalmente largo, los segundos son normalmente leídos por informantes seleccionados em condiciones precisas y bien controladas.”

⁴⁸ Xu (2010) define fala de laboratório como aquela ‘gravada em um laboratório, normalmente por meio da leitura em voz alta de *scripts* elaborados previamente. (Minha tradução)

[æ] e [ɑ]. Esta última pode vir a ser realizada como [ɔ], especialmente pelos sujeitos brasileiros. Dessa forma, esperamos equilibrar a influência que a qualidade da vogal venha a exercer nos dados, bem como obter material suficiente para testar tal influência. Ressalte-se que, embora tenhamos optado por utilizar apenas palavras reais, tivemos de lançar mão de vocábulos cuja frequência na língua é baixa, como por exemplo, *teed*, *git* e *dook*. Alguns dos informantes norte-americanos, inclusive, afirmaram desconhecer tais palavras. No entanto, sua inclusão justifica-se pela dificuldade em encontrar palavras reais que se encaixassem dentro do padrão de monossílabos iniciados por oclusivas surdas e sonoras que constituíssem pares mínimos. O *corpus* utilizado no estudo-piloto e que se manteve para o experimento completo segue na Tabela 14.

Vogal seguinte	Bilabiais	Alveolares	Velares
[i]	peach/beach	teed/deed	keek/geek
[ɪ]	pig/big	tick/dick	kit/git
[u ou ʊ]	poot/boot	took/dook	cook/gook
[ɛ]	pet/bet	ted/dead	ket/get
[ʌ]	puff/buff	tuck/duck	cut/gut
[æ]	pack/back	tap/dap	cap/gap
[ɑ]	pop/bop	top/dop	cosh/gosh

Tabela 14: *Corpus* utilizado no presente estudo

3.3 OS SUJEITOS

Os dados desta pesquisa foram coletados junto a sujeitos pertencentes a três grupos diferentes: aprendizes de inglês falantes de PB (grupo A), falantes proficientes de inglês cuja L1 é o PB (grupo P) e falantes nativos de inglês (grupo N). Todos alegaram não possuir quaisquer problemas de fala ou audição.

O experimento completo contou com 18 informantes, 6 de cada grupo. O número de sujeitos do sexo feminino e masculino foi o mesmo dentro de cada um dos grupos. Sendo a idade outro fator de interferência (CELESTE; TEIXEIRA, 2009), optamos por coletar dados junto a sujeitos com até 44 anos.

O grupo N constituiu-se de 6 falantes nativos de inglês. São todos originários dos Estados Unidos e, à época da coleta, estudavam português como língua estrangeira no CELIN e encontravam-se há menos de um ano no país. Muito embora reconheçamos que o ideal seria gravar a produção de sujeitos que não tivessem contato algum com o PB, a impossibilidade de

realizar tais gravações motivou a utilização do procedimento aqui adotado. Importante mencionar ainda que gostaríamos de ter encontrado sujeitos que falassem o mesmo dialeto de inglês, o que não foi possível. Tomou-se o cuidado, entretanto, de selecionar sujeitos em estágio inicial de aprendizado da língua portuguesa. Suas idades variaram entre 22 e 44 anos. O quadro 1 mostra os seguintes dados de todos os informantes do grupo N: sexo, idade, naturalidade, tempo de estada no Brasil e nível de português no CELIN.

Informante	Sexo	Idade	Naturalidade	Tempo de estada no Brasil	Nível de português no CELIN
N1	M	32	Tallahassee, Flórida (EUA)	8 meses	básico
N2	F	35	Los Angeles, Califórnia (EUA)	3 meses	básico
N3	M	36	Bar Harbor, Maine (EUA)	5 meses	básico
N4	F	22	Richmond, Virgínia (EUA)	3 meses e meio	básico
N5	M	22	Chester, South Carolina (EUA)	2 meses	básico
N6	F	44	Boise, Idaho (EUA)	6 meses	básico

Quadro 1: Perfil dos informantes do grupo N.

Os 6 brasileiros falantes proficientes de inglês são todos formados em Letras Inglês pela Universidade Federal do Paraná e, com exceção da informante P6, são ou foram professores de língua inglesa no CELIN. À época da coleta, suas idades variavam entre 25 e 34 anos. Além de serem formados em Letras Inglês, ensinavam a língua em institutos de idiomas há pelo menos 5 anos, predicados que a pesquisadora julgou inicialmente suficientes para atestar a proficiência desses sujeitos na língua. No entanto, após o exame de qualificação, sugeriu-se que os professores selecionados tivessem passado por exames internacionais de proficiência, a fim de que sua competência na língua fosse mais claramente comprovada. O ideal seria que todos os sujeitos tivessem passado pelo mesmo exame de proficiência, mas não conseguimos obter tal uniformidade. Além disso, embora todos os sujeitos tenham sido submetidos a testes de proficiência, alguns dos certificados não correspondiam ao atual nível de inglês do sujeito por terem sido obtidos muito anteriormente à data da coleta. De toda forma, será reportado apenas o nível comprovado de inglês do sujeito. A fim de uniformizar o nível obtido pelos diferentes exames optamos por fazer a correspondência dos resultados por

meio do Quadro Comum Europeu de Referência para Línguas, que elenca as seguintes categorias de usuários de uma língua estrangeira em ordem crescente de proficiência: A1 e A2 (usuário elementar); B1 e B2 (usuário independente) e C1 e C2 (usuário proficiente). Um dos sujeitos desta pesquisa possui o nível C2, três sujeitos têm o nível C1 e dois sujeitos têm comprovação do nível B2.

O quadro 2 apresenta os seguintes dados em relação ao grupo P: sexo, idade, naturalidade, tempo que leciona inglês, testes ou certificados que comprovam nível de inglês e sua correspondência no Quadro Comum Europeu de Referência para Línguas.

Informante	Sexo	Idade	Naturalidade	Tempo que leciona inglês	Permanência em país de língua inglesa (tempo)	Teste/Certificado Internacional	Correspondência no Quadro Comum Europeu
P1	M	25	Curitiba, PR	5 anos	Sim (3 meses)	TOEFL ITP (score 667)	C1
P2	F	25	São Sebastião, SP	7 anos	Sim (1 ano)	FCE	B2
P3	M	34	Curitiba, PR	14 anos	Não	TOEFL ITP (score 630)	C1
P4	F	29	Curitiba, PR	7 anos	Sim (18 meses)	Teste da Universidade de Leipzig, aplicado no CELIN	B2
P5	M	30	Terra Roxa, PR	12 anos	Sim (7 meses)	TOEFL (score 613)	C1
P6	F	29	São Paulo, SP	11 anos	Sim (1 mês)	CPE	C2

Quadro 2: Perfil dos informantes do grupo P.

Finalmente, o grupo dos aprendizes foi composto por 6 alunos adultos do Centro de Línguas e Interculturalidade (CELIN) da Universidade Federal do Paraná. Como todos relataram ter estudado inglês nessa instituição desde o nível inicial, é possível estimar que possuíam entre 300 e 400 horas de instrução da língua-alvo. Voltamos a frisar que esse dado é mais confiável do que a denominação pré-intermediário ou intermediário, que varia de acordo com a instituição em que se estuda. Nenhum dos sujeitos havia visitado ou morado em país onde se fala inglês como língua nativa e, portanto, o contato do grupo A com o inglês restringia-se basicamente às aulas de língua no CELIN. O quadro 3 apresenta os seguintes

dados dos sujeitos do grupo A: sexo, idade, naturalidade, horas de instrução em língua inglesa, permanência em país de língua inglesa.

Informante	Sexo	Idade	Naturalidade	Horas de Instrução	Permanência em país de língua inglesa
A1	M	24 anos	Ibaiti, PR	300	Não
A2	F	31 anos	Curitiba, PR	300	Não
A3	M	30 anos	Curitiba, PR	400	Não
A4	F	23 anos	Curitiba, PR	400	Não
A5	M	26 anos	Curitiba, PR	400	Não
A6	F	25 anos	Curitiba, PR	400	Não

Quadro 3: Perfil dos informantes do grupo A.

Os dois primeiros informantes de cada grupo fizeram parte do experimento piloto e, como não houve problemas de qualquer natureza envolvendo a coleta de dados com esses sujeitos, seus dados também foram utilizados para o experimento completo.

Para o estudo-piloto, portanto, foi selecionado um par de informantes de cada grupo (N1, N2, P1, P2, A1 e A2). Para cada par, havia um informante do sexo feminino e um do sexo masculino, uma vez que é notória a interferência da variável gênero no aspecto aqui pesquisado (FLEGE, 1991; WHITESIDE; HENRY; DOBBIN, 2004). A letra designa o grupo a que pertencem e o número diferencia o sujeito dentro do grupo. Optamos por utilizar os números ímpares para os sujeitos do sexo masculino e os números pares, para o sexo feminino.

Os aprendizes brasileiros são ambos paranaenses, sendo A1 um informante do sexo masculino, natural de Ibaiti, com 24 anos de idade e A2 uma informante do sexo feminino, natural de Curitiba, com 31 anos de idade. Na ocasião da coleta, ambos estudavam inglês no Centro de Línguas e Interculturalidade (CELIN) da Universidade Federal do Paraná, no nível denominado pré-intermediário. Como ambos relataram ter estudado inglês nessa instituição desde o nível inicial, é possível estimar que contavam com aproximadamente 300 horas de instrução da língua-alvo, o que acaba por ser um dado mais confiável do que a denominação pré-intermediário, que varia de acordo com a instituição em que se estuda. Nenhum deles havia visitado país no qual o inglês é falado como língua nativa.

Os brasileiros falantes proficientes de inglês (P1 e P2) são formados em Letras Inglês pela Universidade Federal do Paraná e são professores de língua inglesa no CELIN. Ambos tinham 25 anos de idade à época da coleta. O informante P1 é natural de Curitiba e, à época, lecionava inglês há 5 anos. Esteve em país onde o inglês é falado como língua nativa pelo

período de 3 meses. A informante P2 é natural de São Sebastião, em São Paulo, e lecionava inglês há 7 anos. Esteve em país onde o inglês é falado como língua nativa pelo período de um ano.

Os informantes nativos de inglês são ambos dos Estados Unidos. O informante N1 é natural de Tallahassee, na Flórida, e a informante N2 é de Los Angeles, na Califórnia. À época da coleta, a idade dos informantes era de 32 e 25 anos, respectivamente. Ambos estavam no Brasil há menos de um ano e estudavam português como língua estrangeira em nível básico.

3.4 OS PROCEDIMENTOS DE COLETA

Primeiramente, as gravações dos sujeitos ocorreram no Laboratório de Estudos Fônicos da UFPR (LEFON), que conta com cabine com tratamento acústico. Para a captação dos dados, foi utilizado um microfone Shure KSM 27 e o programa *Audacity*. Os dados foram capturados a uma taxa de amostragem de 44100Hz. Posteriormente, devido a problemas técnicos no LEFON, as gravações finais tiveram de ser feitas no laboratório da Universidade Tecnológica Federal do Paraná, onde também há cabine com tratamento acústico. Nesse local, também foi utilizado o programa *Audacity*, sendo os dados também capturados a uma taxa de amostragem de 44100Hz.

Inicialmente, a pesquisadora explicou brevemente, na língua nativa do sujeito, como se daria a gravação e solicitou a leitura das instruções e o preenchimento das informações pessoais em uma folha que lhe foi entregue (Apêndices A, B e C). Em seguida, foi feito um pequeno teste para que o sujeito se familiarizasse com a maneira como se conduziria a gravação. Por fim, deu-se a gravação completa com pequenos intervalos entre as 5 repetições, a fim de que o processo não se tornasse muito cansativo. Os sujeitos não foram informados dos objetivos da pesquisa.

A cada repetição, os participantes liam 42 frases-veículo, que diferiam apenas com relação à palavra-alvo. As frases foram apresentadas para a leitura com um intervalo fixo de 4 segundos entre uma e outra, o que foi possível com a utilização do programa *PowerPoint* para a apresentação. O intervalo de 4 segundos foi estipulado pela pesquisadora por considerá-lo um tempo confortável para a transição na leitura das frases, o que foi comprovado durante as gravações. A cada repetição, a ordem das sentenças era diferente, tendo sido pré-aleatorizada por meio do programa *PowerPoint*. Ao final da gravação, feitas 5 repetições das 42 frases,

cada participante havia contribuído com 210 frases. Quando o sujeito precisava repetir a frase, ele não o fazia imediatamente, mas no final da seção. Tal procedimento havia sido esclarecido de antemão, de modo que o ritmo da gravação não era interrompido e a pesquisadora simplesmente anotava a frase para posterior repetição. Assim, o número de *tokens* não foi prejudicado e manteve-se o mesmo para cada palavra-alvo.

Como observado no *corpus* da tabela 14, na seção 3.2, o número de palavras-alvo iniciadas com consoantes surdas e sonoras é o mesmo, ou seja, dos 210 *tokens* gravados por cada informante, 105 deles são iniciados por consoantes surdas e 105 por sonoras. Sendo 6 o número de participantes em cada grupo, tivemos um total de 1260 *tokens* por grupo, perfazendo um total de 3780 dados analisados. Nenhum dado precisou ser descartado, já que a pesquisadora esteve monitorando cada uma das gravações e quando houve algum problema, foi pedido ao sujeito que relese a sentença no final da seção.

3.5 A ANÁLISE DOS DADOS

A análise dos dados foi feita por intermédio do *software* Praat (Boersma e Weenick, 2012), versão 5.3.03. Após recortarmos todas as sentenças através do programa mencionado, foi medida a duração absoluta do VOT de cada consoante oclusiva inicial, bem como a duração da frase-veículo. Houve, inicialmente, a tentativa de se calcular a duração relativa levando em consideração a palavra-alvo. Entretanto, tal método se mostrou ineficaz, pois em certos dados era impossível precisar o limite entre o último segmento da palavra-alvo e a oclusão do segmento seguinte. A Figura 7, a seguir, apresenta a medição quando feita de maneira inequívoca, ao passo que a Figura 8 atesta a ocasional dificuldade de medição.

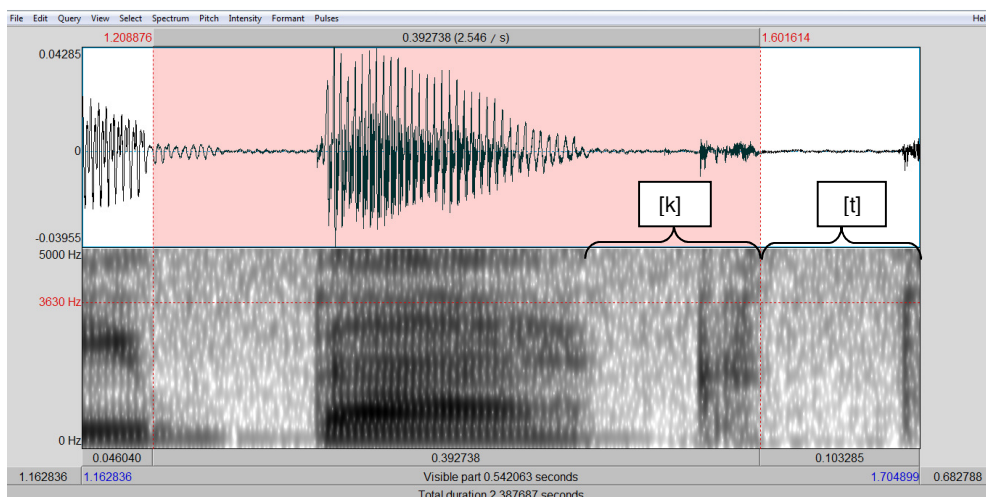


Figura 7: Oscilograma e espectrograma da palavra *back*. Entre as linhas pontilhadas, destaca-se a duração da palavra-alvo. Logo em seguida, é possível divisar a oclusão proveniente do [t] na expressão to me que completa a sentença-veículo. Fonte: *corpus* da autora.

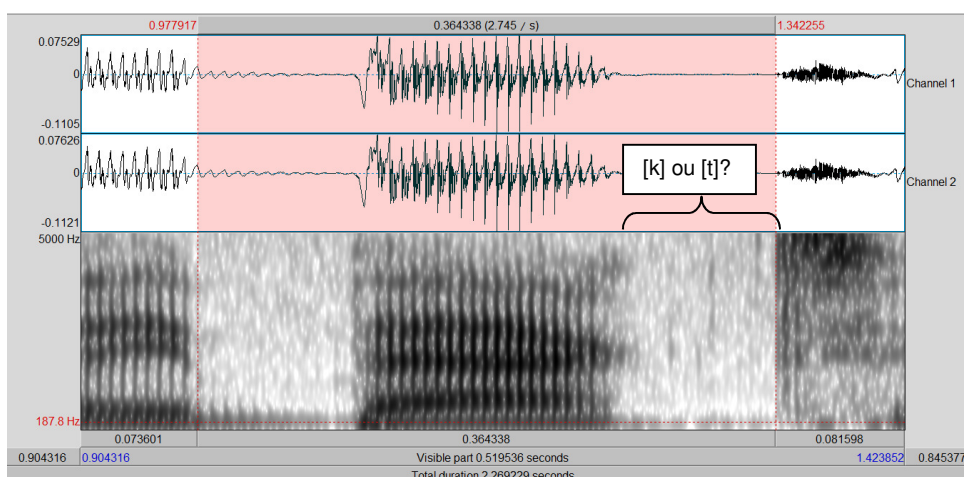


Figura 8: Oscilograma e espectrograma da palavra *back*. Note que não há como saber o limite entre a duração da palavra-alvo e a oclusão proveniente do [t] em to me. Fonte: *corpus* da autora.

A extração do VOT, como já observamos na seção 1.2.2, pode ser feita mediante a utilização de diferentes critérios. É importante, entretanto, que fique clara a necessidade de seguir um único padrão para a extração da medida no decorrer de toda a pesquisa. Optamos aqui pelos seguintes parâmetros:

- (a) para oclusivas surdas, foi coletada a duração em milissegundos entre o *burst* e o primeiro pulso regular da vogal seguinte. Na presença de *bursts* múltiplos, a medição ocorreu a partir do último *burst* da sequência (Fig. 9);

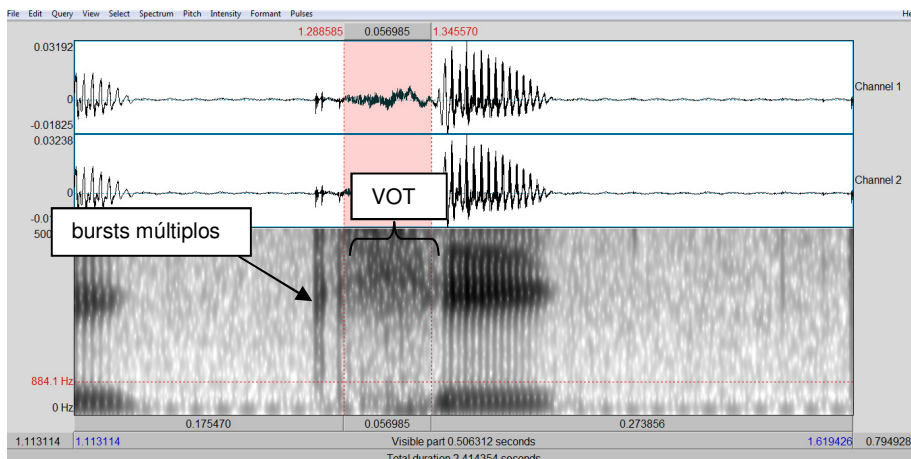


Figura 9: Oscilograma e espectrograma da palavra *keek*. A medição de VOT começa a partir do último burst da sequência. Fonte: *corpus* da autora.

(b) para oclusivas sonoras, mediu-se o segmento compreendido entre o início da barra de sonoridade presente na oclusão até o registro do *burst*. Normalmente, deparamo-nos com uma das seguintes situações: a barra de sonoridade ocupa todo o período da oclusão (Fig. 10); a barra de sonoridade pertencente ao segmento encontra-se imediatamente anterior ao *burst* (Fig.11) ou simplesmente não há barra de sonoridade durante a oclusão e o VOT é marcado como zero ou de intervalo curto (Fig. 12).

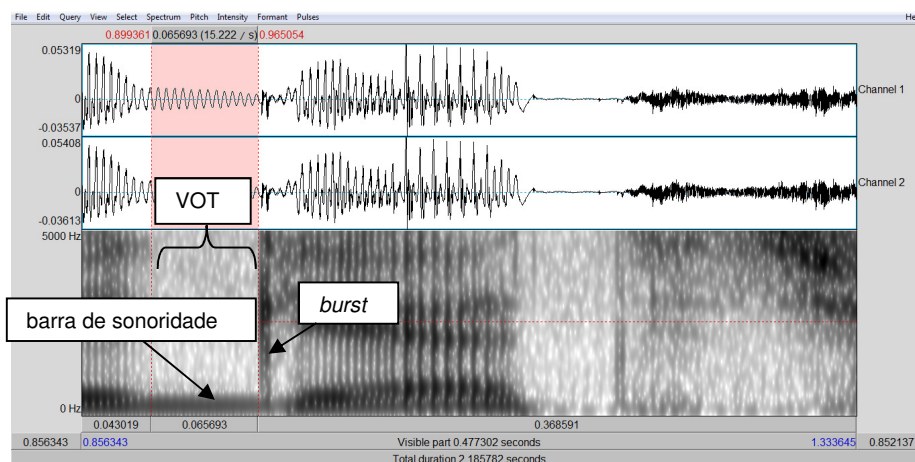


Figura 10 Oscilograma e espectrograma da palavra *get*. A medição de VOT começa a partir do início da barra de sonoridade até o registro do burst (-66ms). Fonte: *corpus* da autora.

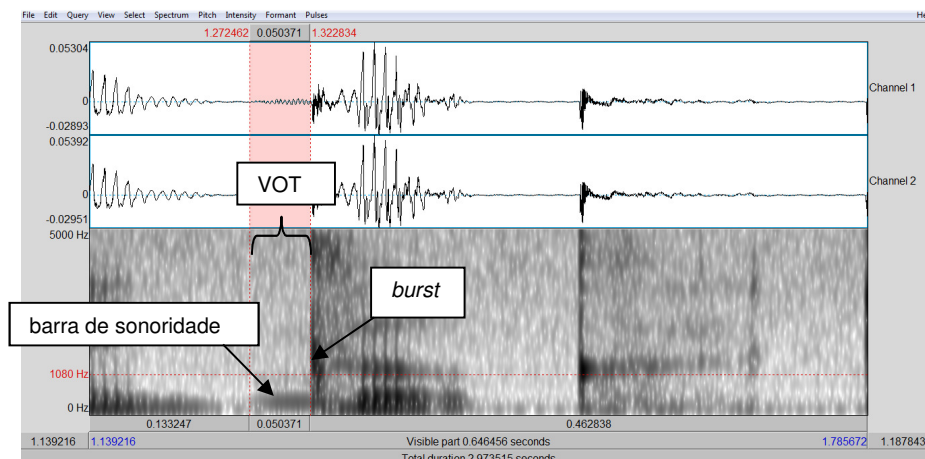


Figura 11: Oscilograma e espectrograma da palavra *gook*. A medição de VOT começa a partir do início da barra de sonoridade até o registro do burst (50ms). Fonte: *corpus* da autora.

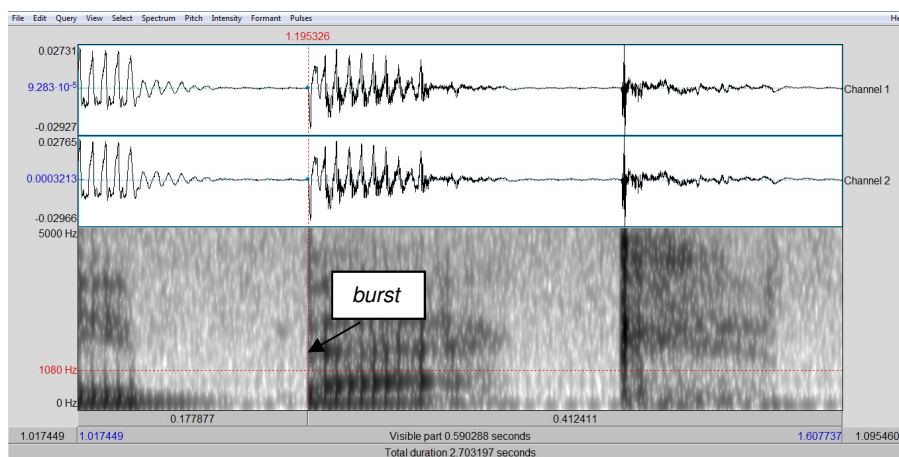


Figura 12: Oscilograma e espectrograma da palavra *buff*. Não há barra de sonoridade imediatamente anterior ao burst (VOT=zero). Fonte: *corpus* da autora.

Depois de medida a duração absoluta do VOT presente nos segmentos, foi calculada a duração relativa do VOT, ou seja, a porcentagem de sua duração em relação à frase-veículo. Foi essa, inclusive, nossa opção para a apresentação dos dados, uma vez que acreditamos ser a duração relativa mais acurada em relação à duração absoluta, pois corrige quaisquer vieses que venham a existir por conta da taxa de elocução dos informantes.

3.6 A ANÁLISE ESTATÍSTICA

Para a análise estatística dos dados, o programa utilizado foi o SPSS 18.0 (SPSS, Chicago, IL). Devido ao número reduzido de sujeitos em cada grupo, a presunção de

normalidade de distribuição não pode ser testada. Portanto, todos os testes utilizados são não paramétricos. O nível de significância adotado para todos foi de $\alpha = 0.05$.

Para comparar a produção das oclusivas por ponto de articulação dentro de um mesmo grupo, ou seja, analisar se havia diferença significativa na produção de [p], [t] e [k] em cada um dos grupos, foi utilizado o teste de Friedman. Esse teste compara dados contínuos com três ou mais opções de resposta e é uma alternativa aos testes paramétricos quando estamos diante de um design intrassujeitos.

Havendo diferença de produção, aplicou-se o teste de Wilcoxon para detectar onde estava a diferença, se entre [p] e [t], [p] e [k] ou [t] e [k]. Tal teste é também utilizado para a comparação de dados contínuos quando temos um *design* intrassujeitos, porém apenas dois parâmetros relacionados.

Finalmente, para comparar a produção dos grupos entre si, o teste de Mann-Whitney foi aplicado entre diferentes pares, ou seja, grupo N e A, grupo N e P, grupo P e A. Mann-Whitney é a alternativa não paramétrica para analisar dados contínuos em dois grupos independentes, ou seja, quando temos um design intersujeitos.

3.7 CONCLUSÃO

Como visto, em nossa pesquisa tencionamos analisar a produção de oclusivas iniciais do inglês por falantes de PB em contraste com a produção de falantes nativos de inglês. Nosso intuito será essencialmente analisar a influência do nível de proficiência na produção dos falantes de PB, bem como a interferência dos seguintes fatores na produção das oclusivas: ponto de articulação, altura da vogal seguinte e posição ântero-posterior da vogal seguinte.

Cabe reiterar nossa opção pela utilização de um *corpus ad hoc* especialmente desenvolvido para este estudo, em detrimento de fala espontânea, uma vez que o primeiro nos possibilita maior controle sobre as variáveis estudadas e garantia de que os dados necessários para a análise serão produzidos. É importante ainda insistir no fato de que, no decorrer deste trabalho, os resultados serão apresentados em sua duração relativa, ou seja, a porcentagem que o VOT ocupa em relação à frase-veículo. Muito embora a maioria dos trabalhos da área apresente seus resultados em duração absoluta, inclusive os aqui resenhados no capítulo 2, é preciso que se transgrida tal padrão, uma vez que a duração relativa nos permite corrigir os possíveis vieses concernentes à taxa de elocução dos informantes. Sabemos, entretanto, que

alguns pesquisadores estarão interessados em ter acesso a nossos resultados em sua duração absoluta, inclusive para fins de comparação com estudos já feitos, e, portanto, tais resultados estarão disponibilizados nos apêndices deste trabalho.

4. O ESTUDO-PILOTO

Com o objetivo de testar a metodologia descrita no Capítulo 2, um estudo-piloto foi conduzido entre novembro de 2011 e março de 2013. Os dados utilizados neste piloto foram coletados junto a sujeitos representantes de três grupos distintos: dois brasileiros aprendizes de inglês como língua estrangeira (A1 e A2), dois brasileiros falantes proficientes de inglês como língua estrangeira (P1 e P2) e dois falantes nativos de inglês norte-americano (N1 e N2), que aqui representam o grupo-controle. Seus perfis encontram-se descritos na seção 3.3.

O *corpus* utilizado está descrito na seção 3.2 e os dados foram gravados no Laboratório de Estudos Fônicos da UFPR (LEFON), que é acusticamente tratado. A análise dos dados foi feita por intermédio do software *Praat*. (Boersma e Weenick, 2012). Todos os dados do piloto serão apresentados em sua duração relativa, o que quer dizer que tomamos a porcentagem que a duração absoluta do VOT ocupa na sentença-veículo. Tal medida objetiva neutralizar quaisquer vieses que venham a ocorrer devido à taxa de elocução dos informantes. Mais detalhes sobre essas escolhas encontram-se na seção 2.5.

Feitas essas observações, passemos a reportar os resultados do estudo-piloto. Para sua apresentação, optamos por organizá-los levando em conta a influência das seguintes variáveis na duração do VOT de oclusivas surdas e sonoras: o nível de proficiência do falante, o ponto de articulação da consoante, a altura da vogal seguinte e a posição ântero-posterior da vogal seguinte.

4.1 VOT E NÍVEL DE PROFICIÊNCIA

Os resultados do piloto corroboram pesquisas anteriores que atestam que falantes nativos de PB produzem as oclusivas surdas iniciais do inglês com níveis de aspiração inferiores aos de um falante nativo da língua (MAJOR, 1992; SANCIER; FOWLER, 1997; ROCCA, 2003; ZIMMER, 2004; COHEN, 2004; REIS; NOBRE-OLIVEIRA, 2007; ALVES, 2011; FRANÇA, 2011; SCHWARTZHAUPT, 2012), ao mesmo tempo em que apontam para uma associação positiva entre o grau de proficiência do aprendiz e sua capacidade crescente em aproximar-se dos sons da língua-alvo (ZIMMER, 2004; FRANÇA, 2011). A Tabela 15 apresenta os resultados encontrados no piloto quanto à duração relativa de VOT exibida pelos três grupos analisados.

Oclusivas surdas			
	Nativos	Proficientes	Aprendizes
[p]	4,66%	2,29%	0,70%
[t]	5,39%	3,76%	1,31%
[k]	5,19%	4,74%	2,43%
Média	5,04%	3,60%	1,48%

Tabela 15: Média da duração relativa de VOT de consoantes oclusivas surdas para nativos, falantes proficientes e aprendizes.

Com relação à média geral, a duração de VOT das oclusivas surdas produzidas pelos falantes proficientes foi 59% mais longa em comparação aos aprendizes. Logo, os resultados preliminares deste piloto parecem indicar a existência de um refinamento de pronúncia quanto à produção das oclusivas surdas conforme o domínio da língua se consolida. No gráfico 1 são apresentados os resultados por informante e por grupo.

(a) Duração relativa de VOT por informante.

(b) Duração relativa de VOT por grupo.

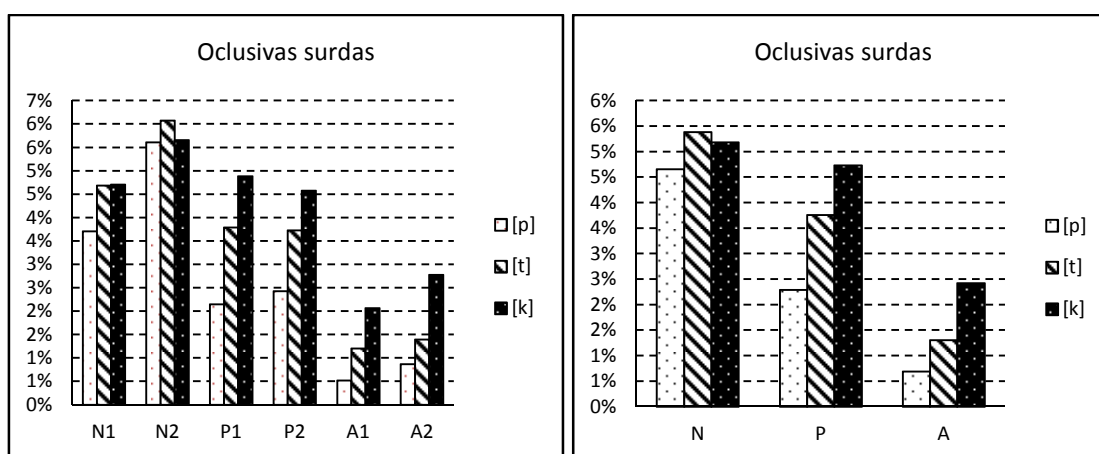


Gráfico 1: Duração relativa de VOT das consoantes oclusivas surdas por informante e por grupo.

No que tange às sonoras, os dados dos falantes nativos não exibiram um padrão uniforme na duração do VOT das oclusivas, apresentando características diferentes entre informantes e mesmo nos dados de um único informante. O informante N1, por exemplo, exibiu basicamente dois tipos de realização de VOT para as oclusivas sonoras: em 46% dos dados, suas oclusivas tiveram VOT com vozeamento total, ou seja, a barra de vozeamento apareceu em todo o período de oclusão até o momento do *burst*, possivelmente devido ao contexto intervocálico propiciado pela frase-veículo “Say _____ to me”; e em 43% dos dados, o VOT foi realizado com intervalo curto, ou seja, valores positivos próximos a zero. Quanto aos demais dados desse informante, apurou-se VOT parcialmente vozeado em 3% dos

dados, com valores em torno de 40ms; e VOT zero em 8% dos dados. Já a informante N2 não apresentou tamanha variação, sendo que apenas um de seus dados exibe VOT zero e os demais foram realizados com VOT de intervalo curto (*short lag*). Nestes, a duração do VOT teve em média 14,54ms. Observamos, portanto, que a quase totalidade dos dados dessa informante se assemelhou muito à realização das oclusivas surdas por falantes de PB, já que não houve pré-vozeamento e o VOT pode ser classificado como levemente aspirado.

Os falantes proficientes também apresentaram realizações variadas na produção das oclusivas sonoras. Ambos, assim como o informante N1, exibiram oclusivas cujo VOT variou essencialmente entre intervalo curto e vozeamento total durante a oclusão. Com relação aos aprendizes, todos os dados de A1 foram realizados com vozeamento total durante a oclusão, bem como 93% dos dados de A2, sendo os demais realizados com VOT zero. Nenhum dos dados dos aprendizes apresentou VOT de intervalo curto. O Gráfico 2 traz um resumo das diferentes realizações de consoantes sonoras para cada informante.

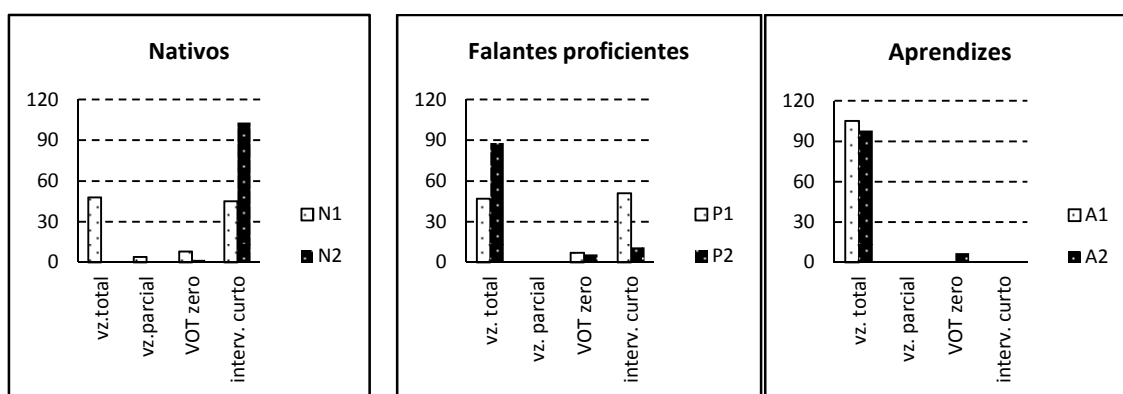


Gráfico 2: Diferentes realizações das oclusivas sonoras por informante no estudo-piloto.

Analisando os resultados do piloto, também se pode sugerir refinamento na produção dos falantes de PB quanto à produção das oclusivas sonoras conforme o domínio da língua se consolida, devido à diminuição de dados com VOT de vozeamento total, característica mais persistente na produção de falantes de PB. Enquanto os aprendizes exibem 97% dos dados com oclusivas cujo VOT apresenta vozeamento total e 3% com VOT zero, a produção dos falantes proficientes se divide entre oclusivas com vozeamento total (64%), VOT zero (6%) e intervalo curto (30%), o que se aproxima mais da diversidade apresentada pelos informantes nativos que, de forma geral, apresentam 23% das oclusivas produzidas com VOT de vozeamento total, 2% de vozeamento parcial, 5% de VOT zero e 70% das oclusivas sonoras com VOT de intervalo curto.

Diante das diferentes realizações das oclusivas sonoras pelos informantes do piloto, a saída encontrada para analisar esses dados é adotar o sistema utilizado por Lisker e Abramson (1964) de apresentar os dados de pré-vozeamento e intervalo curto separadamente. Se assim o fizermos, teremos o panorama exibido no Gráfico 3.

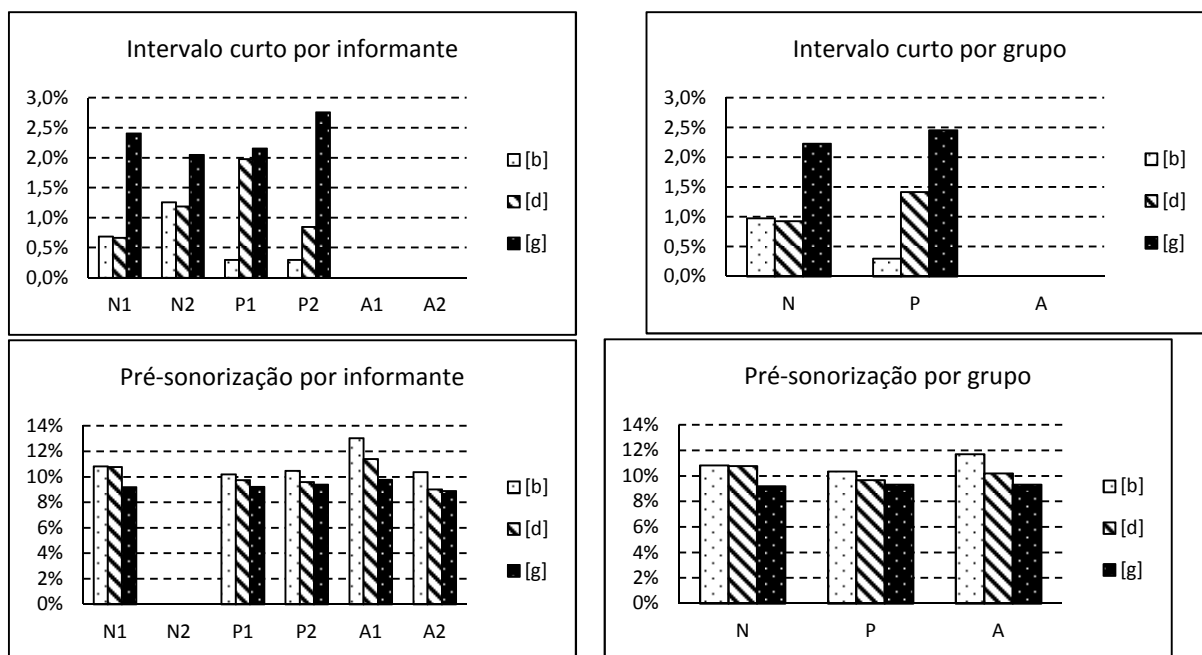


Gráfico 3: Duração relativa de VOT de consoantes oclusivas sonoras por informante e por grupo.

A desvantagem de separarmos as oclusivas sonoras em dois grupos distintos é a impossibilidade de tratamento estatístico para os dados no experimento, uma vez que haverá lacunas na realização das consoantes para ambos os grupos e o número de sujeitos será ainda mais reduzido. Assim, no experimento, tomamos a decisão de reportar os resultados sem conduzirmos análise estatística para as oclusivas sonoras. Na tabela 16, temos os dados de cada um dos três grupos, separados por ponto de articulação e levando em conta as duas realizações diferentes de oclusivas sonoras.

	Nativos		Falantes proficientes		Aprendizes	
	Int.curto	Pré-voz	Int.curto	Pré-voz	Int.curto	Pré-voz
[b]	0,98%	10,85%	0,3%	10,36%	0%	11,72%
[d]	0,93%	10,8%	1,42%	9,68%	0%	10,23%
[g]	2,23%	9,21%	2,46%	9,33%	0%	9,35%

Tabela 16: Média da duração relativa de VOT de consoantes oclusivas sonoras para nativos, falantes proficientes e aprendizes.

4.2 VOT E O PONTO DE ARTICULAÇÃO

Com relação às oclusivas surdas, tem sido atestado (LISKER ;ABRAMSON, 1964; CHO; LADEFOGED, 1999; KLEIN, 1999; COHEN, 2004; SCHWARTZHAUPT, 2012) que os índices de VOT tendem a aumentar quanto mais posterior é o ponto de articulação da consoante.

Para os dados deste estudo-piloto, os grupos P e A exibiram aumento progressivo na duração do VOT conforme o ponto de articulação se tornava mais posterior. Entretanto, essa relação não se confirmou para o grupo N, no qual a alveolar superou a velar em duração de VOT.

Assim, os dados analisados no piloto parecem corroborar tal influência do ponto de articulação sobre duração do VOT das oclusivas surdas, embora não completamente. Devemos lembrar, no entanto, que os resultados deste piloto referem-se aos dados de um número reduzido de sujeitos e, além disso, não foram submetidos à quantificação via análise estatística, o que faz com que quaisquer considerações aqui feitas devam ser relativizadas. Testes estatísticos serão realizados no experimento para observar se há diferença significativa na duração de VOT de [p], [t] e [k] para os grupos N, P e A.

Quanto às oclusivas sonoras, optamos por considerar as duas realizações em separado. Nos dados de intervalo curto, observamos que a duração relativa de VOT nas oclusivas velares é consideravelmente mais longa do que nos demais pontos. Isso ocorre para todos os informantes exceto para P1, que apresenta VOT quase tão longo no ponto alveolar (1,98% contra 2,16%). Além disso, vemos que, para N1 e N2, a duração relativa de VOT no ponto bilabial é apenas levemente superior àquela encontrada no ponto alveolar. Vale lembrar que os aprendizes não apresentam dados de oclusivas sonoras com VOT de intervalo curto.

Nos dados de pré-sonorização, todos os informantes exibem valores progressivamente menores de VOT, quanto mais posterior é o ponto de articulação. Em outras palavras, a bilabial exhibe VOT maior que a alveolar, que exhibe VOT maior que a velar. Isso se dá de maneira mais acentuada com os aprendizes e caracteriza justamente o oposto do que ocorre quando analisamos as consoantes surdas dos grupos P e A.

4.3 VOT E ALTURA DA VOGAL SEGUINTE

Há indícios de que, na realização de oclusivas surdas, os índices de VOT alcancem maior duração quando a oclusiva encontra-se diante de vogais altas (KLATT, 1975; YAVAS, 2008; FRANÇA, 2011).

Analisando os dados do piloto, nota-se que, entre os nativos, a hipótese de que a oclusiva diante de vogal alta terá o VOT mais longo cai por terra. O informante N1 exibiu oclusivas com VOT relativo mais longo diante de [æ] (5,31%) do que diante de [i] (4,30%), ao passo que a diferença no VOT das oclusivas para a informante N2 foi mínima: 5,96% diante de [æ] e 6,05% diante de [i]. Entre os falantes proficientes, também não houve comportamento uniforme. Enquanto o informante P1 apresentou oclusivas com VOT consideravelmente mais longo diante de [i], 4,73% contra 3,42%; a diferença para P2 foi mínima e o VOT mais longo ocorre diante de [æ], 3,72% contra 3,53%. Apenas entre os aprendizes observamos um comportamento uniforme: para ambos o VOT foi mais longo diante de [i]. Para A1 tivemos 1,33% contra 1,01% e para A2 a diferença foi ainda maior, de 1,96% para 0,76%. Em suma, o VOT das oclusivas surdas teve maior duração diante de vogal alta para 4 dos 6 informantes, ou seja, tal tendência se confirmou para ambos os informantes aprendizes, bem como para P1 e N2.

É interessante apontar que embora Schwartzhaupt (2012) não tenha encontrado diferenças significativas na produção de nativos com relação à qualidade da vogal seguinte, ele observa que o ponto de articulação pode exercer influência concomitante nesse aspecto, o que precisa ser confirmado com a análise de um maior número de sujeitos. Em seu estudo, o teste MANOVA apontou diferença significativa quanto à qualidade da vogal para bilabiais [$F(3,76) = 5,316$; $p < 0,05$] e para velares [$F(3,76) = 9,737$; $p < 0,05$] na produção de falantes proficientes de inglês. Entretanto, com os valores de VOT relativo, a diferença para velares não se manteve significativa [$F(3,76) = 1,043$; $p = 0,378$]. De toda forma, para nativos e brasileiros de seu estudo, os resultados sugerem maiores valores de VOT diante de vogal posterior alta para [p] e [t] e diante de vogal anterior alta para [k]. Diante disso, no experimento pretendemos analisar a influência da vogal seguinte separando os dados por ponto de articulação, com intuito de trazer mais precisão à nossa análise.

O Gráfico 4 apresenta os resultados apurados no piloto, por ora com os três pontos de articulação agrupados. Fica claro que, a produção dos brasileiros de forma geral, tem oclusivas com os maiores valores de VOT diante da vogal anterior alta, o que não pode ser

dito da produção dos nativos. Um maior número de sujeitos e a análise estatística dos dados nos permitirão considerações mais precisas.

(a) Duração relativa de VOT por informante. (b) Duração relativa por grupo.

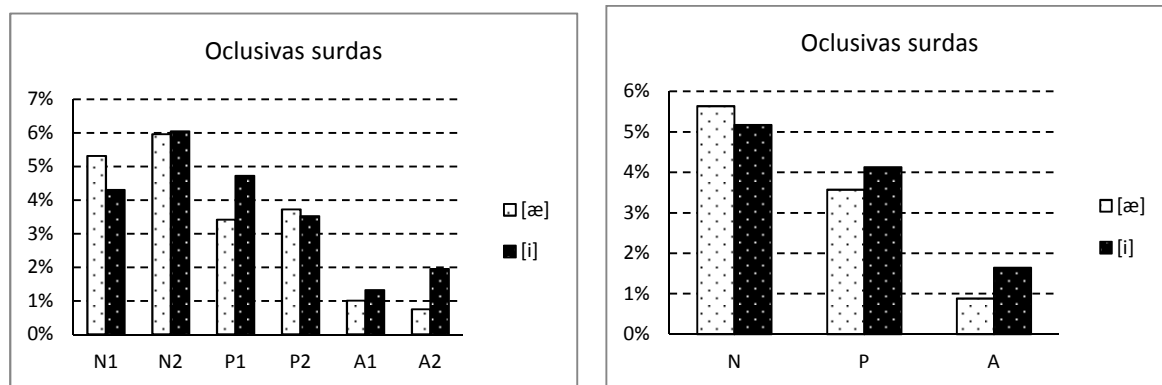


Gráfico 4: Duração relativa de VOT de consoantes oclusivas surdas por informante e por grupo diante das vogais alta [i] e baixa [æ].

Quanto às oclusivas sonoras, a análise neste piloto apresentou a falha de combinar os dados de oclusivas com VOT negativo (pré-vozeado) e VOT positivo (intervalo curto). Posteriormente, chegamos à conclusão de que, no experimento completo, seria mais adequado separar as oclusivas com pré-vozeamento das oclusivas com intervalo curto, já se trata de duas realizações de natureza e durações muito distintas e que, portanto, devem ser observadas de maneira independente. Além disso, diante dos resultados obtidos por Schwartzhaupt (2012) para oclusivas surdas, optaremos por separar os dados por ponto de articulação, já que essa variável também poderá exercer influência sobre a duração do VOT com relação à altura da vogal. Desse modo, por acreditar que a análise do piloto não foi efetuada com o rigor necessário, optamos por rever nossa metodologia e apresentar resultados apenas no capítulo 5, com a condução do experimento.

4.4 VOT E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL SEGUINTE

Os resultados encontrados neste estudo-piloto corroboram apenas em parte aqueles obtidos por Alves e Dias (2010), que apontam VOT de maior duração diante da vogal anterior [i] do que diante da posterior [u] no PB. Constatou-se que, em relação aos informantes N1 e N2, o VOT das oclusivas diante da vogal posterior superou em duração o VOT das oclusivas diante de vogal anterior, o que, portanto, não coincide com os resultados de Alves e Dias

(2010). Para N1, o VOT relativo diante de vogal posterior foi de 3,85% e diante de anterior foi 3,5%. Para N2, esses valores foram de 5,82% e 5,1%, respectivamente. A mesma tendência foi seguida pelo informante P1, que exibiu VOT de 4,16% diante de [ʊ] e 3,6% diante de [ɪ]. Já para a informante P2, o VOT de oclusivas diante da vogal anterior superou levemente aquele encontrado diante da vogal posterior (4,07% contra 4,02%). Por fim, ambos os aprendizes exibiram aumento considerável do VOT diante da vogal anterior. Para A1, os valores são de 1,16% contra 2,6% e para A2 são de 1,99% contra 2,2%. Portanto, o VOT das oclusivas surdas foi maior diante de [ɪ] para 3 dos 6 informantes. Quando analisamos as médias por grupo, observamos que os grupos N e P apresentaram VOT mais longo diante da vogal posterior, ao passo que o grupo A exibe VOT mais longo diante da vogal anterior. Tais resultados serão posteriormente separados por ponto de articulação, tendo em vista as considerações feitas por Schwartzhaupt (2012) e mencionadas na seção 4.2.1.

O Gráfico 5 apresenta os resultados do piloto.

(a) Duração relativa de VOT por informante. (b) Duração relativa de VOT por grupo.

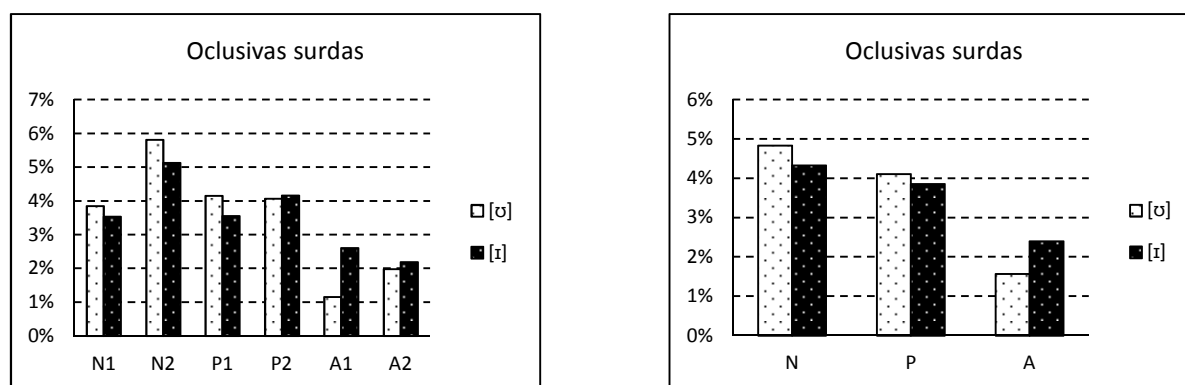


Gráfico 5: Duração relativa de VOT de consoantes oclusivas surdas por informante e por grupo diante das vogais anterior [ɪ] e posteriores [ʊ].

Quanto às oclusivas sonoras, valem as mesmas observações feitas com relação à influência da altura da vogal seguinte na duração do VOT, ou seja, chegamos à conclusão de que, no experimento completo, será mais adequado separar as oclusivas com pré-vozeamento das oclusivas com intervalo curto, bem como separar os dados por ponto de articulação, já que essa variável também poderá exercer influência sobre a duração do VOT, segundo os resultados obtidos por Schwartzhaupt (2012).

4.5 CONCLUSÃO

O estudo-piloto nos permitiu testar a metodologia planejada e também repensar alguns questionamentos em foco nesta pesquisa. O único ajuste metodológico necessário foi o já mencionado na seção 3.3 e diz respeito aos sujeitos do grupo P, que inicialmente foram selecionados mediante dois critérios: formação em Letras Inglês pela Universidade Federal do Paraná e experiência de pelo menos 5 anos em institutos de idiomas. Entretanto, após a banca de qualificação, decidimos acatar a sugestão de aperfeiçoar ainda mais nossos critérios incluindo um item que fornece uma garantia a mais com relação ao domínio de língua do grupo P: a exigência de que seus sujeitos tenham passado por algum exame internacional de proficiência na língua inglesa. Como ambos os informantes do piloto já eram detentores de certificados internacionais, não houve a necessidade de descartá-los do experimento; no entanto, tomou-se a precaução de selecionar os demais sujeitos levando-se em conta esse terceiro critério, além dos dois inicialmente aventados.

O piloto nos permitiu ainda amadurecer nossos questionamentos, que passaram de um momento inicial, no qual queríamos mensurar o quanto faltava para que a produção de um aprendiz se equiparasse a de um falante nativo de inglês, para um momento seguinte, no qual resolvemos inserir o grupo de falantes proficientes e passar a comparar as produções tendo em vista critérios mais elaborados, ou seja, a influência do nível de proficiência dos falantes de PB, do ponto de articulação e da qualidade da vogal seguinte na duração do VOT das oclusivas. Além disso, com o estudo-piloto, chegamos à conclusão de que seria necessária a separação por ponto de articulação quando da análise da influência da altura e posição ântero-posterior da vogal seguinte, pois a resposta poderia ser diferente para cada ponto em relação a essas variáveis.

5. O EXPERIMENTO

Passemos agora a reportar a realização do experimento. O *corpus*, descrito na seção 3.2, foi mantido tal qual o concebemos inicialmente. Quanto aos critérios para a seleção dos sujeitos, o único ajuste realizado diz respeito à seleção dos informantes que constituíram o grupo P (falantes proficientes), cujo grau de competência na língua inglesa passou a ser verificado mediante teste ou certificação internacional. Todos os sujeitos do grupo P, portanto, eram portadores de algum certificado internacional e o nível mínimo atestado foi B2, de acordo com o Quadro Comum Europeu de Referência para Línguas. Mais detalhes encontram-se na seção 3.3 deste trabalho. Os dados utilizados neste experimento foram coletados junto a sujeitos representantes de três grupos distintos: seis brasileiros aprendizes de inglês como língua estrangeira (grupo A), seis brasileiros falantes proficientes de inglês como língua estrangeira (grupo P) e seis falantes nativos de inglês norte-americano (grupo N).

A gravação dos dados ocorreu em parte no Laboratório de Estudos Fônicos da UFPR (LEFON) e em parte no laboratório da Universidade Tecnológica Federal do Paraná, ambos possuindo cabine com tratamento acústico. A análise dos dados foi feita por intermédio do software *Praat*. (Boersma e Weenick, 2012). Todos os dados do experimento são apresentados em sua duração relativa, a fim de neutralizar quaisquer vieses que venham a ocorrer devido à taxa de elocução dos informantes. As seções 3.4 e 3.5 apresentam mais detalhes sobre os procedimentos de coleta e análise dos dados. Os dados em sua duração absoluta poderão ser conferidos nos Apêndices D a K.

Passemos agora a reportar os resultados do experimento, que serão apresentados levando em conta a interferência dos seguintes fatores na duração de VOT das oclusivas surdas e sonoras: nível de proficiência do sujeito, ponto de articulação da consoante, altura da vogal seguinte e posição ântero-posterior da vogal seguinte. O ponto de articulação tem um papel de destaque em nosso estudo, pois além de ser objeto de uma seção exclusiva, também estará envolvido na análise dos demais fatores, uma vez que nível de proficiência, altura e posição da vogal seguinte terão suas análises separadas por ponto. Na conclusão, retomamos as perguntas e hipóteses desta pesquisa à luz dos resultados obtidos. Saliente-se que os testes estatísticos somente foram rodados com os dados relativos às oclusivas surdas e, portanto, a elaboração e testagem de hipóteses só foi possível para esse grupo de consoantes. Quanto às sonoras, apresentamos e discutimos os resultados encontrados sem a testagem de hipóteses.

Tal limitação se deve ao fato de que, na coleta e análise dos dados, encontramos sonoras com duas realizações distintas, o que fez com que tivéssemos de separar as sonoras em dois grupos e, portanto, a quantidade de dados em cada grupo ficou reduzida demais para que fosse submetida a testes estatísticos.

5.1 VOT E O NÍVEL DE PROFICIÊNCIA

Nesta seção tratamos dos resultados encontrados tendo em vista o nível de proficiência dos falantes de PB. Os grupos de falantes proficientes (grupo P) e aprendizes (grupo A) terão suas produções comparadas entre si e com o grupo controle dos nativos (grupo N). Os resultados serão apresentados por ponto de articulação, uma vez que essa variável provou exercer influência de destaque na apuração dos dados. Nesta seção, assim como nas seguintes, iniciamos com a análise das consoantes surdas, para, em seguida, abordar as sonoras.

5.1.1 OCLUSIVAS SURDAS E O NÍVEL DE PROFICIÊNCIA

As médias de duração relativa de VOT (%) e o desvio padrão (DP) verificados para cada um dos três grupos seguem na tabela 17. Para cada grupo, analisamos 210 *tokens* por consoante. Valores absolutos de VOT podem ser encontrados no Apêndice D. Para facilitar a visualização dos dados, traremos as médias também em forma de gráfico no decorrer de todo este capítulo. Logo, à tabela 17, segue-se o gráfico 6.

Oclusivas surdas – duração relativa									
Consoante	[p]			[t]			[k]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	210	4,76	1,85	210	5,40	1,67	210	5,24	1,51
Grupo P	210	2,13	2,84	210	3,72	1,62	210	4,57	1,63
Grupo A	210	1,08	0,66	210	2,03	1,48	210	3,06	2,07

Tabela 17: Média e desvio padrão da duração relativa de VOT das oclusivas surdas por ponto de articulação.

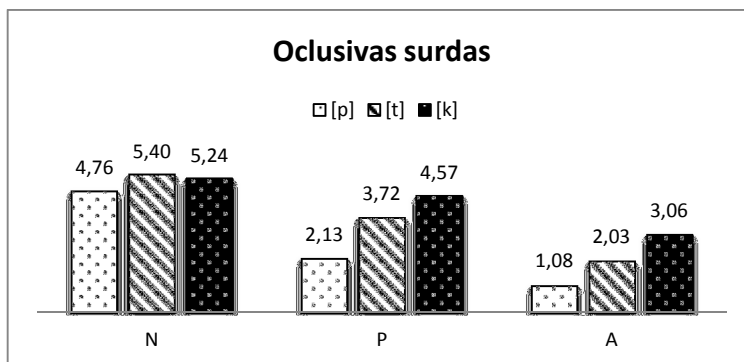


Gráfico 6: Média da duração relativa de VOT das consoantes oclusivas surdas por ponto de articulação.

Para a condução dos testes estatísticos, adotamos o seguinte procedimento: inicialmente rodamos o teste de Friedman com o intuito de observar se haveria diferença significativa na produção das três consoantes para cada grupo. Constatamos não haver diferença significativa entre nenhum dos pontos para o grupo N ($X^2(2, N=6) = 4,333, p = 0,115$). Todavia, para os grupos P e A, a diferença na produção das consoantes foi estatisticamente significativa. Para ambos o resultado foi ($X^2(2, N=6) = 12,000, p = 0,002$). Em seguida, aplicamos o teste de Wilcoxon para os grupos P e A e verificamos haver diferença significativa entre todos os pontos de articulação, tanto para o grupo P quanto para o A. Novamente os resultados foram os mesmos para os dois grupos: entre cada um dos contrastes analisados, ou seja, [p] e [t], [p] e [k] e [t] e [k] encontramos ($Z = -2,201, p = 0,028$), tendo a bilabial a menor duração relativa de VOT e a velar a maior, em ambos os grupos.

Assim sendo, optamos por analisar a diferença de produção entre os grupos organizando os dados por ponto de articulação. Nessa etapa, utilizamos o teste de Mann-Whitney para compararmos dois grupos de cada vez. A primeira comparação foi entre o grupo N e o grupo A; a segunda entre o grupo N e o grupo P; e finalmente comparamos a produção dos grupos P e A.

Como o esperado, a média do grupo N superou a do grupo A, havendo diferença significativa na produção dos três pontos de articulação. Para [p] obtivemos ($Z = -2,882, p = 0,004$), para [t] ($Z = -2,722, p = 0,006$) e para [k] ($Z = -2,722, p = 0,006$).

Entre os grupos N e P, a média de N também foi maior em todos os pontos. No entanto, a diferença foi significativa somente para o ponto bilabial ($Z = -2,722, p = 0,006$), mas não para os pontos alveolar ($Z = -1,601, p = 0,109$) e velar ($Z = -1,281, p = 0,200$).

Finalmente, a média de duração de VOT do grupo P superou a do grupo A, sendo observada diferença significativa para os três pontos de articulação, com os seguintes

resultados: para [p] ($Z = -2,242$, $p = 0,025$), para [t] ($Z = -2,082$, $p = 0,037$) e para [k] ($Z = -2,242$, $p = 0,025$).

5.1.2 OCLUSIVAS SONORAS E O NÍVEL DE PROFICIÊNCIA

Como já mencionado na seção 2.4.3, as oclusivas sonoras do inglês caracterizam-se por possuir um VOT com duração absoluta entre -20 ms até +20 ms. Valores negativos ocorrem quando o vozeamento precede o *burst* em alguns milissegundos (KENT; READ, 1992). Por outro lado, o PB conta com oclusivas sonoras cujo VOT é primordialmente negativo, com valores bem acima dos -20ms apontados para o inglês, atingindo em média -100ms (REIS; NOBRE-OLIVEIRA, 2007; BONATTO, 2007).

Nos dados deste experimento, as oclusivas sonoras tiveram sua realização variando basicamente de duas maneiras: ora o início do vozeamento antecedia o *burst*, resultando em VOT de valores negativos; ora o início do vozeamento ocorria concomitantemente ao momento da soltura ou alguns segundos depois dela, resultando em VOT = zero ou de valores positivos próximos a zero, o que é considerado VOT de intervalo curto (CHO; LADEFOGED, 1999). As variações ocorreram tanto de um informante para outro, como na produção de um mesmo informante. Quando o VOT teve valores negativos, normalmente a barra de vozeamento foi observada durante todo o período de oclusão, o que pode ter sido propiciado pelo contexto intervocálico, uma vez que a palavra-alvo foi inserida na frase-veículo “Say_____ to me”.

A diferença entre os três grupos foi registrada basicamente na proporção de dados realizados conforme um ou outro padrão de vozeamento. Na produção do grupo N, houve predominância de dados com VOT de intervalo curto. No entanto, não se pode dizer que essa predominância foi uniforme para todos os sujeitos do grupo. O informante N3, por exemplo, teve 91 dos 105 dados de sonoras produzidos com pré-vozeamento e apenas 14 dados com VOT de intervalo curto; ao passo que a informante N4 apresentou todos os 105 registros de sonoras com intervalo curto.

Por outro lado, nos grupos P e A, houve maior registro de dados com pré-vozeamento. Percebe-se, no entanto, que os informantes do grupo P ainda apresentam maior variação na produção das oclusivas do que os informantes do grupo A. O gráfico 7 mostra como se dá a realização das oclusivas sonoras para os três grupos, separando as produções por informante.

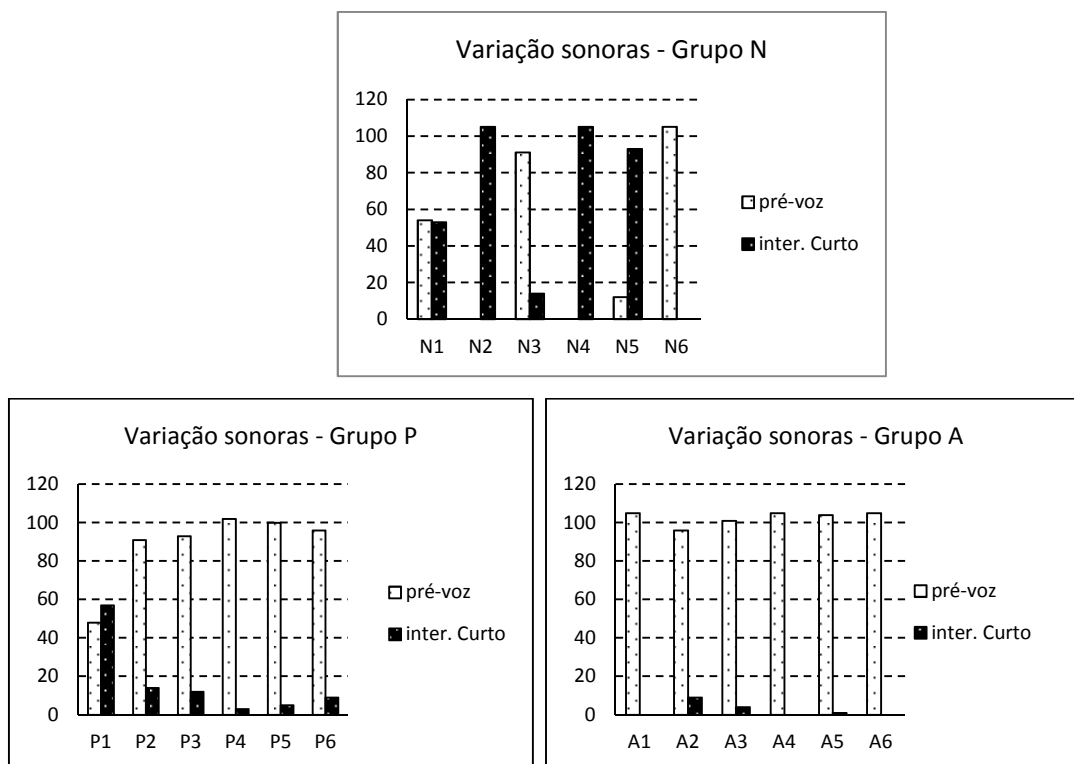


Gráfico 7: Variação das oclusivas sonoras por informante no experimento.

Assim como sugerem os resultados do piloto, o experimento confirma que podemos estar diante de um refinamento na produção dos falantes de PB quanto à produção das oclusivas sonoras conforme o domínio da língua se consolida, devido à diminuição de dados com VOT de pré-vozeamento, característica mais persistente na produção de falantes de PB. Considerando a produção por grupo, temos que 98% dos dados do grupo A apresentam pré-vozeamento, enquanto que essa porcentagem cai para 84% no grupo P e fica em 41% no grupo N.

Uma vez compreendida a distribuição dos dados relativos às oclusivas sonoras, passemos a investigar como se deu a duração do VOT dessas consoantes, assim como fizemos com as surdas. Salientamos, entretanto, que apresentaremos os resultados de pré-vozeamento e intervalo curto separadamente, sistema adotado inclusive por Lisker e Abramson (1964) quando se depararam com a mesma variação nas produções de seus informantes. Segue, primeiramente, a discussão dos resultados relativos aos dados nos quais a oclusiva sonora foi realizada com VOT de valores negativos, ou seja, com pré-vozeamento. A tabela 18 e o gráfico 8 apresentam os resultados quanto à duração relativa de VOT. Os valores absolutos podem ser encontrados no apêndice E.

Oclusivas sonoras com pré-vozeamento									
Consoante	[b]			[d]			[g]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	88	9,91	1,73	83	8,62	1,90	89	8,21	2,52
Grupo P	181	11,89	2,63	166	11,08	2,73	183	10,99	2,68
Grupo A	202	10,54	2,46	206	8,85	2,55	208	8,45	2,06

Tabela 18: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com pré-vozeadas por ponto de articulação.

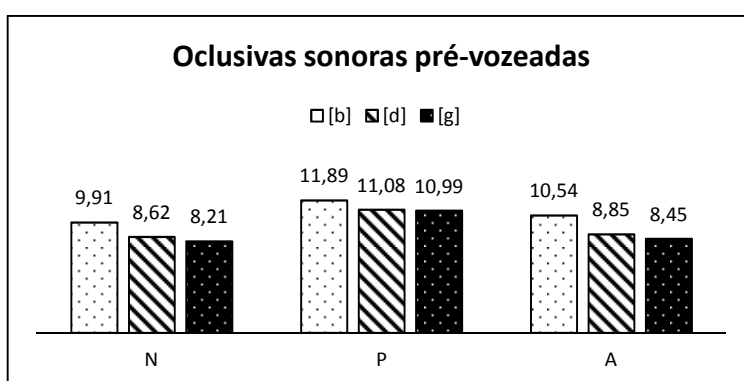


Gráfico 8: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação

Observamos que a duração de VOT na produção do grupo N é menor para todos os pontos de articulação e, curiosamente, aproxima-se mais da produção do grupo A do que do grupo P, que apresenta os maiores valores de VOT. Outro fator a ser notado é que as médias de duração relativa ficam entre 8,21 e 11,89, variação menor se comparada com a existente para a produção das oclusivas surdas, cujos valores estiveram entre 1,08 e 5,24.

Passemos agora aos dados nos quais as oclusivas sonoras são realizadas com VOT de intervalo curto. Voltamos a destacar que o maior número de ocorrências desse tipo de realização se deu na produção dos informantes do grupo N. A tabela 19 e gráfico 9 apresentam as médias de duração relativa de VOT. Valores absolutos podem ser encontrados no apêndice F.

Oclusivas sonoras com intervalo curto									
Consoante	[b]			[d]			[g]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	122	0,84	0,57	127	1,20	0,54	121	1,99	0,67
Grupo P	29	0,53	0,58	44	1,37	0,72	27	2,01	0,55
Grupo A	8	0,19	0,13	4	0,95	0,20	2	1,06	0,26

Tabela 19: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.

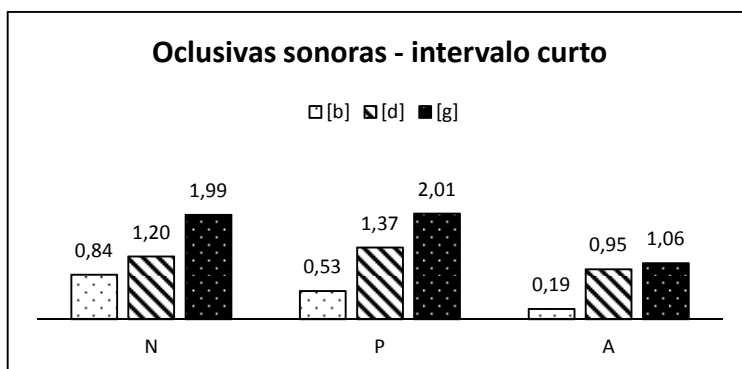


Gráfico 9: Média da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.

Esperaríamos aqui um comportamento semelhante àquele observado na produção das consoantes surdas, já que também temos valores positivos de VOT, embora em duração mais reduzida. Poder-se-ia supor, por exemplo, que o VOT positivo partisse de um ponto zero com os aprendizes, tivesse um valor intermediário com os falantes proficientes, para finalmente alcançar uma duração mais longa com os nativos. Todavia, temos que, nos dados de intervalo curto, o VOT dos falantes proficientes supera em duração o VOT dos nativos nos pontos alveolar (1,37 contra 1,20) e velar (2,01 contra 1,99). Apenas com relação às bilabiais, o grupo N supera o grupo P (0,84 contra 0,53).

Isso nos leva a pensar que, no processo de aquisição, parte-se de um momento inicial, no qual os índices de VOT se assemelham à L1 (VOT negativo), para um momento seguinte, no qual o domínio mais amplo da língua permite aos falantes proficientes perceber e produzir VOT de intervalo curto. Porém, em tal tarefa, apresentam VOT mais longo do que o próprio grupo-controle, o que pode ser indício de que, na tentativa de coordenar os gestos articulatórios para se aproximar da forma-alvo, haja inicialmente uma amplificação do detalhe que é mais notadamente percebido para que, quiçá em um momento posterior, o faseamento dos gestos venha a ajustar.

5.2. VOT E O PONTO DE ARTICULAÇÃO

Nesta seção apresentamos os resultados encontrados com relação à influência exercida pelo ponto de articulação na duração de VOT de consoantes oclusivas surdas e sonoras. Os resultados serão apresentados separadamente para cada um dos três grupos pesquisados: nativos (grupo N), falantes proficientes (grupo P) e aprendizes (grupo A). Como

na seção 5.1, iniciamos nossa análise com as consoantes surdas, para, em seguida, abordar as sonoras.

5.2.1 OCLUSIVAS SURDAS E O PONTO

Os resultados deste experimento indicam que o ponto de articulação exerce influência na duração do VOT para as consoantes oclusivas surdas, tanto na produção de nativos quanto de falantes de PB. Os dados aqui analisados já foram reportados na seção 5.1, embora o enfoque tenha sido o nível de proficiência. Portanto, reproduzo novamente na tabela 20 e gráfico 10 as informações sobre as médias de duração relativa de VOT por ponto de articulação para os grupos N, P e A, a fim de poder discuti-las sob essa nova perspectiva.

Oclusivas surdas – duração relativa									
Consoante	[p]			[t]			[k]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	210	4,76	1,85	210	5,40	1,67	210	5,24	1,51
Grupo P	210	2,13	2,84	210	3,72	1,62	210	4,57	1,63
Grupo A	210	1,08	0,66	210	2,03	1,48	210	3,06	2,07

Tabela 20: Média e desvio padrão da duração relativa de VOT das oclusivas surdas por ponto de articulação.

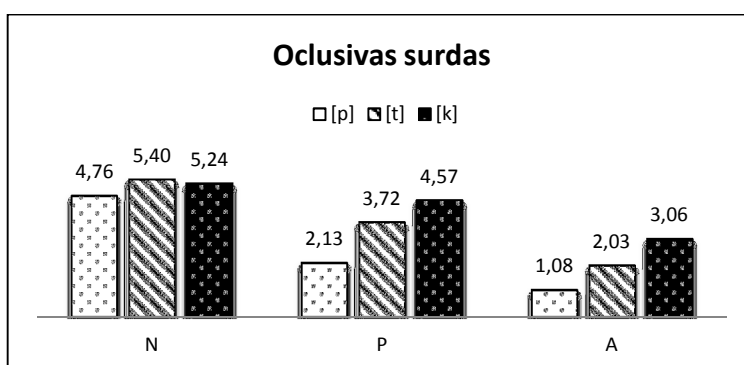


Gráfico 10: Média da duração relativa de VOT das consoantes oclusivas surdas por ponto de articulação.

Observa-se que, para os grupos de falantes de PB (P e A), quanto mais posterior o ponto de articulação, mais longa é a duração de VOT, o que corrobora pesquisas anteriores (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1999; YAVAS, 2008; STEIN, 2011; FRANÇA, 2011; SCHWARTZHAUPT, 2012).

No entanto, essa relação não se mantém para o grupo N, no qual os valores de VOT das alveolares superam o das velares. Ao investigarmos se tal resultado pode ter-se devido à produção de um informante em particular, observamos que os valores de VOT das alveolares superam os das velares para os informantes N2, N3, N4 e N5, ou seja, mais da metade dos informantes desse grupo. Inclusive, a produção de N4 contraria completamente o esperado e a duração de VOT em seus dados diminui conforme o ponto de articulação se posterioriza. O gráfico 11 mostra as produções do grupo N por informante.

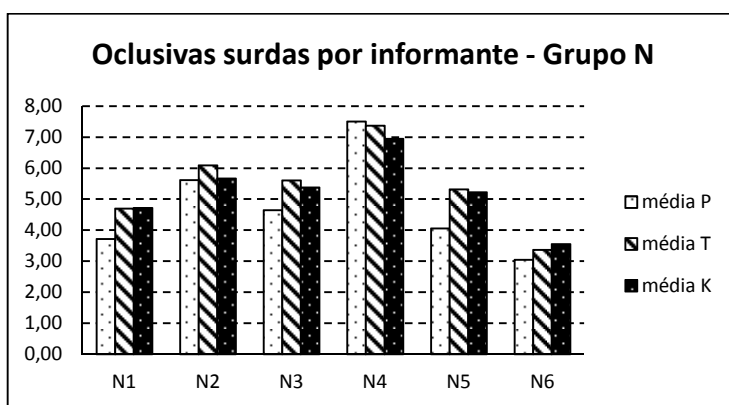


Gráfico 11: Duração relativa das oclusivas surdas por informante no grupo N.

Fica claro ao observarmos o gráfico 8 que, de forma geral, a duração de VOT não muda substancialmente de um ponto a outro. Quando submetemos os resultados do grupo N à análise estatística, observamos que não houve efetivamente diferença significativa na duração de VOT entre os três pontos. Foi aplicado o teste de Friedman, com nível de significância 0,05 e constatamos não haver diferença significativa entre nenhum dos pontos para o grupo N ($X^2(2, N=6) = 4,333, p = 0,115$).

Para os grupos P e A, entretanto, a produção é uniforme para todos os informantes, que exibem duração progressivamente mais longa de VOT conforme o ponto de articulação se torna mais posterior. Além disso, observamos que a diferença é maior entre os pontos, conforme o gráfico 12.

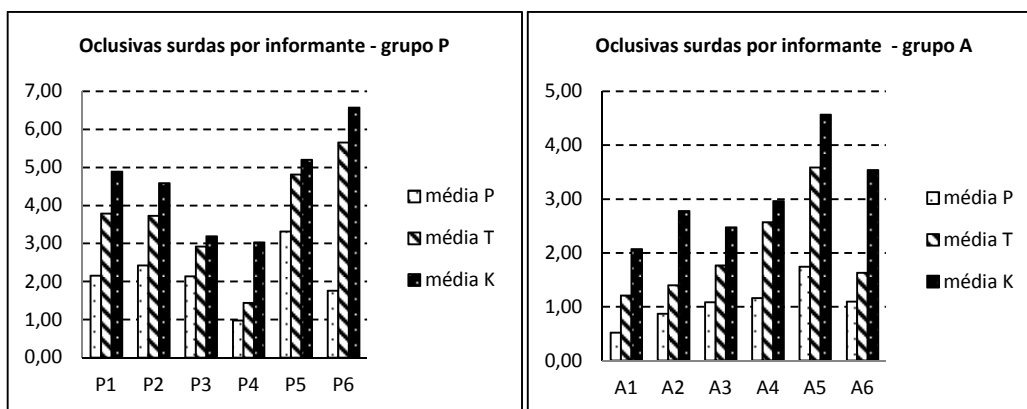


Gráfico 12: Duração das oclusivas surdas por informante nos grupos P e A.

A partir da aplicação do teste de Friedman, a diferença na produção das consoantes foi estatisticamente significativa para os grupos P e A. Para ambos o resultado foi ($X^2(2, N=6) = 12,000, p = 0,002$). Com o teste de Wilcoxon, verificamos haver diferença significativa entre todos os pontos de articulação, tanto para o grupo P quanto para o A. Novamente os resultados foram os mesmos para os dois grupos: entre cada um dos contrastes analisados, ou seja, [p] e [t], [p] e [k] e [t] e [k] encontramos ($Z = -2,201, p = 0,028$), tendo a bilabial a menor duração relativa de VOT e a velar a maior.

5.2.2 OCLUSIVAS SONORAS E O PONTO

Agora passemos a analisar como se dá a produção das oclusivas sonoras para os três grupos com relação ao ponto de articulação. Como já esclarecido na seção 5.1.2, iremos analisar as duas realizações de oclusivas sonoras em separado. Iniciaremos com as oclusivas sonoras com VOT pré-vozeado. Salientamos que o grupo que mais exibiu dados dessa natureza foi o grupo P (98%), seguido do grupo A (84%) e do grupo N (41%). A tabela 21 e o gráfico 13 levantam os resultados obtidos.

Oclusivas sonoras com pré-vozeamento									
Consoante	[b]			[d]			[g]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	88	9,91	1,73	83	8,62	1,90	89	8,21	2,52
Grupo P	181	11,89	2,63	166	11,08	2,73	183	10,99	2,68
Grupo A	202	10,54	2,46	206	8,85	2,55	208	8,45	2,06

Tabela 21: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação.

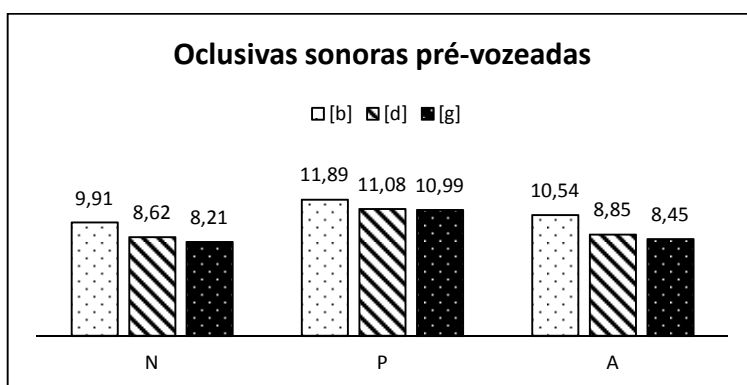


Gráfico 13: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas por ponto de articulação

Observamos uniformidade de comportamento nos três grupos: quanto mais anterior o ponto de articulação, maior a duração do VOT, resultado que é justamente o oposto do encontrado para as oclusivas surdas. Entretanto, a diferença entre os pontos não parece ser significativa em nenhum dos grupos. Infelizmente, isso não poderá ser comprovado com a realização de análise estatística, uma vez que o número de dados caiu consideravelmente com a divisão das oclusivas em dois grupos e alguns dos sujeitos simplesmente não apresentaram nenhuma ocorrência de pré-vozeamento.

Nossos resultados se assemelham aos encontrados por Lisker e Abramson (1964) para as oclusivas sonoras com VOT negativo, para as quais foram apurados os seguintes valores de VOT absoluto em palavras isoladas: [b] = -101ms, [d] = -102ms e [g] = -88ms. Em posição inicial de frase, os pesquisadores obtiveram médias de -65ms, -56ms e -45ms, respectivamente. Em posição medial, a média para [b] foi de -63ms. A ausência de valores de VOT negativo para [d] e [g] é justificada pelos pesquisadores pela existência do chamado *unbroken voicing*, ou seja, vozeamento proveniente do segmento precedente que segue sem interrupção até a soltura dos articuladores da oclusiva em questão. Esse fenômeno se deu em todas as ocorrências do [d] e do [g] e, embora tenha ocorrido também com o [b], permitiu que fossem realizadas algumas medições.

Em nosso trabalho, optamos por não levar em consideração a questão do *unbroken voicing*, pois acreditamos que é preciso relatar de que maneira o VOT da consoante sonora se comporta em contexto medial e se isso envolver a ocorrência de *unbroken voicing*, não há porque deixar de relatar o fenômeno, uma vez que ele parte do processo observado quando a consoante em início de palavra está inserida na cadeia da fala.

Na Tabela 22, temos um resumo das médias das oclusivas sonoras com VOT negativo na pesquisa de Lisker e Abramson (1964) com falantes norte-americanos em contraste com as médias de VOT absoluto encontradas nesta pesquisa. Os valores de VOT relativo seguem entre parênteses.

Lisker e Abramson (1964)	[b]	[d]	[g]
Palavras isoladas	-101	-102	-88
Posição inicial	-65	-56	-45
Posição medial	-63		
Presente pesquisa			
Posição medial (grupo N)	-102 (9,91)	-90 (8,62)	-84 (8,21)
Posição medial (grupo P)	-127 (11, 89)	-121 (11,08)	-120 (10,99)
Posição medial (grupo A)	-122 (10,54)	-105 (8,85)	-100 (8,45)

Tabela 22: Valores médios encontrados para o VOT absoluto das oclusivas sonoras pré-vozeadas do inglês norte-americano em Lisker e Abramson (1964) e neste estudo.

Nota-se que, em ambos os estudos, o valor de VOT nas oclusivas velares é sempre o menor e os resultados de Lisker e Abramson (1964) para palavras em posição inicial de frase também apresentam VOT mais longo conforme o ponto se torna mais anterior. Tal tendência também parece estar presente no PB. Melo et al. (2011) encontraram os seguintes valores para o PB: [b] = -101,9, [d] = -95,6 e [g] = -80,4 e embora Klein (1999) tenha apurado praticamente os mesmos valores para [b] = -92,27 e [d] = -92,07, a velar apresenta o VOT mais curto entre as três oclusivas, [g] = -78,69.

Novas pesquisas são necessárias para observar se a tendência aqui observada se confirma mediante um maior número de sujeitos e a aplicação de testes estatísticos apropriados.

Passemos agora à discussão dos dados relativos às oclusivas sonoras com intervalo curto. Vale lembrar que a maioria dos dados produzidos nessas condições pertence ao grupo N (59%), seguido do grupo P (16%) e do grupo A (2%)

A natureza desta produção se assemelha à das oclusivas surdas do inglês, uma vez que em ambas há ausência de vozeamento durante a oclusão e os valores de VOT são positivos, muito embora neste caso, a duração seja consideravelmente menor. Diríamos que, devido aos níveis de aspiração exibidos, os dados das oclusivas sonoras de intervalo curto guardam semelhança maior ainda com as oclusivas surdas do PB. Apesar das semelhanças, a configuração dos dados, especialmente no grupo N, apresentou algumas mudanças interessantes. A tabela 23 e o gráfico 14 indicam os resultados apurados neste estudo.

Oclusivas sonoras com intervalo curto									
Consoante	[b]			[d]			[g]		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
Grupo N	122	0,84	0,57	127	1,20	0,54	121	1,99	0,67
Grupo P	29	0,53	0,58	44	1,37	0,72	27	2,01	0,55
Grupo A	8	0,19	0,13	4	0,95	0,20	2	1,06	0,26

Tabela 23: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.

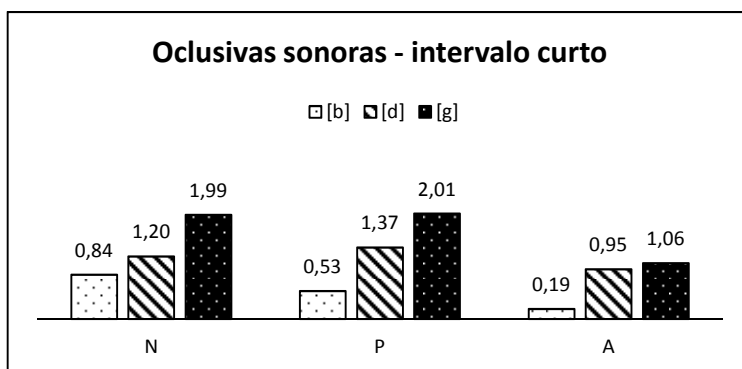


Gráfico 14: Média da duração relativa de VOT das oclusivas sonoras com intervalo curto por ponto de articulação.

Em primeiro lugar, cabe lembrar que, na produção das oclusivas surdas pelo grupo N, os valores de VOT no ponto alveolar superaram os valores apurados para o ponto velar. Além disso, os índices de VOT foram bastante semelhantes para os três pontos de articulação, não havendo, inclusive, diferença estatisticamente significativa entre eles. Por outro lado, para as oclusivas sonoras de intervalo curto, parece haver uma diferença mais expressiva entre um ponto de articulação e o seguinte. Além disso, a duração do VOT aumenta conforme o ponto se torna mais posterior, o que assemelha a produção do grupo N a dos grupos P e A.

Parece-nos, portanto, que entre a produção de uma surda aspirada (intervalo longo) e de uma sonora com VOT positivo (intervalo curto), há algo mais do que a simples diminuição da duração de VOT. A proporção em que essa diminuição é feita parece se alterar, havendo menor diferença entre os pontos de articulação quando a oclusiva é aspirada e outra configuração com relação à influência do ponto de articulação, o que sugere estratégias articatórias diferentes na produção das duas categorias. Infelizmente, o baixo número de ocorrências de dados com oclusivas sonoras de intervalo curto e o fato de que nem todos os informantes do grupo N as produziram, impedirá a condução de análise estatística.

Destacamos ainda a baixa duração de VOT da bilabial com relação aos demais pontos na produção dos grupos P e A. Parece haver, assim como na produção das surdas, maior

dificuldade na coordenação dos gestos que levam à forma-alvo quando se trata desse ponto de articulação.

Comparemos agora na tabela 24 os resultados aqui apurados com os obtidos em três outros estudos.

Inglês norte-americano			
	[b]	[d]	[g]
Lisker e Abramson (posição medial)	4	7	16
Klatt (posição inicial)	11	17	27
Macken e Barton (diferentes posições)	1 a 11	5 a 17	21 a 27
Presente pesquisa (posição medial)			
Grupo N	9	13	22
Grupo P	6	15	22
Grupo A	2	12	14

Tabela 24: Valores médios encontrados para o VOT absoluto das oclusivas sonoras de intervalo curto do inglês norte-americano neste estudo, em Lisker e Abramson (1964), Klatt (1975) e Macken e Barton (1978)

Os valores de VOT absoluto encontrados neste estudo para a produção dos falantes nativos (grupo N) são notadamente maiores do que os de Lisker e Abramson (1964) apurados para a posição medial. Assemelham-se mais àqueles encontrados em Klatt (1975), cuja análise envolveu oclusivas em posição inicial de palavra inseridas na frase veículo “*Say_____ instead.*”. As médias em Klatt (1975) foram de 11ms para [b], 17ms para [d] e 27ms para [g]. Já em Macken e Barton (1978), foram reportados os seguintes valores de VOT para as oclusivas sonoras do inglês: de 1ms a 11ms para [b], 5 a 17ms para [d] e 21 a 27ms para [g]. Reiteramos, portanto, a necessidade de mais estudos, especialmente para verificar se a diferença entre os pontos de articulação nas oclusivas sonoras de intervalo curto é estatisticamente significativa e também para contrastar a produção de oclusivas surdas aspiradas (60-100ms) e oclusivas sonoras com intervalo curto (0-25ms), verificando se o fato de as primeiras estarem na área de retardo longo influencia diferentemente a forma como o nível de VOT e o ponto de articulação interagem.

5.3 VOT E A ALTURA DA VOGAL SEGUINTE

Passemos agora a observar a possível influência da altura da vogal seguinte na duração do VOT para oclusivas surdas e sonoras. Compararemos os dados de consoantes que estão diante da vogal baixa [æ] e da vogal alta [i]. Sabemos que nem sempre a realização das vogais será uniforme; por exemplo, aprendizes tendem a pronunciar palavras como *cap* [kæp] utilizando a vogal média [ɛ]. De toda forma, certa diferença de altura é mantida. Os resultados serão apresentados por ponto de articulação, visto que tal variável parece exercer influência na duração do VOT, conforme já mencionado nas seções 4.2 e 4.3.

5.3.1 OCLUSIVAS SURDAS E A ALTURA DA VOGAL

Tem-se atestado que o VOT de oclusivas surdas costuma ser mais longo diante de vogais altas do que diante de baixas (PORT; ROTUNNO, 1979; OHALA, 1981; YAVAŞ, 2007). Vamos observar se isso se confirma para os dados deste estudo com relação a todos os pontos de articulação.

Optamos por selecionar duas vogais tensas (*tense*), pois a tensão das vogais tem provado influenciar a duração de VOT das consoantes (PORT; ROTUNNO, 1979). Como as demais palavras, foram gravadas cinco repetições para cada um dos 18 informantes. Portanto, temos 90 *tokens* da oclusiva diante de [æ] e 90 *tokens* da oclusiva diante de [i]. Com a divisão por grupos, foram analisados 30 *tokens* para cada grupo. Os resultados encontram-se na tabela 25 e gráfico 15. Os valores absolutos de VOT encontram-se no apêndice G.

Oclusivas surdas e a altura da vogal									
	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[pæ]	30	4,98	2,00	30	2,06	1,18	30	0,64	0,50
[pi]	30	4,55	1,74	30	2,27	1,00	30	1,09	0,65
[tæ]	30	5,78	1,72	30	3,46	1,45	30	1,48	1,38
[ti]	30	5,95	1,13	30	3,84	2,41	30	3,70	1,52
[kæ]	30	5,47	1,37	30	4,53	1,48	30	3,21	4,24
[ki]	30	5,42	1,06	30	5,09	1,94	30	3,85	1,50

Tabela 25: Média da duração relativa de VOT e desvio padrão para as oclusivas surdas com relação à altura da vogal seguinte.

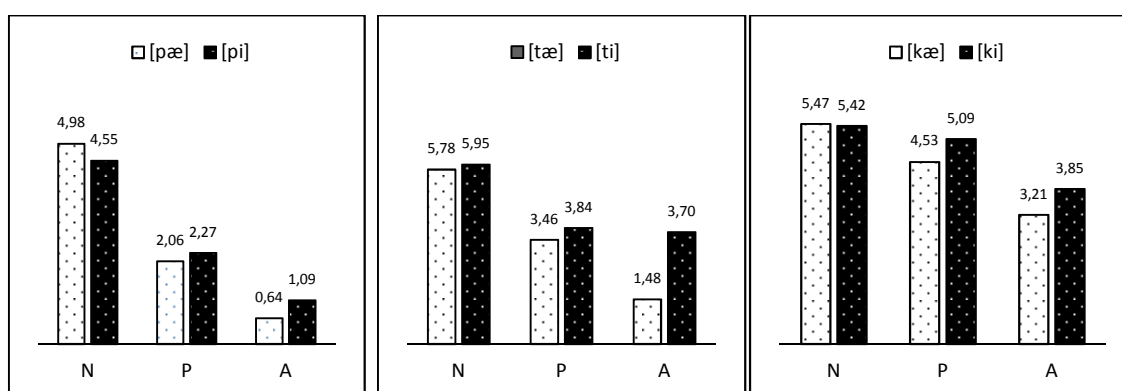


Gráfico 15: Média da duração relativa de VOT das oclusivas surdas com relação à altura da vogal seguinte.

Primeiramente, tratemos da oclusiva bilabial. Notamos que a duração de VOT na bilabial diante de [i] é maior do que diante [æ] na produção de ambos os grupos de falantes de PB (P e A); no entanto o inverso ocorre na produção do grupo N. A princípio, poderíamos pensar em uma eventual transferência do padrão fonético/fonológico da L1 para a L2 na produção do grupo P e A, já que no PB vogais altas tendem a causar VOT mais longo nas oclusivas que as antecedem (ALVES; DIAS, 2010). Todavia, estudos mostram que o VOT das oclusivas surdas do inglês também tem provado ser mais longo diante de vogais altas em comparação com as baixas (SUMMERFIELD, 1975; PORT; ROTUNNO, 1979; OHALA, 1981; FRANÇA, 2011; SCHWARTZHAUPT, 2012).

Para as oclusivas alveolares, a duração diante de [i] é maior nos três grupos, mas é para o grupo A que a diferença realmente se destaca. Uma possível explicação para esse fato é que a produção em inglês dos aprendizes pode estar sendo influenciada pelo processo de palatalização, comum na variedade de PB falada na região do Paraná, de onde são todos os sujeitos. Quando esse processo ocorre, as consoantes alveolares surda [t] e sonora [d] diante de [i] e suas variantes são produzidas como segmentos africados [tʃ] e [dʒ]. Assim, quando medimos o VOT nas consoantes que ocorrem nesse contexto, há o risco de que a medida de duração seja enviesada se o falante palatalizar a consoante em questão. Como o grupo A estaria, em princípio, mais propenso do que o grupo P a transferir características da L1 para a produção em inglês, supomos que esse seja o motivo pelo qual tenhamos flagrado duração consideravelmente mais longa de VOT de [t] diante de [i] para esse grupo.

Finalmente, quanto à oclusiva velar, a configuração dos dados se assemelha àquela observada para a bilabial, ou seja, os grupos P e A produzem a oclusiva com VOT mais longo diante de [i], enquanto o grupo N tem velares mais longas diante de [æ].

Com a aplicação do teste de Wilcoxon, com nível de significância de 0,05, observamos que a diferença entre a duração do VOT diante de vogal alta ou baixa só é significativa para o grupo A em todos os pontos de articulação. Nos demais casos, não houve nenhuma diferença significativa nos grupos N e P para nenhum dos pontos de articulação. Os resultados dos testes podem ser conferidos na tabela 26 e em negrito destacamos quando houve diferença significativa.

Wilcoxon para altura da vogal			
	N	P	A
[pæ]	Z = -0,051, p = 0,959	Z = - 0,751, p = 0,453	Z = - 2,561, p = 0,010
[pi]			

[tæ] [ti]	Z = - 0,977, p = 0,329	Z = - 0,710, p = 0,478	Z = - 2,849, p = 0,004
[kæ] [ki]	Z = -0,051, p = 0,959	Z = - 0,710, p = 0,478	Z = -2,520, p= 0,012

Tabela 26: Resultados do teste de Wilcoxon quanto à influência da altura da vogal seguinte.

Assim, podemos concluir que, na produção dos falantes de PB de forma geral, há uniformidade nos dados para os três pontos de articulação: as oclusivas surdas sempre exibem VOT mais longo diante da vogal alta, muito embora tal diferença seja significativa somente para o grupo A. Quanto ao grupo N, nos pontos bilabial e velar, o VOT é mais longo diante da vogal baixa, embora a diferença não seja significativa em nenhum dos pontos desse grupo.

Nossos resultados não coincidem completamente com aqueles obtidos por Schwartzhaupt (2012), que, na produção dos informantes norte-americanos encontrou VOT mais longo diante de vogal alta para os três pontos de articulação, ao passo que em nosso estudo, a vogal alta foi responsável por VOT mais longo somente no ponto alveolar. No que concerne aos falantes proficientes, nosso estudo encontrou VOT mais longo para todas as oclusivas diante da vogal alta, ao passo que Schwartzhaupt (2012), obteve o mesmo resultado para as oclusivas bilabial e velar, mas não quanto à alveolar, que teve VOT mais longo diante de vogal baixa.

5.3.2 OCLUSIVAS SONORAS E A ALTURA DA VOGAL

Os resultados concernentes às oclusivas sonoras, especialmente as de VOT de intervalo curto, terão de ser relativizados, uma vez que, como mencionado na seção 5.1.2, tivemos de separar os dados em duas categorias. A consequência disso foi que, muitas vezes, não foram produzidas consoantes oclusivas diante de certa vogal em uma das categorias. Por exemplo, no grupo A, não houve nenhuma ocorrência de oclusivas sonoras de intervalo curto diante das vogais [æ] e [i] nos pontos bilabial e velar. O grupo P teve algumas ocorrências, mas podemos dizer que o número é efetivamente baixo para que possamos tirar quaisquer conclusões mais acuradas. O grupo N foi o único que teve um número razoável de ocorrências de oclusivas sonoras com intervalo curto.

Iniciaremos nossa análise com as oclusivas sonoras pré-vozeadas. Embora tenhamos um número reduzido de dados, é possível fazer algumas observações sobre a produção dos três grupos, uma vez que o grupo que produziu menos oclusivas sonoras pré-vozeadas, ainda

assim o fez em 41% dos dados das sonoras. Na tabela 27 temos as médias e o número de ocorrências para cada grupo por ponto de articulação. No apêndice H encontram-se os valores absolutos de VOT e no gráfico 16 é possível visualizar mais claramente os resultados.

Oclusivas sonoras pré-vozeadas e a altura da vogal seguinte

	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[bæ]	11	9,89	1,73	22	10,98	2,56	30	10,06	1,44
[bi]	16	9,13	1,49	27	11,3	2,66	30	10,56	2,06
[dæ]	11	7,51	1,27	23	10,77	2,26	28	8,43	2,26
[di]	8	9,97	1,36	26	11,32	2,54	29	9,31	2,97
[gæ]	13	8,45	1,02	29	9,79	1,68	30	7,45	1,38
[gi]	13	8,31	3,83	26	11,35	1,58	30	9,09	2,01

Tabela 27: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras pré-vozeadas com relação à altura da vogal seguinte

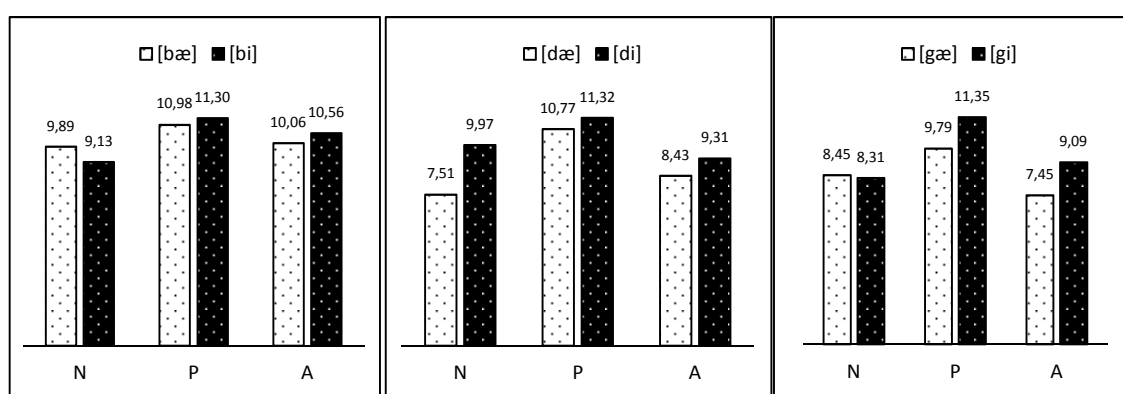


Gráfico 16: Média da duração relativa de VOT das oclusivas sonoras pré-vozeadas com relação à altura da vogal seguinte

Em todos os pontos de articulação, tivemos para as oclusivas sonoras a mesma configuração observada para as surdas. Para a oclusiva bilabial, os maiores valores de VOT foram flagrados diante da vogal alta [i] na produção dos grupos P e A, e diante da vogal baixa [æ] no grupo N. Para a oclusiva alveolar, todos os grupos apresentaram VOT mais longo nas oclusivas diante de [i]. E quanto às velares, os grupos P e A tiveram valores mais longos de VOT nas consoantes diante de [i], enquanto que no grupo N, as velares tiveram VOT mais longo diante de [æ]. Daí se pode supor que a influência que altura da vogal seguinte exerce para as oclusivas é a mesma, independentemente de serem surdas ou sonoras. No entanto, não se pode ir muito além quanto às considerações aqui realizadas, devido à escassez de dados.

Voltemo-nos agora às consoantes oclusivas com VOT de intervalo curto. Como exposto no início desta seção, nossa análise ficará prejudicada pela escassez e mesmo

ausência de ocorrências dessa natureza na produção do grupo P e, especialmente, do grupo A. A análise mais cuidadosa ficará por conta dos dados do grupo N, que apresentou um maior número de dados. Novamente, a tabela 28 traz as médias e o número de ocorrências para cada grupo por ponto de articulação. O gráfico 17 permite uma melhor visualização dos resultados.

Oclusivas sonoras de intervalo curto									
	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[bæ]	19	0,56	0,50	8	0,41	0,40	0	---	---
[bi]	14	0,87	0,42	3	0,21	0,25	0	---	---
[dæ]	19	1,16	0,37	7	1,21	0,63	2	1,06	0,01
[di]	22	1,46	1,63	4	2,52	0,59	1	0,68	---
[gæ]	17	1,7	0,61	1	1,67	---	0	---	---
[gi]	17	2,51	1,00	4	1,86	0,70	0	---	---

Tabela 28: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à altura da vogal seguinte

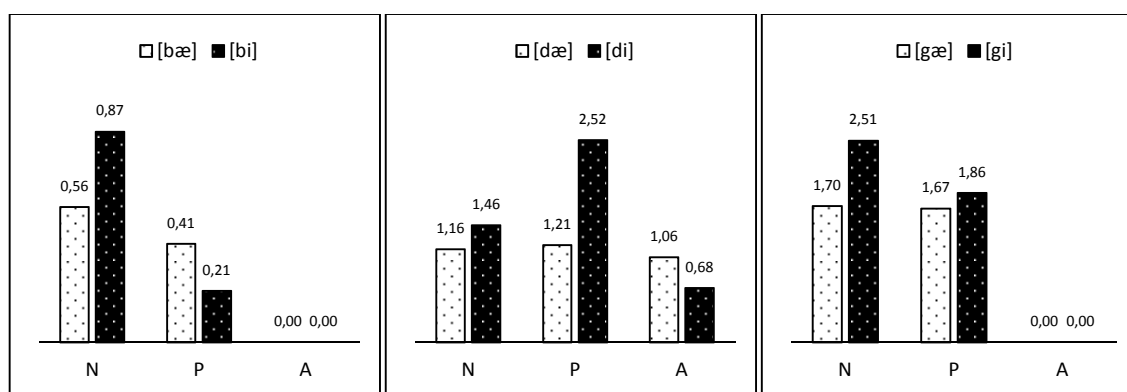


Gráfico 17: Média da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à altura da vogal seguinte

A expectativa quanto à produção das oclusivas sonoras com intervalo curto era a de que a influência da altura da vogal seguinte inspirasse comportamento semelhante àquele apresentado nas oclusivas surdas, já que ambas exibem VOT positivo, embora em níveis diferentes. No entanto, para o grupo N, a duração de VOT das oclusivas em todos os pontos de articulação foi maior diante da vogal alta [i], o que faz com que tais resultados se assemelhem mais àqueles dos grupos P e A para oclusivas surdas. Esse é, inclusive, o segundo ponto de identificação que flagramos entre as oclusivas surdas produzidas pelos falantes de PB e as oclusivas sonoras com intervalo curto produzidas pelos nativos. O primeiro foi flagrado quando analisamos o ponto de articulação: embora na produção das surdas, o grupo N tenha exibido maiores valores de VOT nas oclusivas alveolares e a diferença entre os três pontos não tenha sido significativa, quando da produção das oclusivas

sonoras com intervalo curto, tivemos gradação progressiva dos pontos bilabial, alveolar e velar e a diferença entre os pontos passou a ser maior. Isso nos leva a supor que oclusivas com VOT de intervalo curto estão sujeitas a um processo diferente daquele existente para oclusivas de intervalo longo.

Os grupos P e A também apresentaram diferença com relação aos resultados obtidos para as oclusivas surdas nos seguintes casos: no ponto bilabial, o grupo P teve VOT mais longo diante de vogal posterior e no ponto alveolar, o grupo A também apresentou VOT mais longo diante da posterior. Ressalte-se, entretanto, que o número de dados para esses grupos são realmente escassos para que possamos ir além em nossas considerações.

5.4 VOT E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL SEGUINTE

Schwarzhaupt (2012) relata que, embora a literatura comumente indique a altura da vogal seguinte como responsável por mudanças na duração do VOT (YAVAS; WILDERMUTH, 2006; YAVAS, 2008; FRANÇA, 2011), há indícios de que a posição ântero-posterior da vogal seguinte desempenhe um papel igualmente importante na duração do VOT. Alves e Dias (2010) apuram maiores valores de VOT diante da vogal anterior [i] do que diante de [u], o que, no entanto, não se confirma totalmente em Schwarzhaupt (2012).

Cumprido, portanto, analisar a relação entre a duração de VOT das oclusivas com relação à posição ântero-posterior da vogal seguinte em nossos dados. Analisamos aqui os dados que possuem oclusivas surdas diante da vogal anterior [i] contrastando-os com dados nos quais as oclusivas estão diante da vogal posterior [u]. Optamos por selecionar duas vogais breves (*lax*), que, em comparação com as tensas (*tense*), estão normalmente associadas a VOT mais curto nas oclusivas que as precedem (PORT; ROTUNNO, 1979). Entretanto, é possível que, ocasionalmente, esses segmentos venham a ser realizados como [i] e [u], especialmente entre o grupo A.

5.4.1 OCLUSIVAS SURDAS E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL

Na tabela 29 encontramos os valores relativos de VOT (%) e o desvio padrão apurado para as consoantes oclusivas surdas com relação à posição ântero-posterior da vogal seguinte. Os valores absolutos de VOT encontram-se no apêndice I. O gráfico 18 repete as médias para melhor visualização.

Oclusivas surdas e a posição ântero-posterior da vogal									
	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[pʊ]	30	4,88	1,64	30	3,21	1,32	30	1,53	0,75
[pɪ]	30	4,74	1,89	30	1,90	1,21	30	1,00	0,55
[tʊ]	30	5,12	1,68	30	3,79	1,50	30	1,97	1,09
[tɪ]	30	5,08	1,53	30	3,54	1,56	30	4,00	1,42
[kʊ]	30	5,19	1,15	30	4,43	2,12	30	2,97	1,47
[kɪ]	30	5,03	1,88	30	5,33	1,61	30	3,60	1,40

Tabela 29: Média e desvio padrão da duração relativa de VOT para as oclusivas surdas com relação à posição ântero-posterior da vogal seguinte.

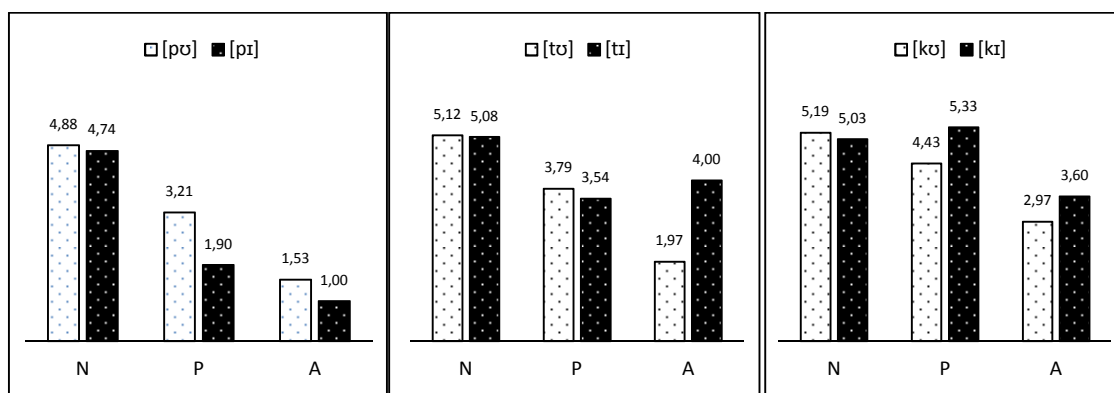


Gráfico 18: Média da duração relativa de VOT para as oclusivas surdas com relação à posição ântero-posterior da vogal seguinte.

Observamos que, para a oclusiva bilabial surda, a duração de VOT é maior diante da vogal posterior na produção dos três grupos, embora a diferença se acentue nos grupos de falantes de PB, especialmente no grupo P.

Para o ponto alveolar, a vogal posterior [ʊ] associa-se aos valores mais longos de VOT para os grupos N e P. Entretanto o grupo A exhibe valores consideravelmente mais longos de VOT nas oclusivas diante de [ɪ]. É preciso, no entanto, relativizar os resultados quanto à produção do grupo A para a sequência [tɪ], pois como já observado na seção 5.3.1, tal fato pode estar relacionado ao fenômeno da palatalização da oclusiva alveolar diante de [i]

e variantes, que ocorre no português paranaense e pode estar sendo transferido da L1 para a L2, resultando em valores de VOT equivocadamente mais longos.

No caso da oclusiva velar, os valores mais longos de VOT dos grupos P e A passam a ser associados à vogal anterior [ɪ], enquanto que o grupo N continua produzindo consoantes com maior VOT diante da vogal posterior [ʊ].

Assim, temos que a bilabial é sempre produzida com VOT mais longo diante da vogal posterior, a alveolar tem VOT mais longo também diante da posterior para os grupos N e P e a velar tem VOT mais longo diante da vogal anterior nos grupos P e A.

Mediante a condução do teste de Wilcoxon, com nível de significância de 0,05, apuramos diferença significativa em todos os pontos de articulação para o grupo A e nos pontos bilabial e velar para o grupo P. Os resultados seguem na tabela 30.

Wilcoxon para a posição da vogal			
	N	P	A
[pʊ] [pɪ]	Z = -0,729, p = 0,466	Z = -3,051, p = 0,002	Z = -3,234, p = 0,001
[tʊ] [tɪ]	Z = -0,068, p = 0,946	Z = -0,592, p = 0,554	Z = -4,304, p = 0,000
[kʊ] [kɪ]	Z = -0,068, p = 0,946	Z = -2,141, p = 0,032	Z = -2,163, p = 0,031

Tabela 30: Resultados do teste de Wilcoxon para oclusivas surdas quanto à posição ântero-posterior da vogal seguinte.

Os resultados encontrados coincidem em parte com aqueles verificados em Schwartzhaupt (2012). Na produção dos sujeitos norte-americanos, o autor constata VOT mais longo nas oclusivas bilabial e alveolar diante de vogal posterior, porém na velar o VOT é mais longo diante de vogal anterior. Em nosso estudo, o ponto posterior foi responsável por VOT mais longo nos três pontos de articulação do grupo N. Nos dados de falantes proficientes, houve coincidência entre nossos resultados e os encontrados por Schwartzhaupt (2012), pois para ambos os estudos foi verificado VOT mais longo para oclusivas diante de vogal posterior nos pontos bilabial e alveolar, mas diante de anterior no ponto velar.

5.4.2 OCLUSIVAS SONORAS E A POSIÇÃO ÂNTERO-POSTERIOR DA VOGAL

Iniciamos nossa análise com os resultados apurados para as consoantes oclusivas com VOT pré-vozeado. Os resultados seguem na tabela 31 e no gráfico 19. Como nos demais casos, os resultados com VOT absoluto podem ser encontrados no apêndice J.

Oclusivas sonoras com pré-vozeamento e a posição ântero-posterior da vogal

	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[bʊ]	15	10,1	1,26	26	12,03	3,80	30	10,74	2,90
[bɪ]	13	10,45	1,53	29	12,32	2,26	30	11,3	1,87
[dʊ]	13	9,52	1,68	23	11,44	2,49	30	8,73	2,68
[dɪ]	12	8,73	2,75	22	11,1	3,61	30	8,97	2,46
[gʊ]	13	6,83	3,26	24	13,12	2,71	30	8,91	2,04
[gɪ]	14	10,2	1,25	26	11,79	2,50	29	9,34	1,98

Tabela 31: Média e desvio padrão das oclusivas sonoras pré-vozeadas com relação à posição ântero-posterior da vogal seguinte.

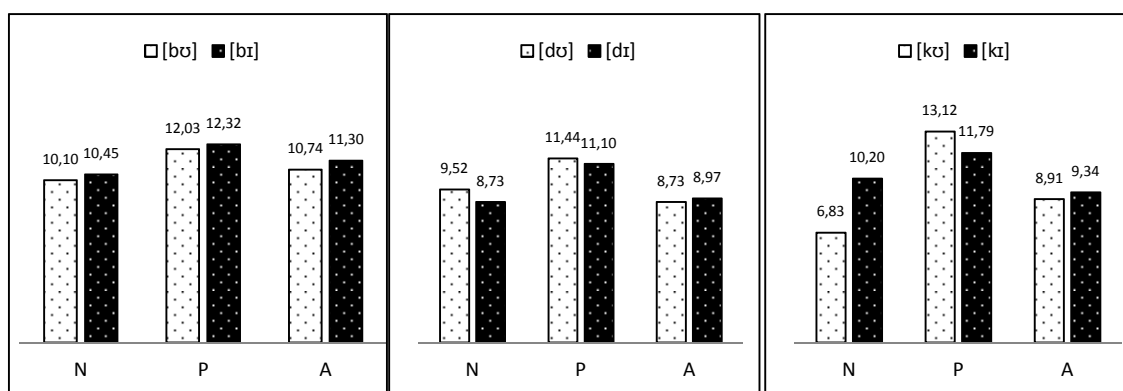


Gráfico 19: Média da duração relativa de VOT para oclusivas sonoras pré-vozeadas com relação à posição ântero-posterior da vogal seguinte.

Com relação à bilabial, verificamos que a duração mais longa de VOT ocorreu diante da vogal anterior para todos os grupos. Tal tendência se manteve para a alveolar nos grupos P e A, mas não no grupo N. Por fim, a vogal anterior também pareceu motivar duração mais longa de VOT nas velares dos grupos N e A, mas dessa vez, não no grupo P. Comparando tais resultados com aqueles observados para as surdas, houve discrepância quanto às bilabiais e velares nos grupos N e P e quanto às bilabiais no grupo A.

Constatamos, portanto, que a resposta obtida para as sonoras pré-vozeadas não coincidiu com a das surdas, ao contrário do que observamos quanto à altura da vogal seguinte, quando houve coincidência nos efeitos observados tanto para surdas quanto para sonoras pré-vozeadas em todos os pontos de articulação. Podemos supor que o vozeamento aparentemente causou diferentes repercussões na maneira como a posição ântero-posterior da vogal seguinte influenciou o VOT.

No que tange às oclusivas sonoras com intervalo curto, novamente nos encontramos diante da limitação de não possuir ocorrências em número suficiente que nos permita fazer generalizações sobre a influência da posição da vogal seguinte, especialmente com relação aos grupos P e A.

Diante disso, teceremos apenas algumas observações sobre a produção do grupo N, que contou com um número mais expressivo de *tokens* na categoria de sonoras com intervalo curto. Os resultados seguem na tabela 32 e no gráfico 20. Valores absolutos de VOT podem ser encontrados no apêndice K.

Oclusivas sonoras com VOT de intervalo curto e a posição ântero-posterior da vogal									
	N			P			A		
	Tokens	Média	DP	Tokens	Média	DP	Tokens	Média	DP
[bʊ]	15	1,28	0,71	4	0,72	0,25	0	---	---
[bɪ]	17	0,95	0,51	1	0	---	0	---	---
[dʊ]	17	1,3	0,70	7	1,1	0,49	0	---	---
[dɪ]	18	1,26	0,46	8	2,52	0,36	0	---	---
[gʊ]	17	2,11	0,50	6	2,2	0,52	0	---	---
[gɪ]	16	1,79	0,50	4	2	0,59	1	1,05	---

Tabela 32: Média e desvio padrão da duração relativa de VOT das oclusivas sonoras de intervalo curto com relação à posição ântero-posterior da vogal seguinte.

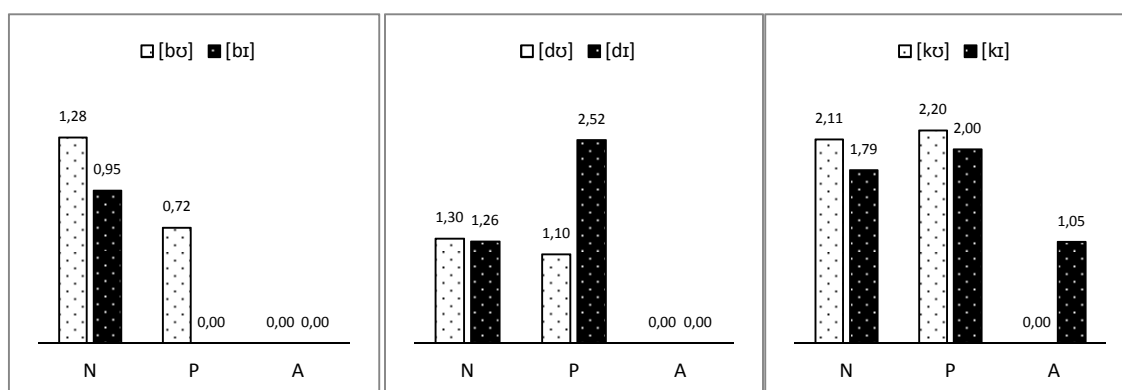


Gráfico 20: Média da duração relativa de VOT para oclusivas sonoras de intervalo curto com relação à posição ântero-posterior da vogal seguinte.

Para o grupo N, verificamos que a duração de VOT é mais longa diante da vogal posterior [ʊ] do que diante da anterior [ɪ] para todos os pontos de articulação. Esse foi o mesmo padrão observado para as oclusivas surdas do grupo N, porém não foi o mesmo com relação às sonoras pré-vozeadas, para as quais a vogal posterior só pareceu acarretar VOT mais longo quando diante da oclusiva alveolar. De forma geral, no entanto, nossos resultados quanto ao grupo N sugerem que a posição posterior da vogal seguinte é responsável por um aumento na duração do VOT da oclusiva. Tal aumento, não é aparentemente muito expressivo, o que só poderia ser confirmado de fato com um maior número de dados e a aplicação de testes estatísticos.

Para o grupo P, a duração de VOT também foi maior diante da vogal posterior para os pontos bilabial e velar, porém quanto ao ponto alveolar, a diferença foi consideravelmente

maior no VOT da oclusiva diante de vogal anterior. Finalmente, houve apenas uma ocorrência de sonora com VOT de intervalo curto na produção do grupo A, o que nos impede de avaliar o grupo para esta categoria.

5.5 CONCLUSÃO

Destacaremos agora as conclusões que consideramos as mais relevantes após a condução do experimento e a análise dos resultados obtidos.

Quanto ao nível de proficiência, observamos que a produção das oclusivas surdas por falantes de PB difere significativamente do estágio em que são aprendizes para o ponto em que são falantes proficientes da língua. Essa diferença se manifesta na produção de oclusivas com VOT de maior duração e que, portanto, aproximam-se mais da forma-alvo, representada aqui pelo grupo dos nativos norte-americanos. Além disso, notamos que as bilabiais são as oclusivas surdas que mais se distanciam da forma-alvo, o que nos leva a sugerir que a instrução explícita da aspiração do [p] seja priorizada para os falantes de PB.

Nas oclusivas sonoras, obtivemos dados de duas naturezas: oclusivas com VOT negativo, mais abundantes na produção dos falantes de PB, e oclusivas com VOT positivo de intervalo curto, mais recorrentes na produção dos nativos norte-americanos. Diante disso, resolvemos verificar se os dados de VOT positivo, fossem eles de intervalo curto ou de intervalo longo, apresentariam comportamento semelhante. Assim, ao compararmos a influência do ponto de articulação, constatamos que nos casos em que o VOT exibia intervalo curto, houve aumento progressivo na duração do VOT conforme o ponto se tornava mais posterior e essa diferença foi considerável da bilabial para a alveolar e da alveolar para a velar. Todavia, nas oclusivas com VOT de intervalo longo, não houve o mesmo aumento progressivo e tampouco a diferença significativa na duração do VOT entre os pontos.

Mais um indício de que as duas categorias de VOT positivo se comportam de maneira diferente foi encontrada quando analisamos a influência da altura da vogal seguinte. A princípio, nossa expectativa era a de que a influência da altura da vogal seguinte se desse de maneira semelhante tanto para as oclusivas com VOT de intervalo longo quanto de intervalo curto, já que ambas têm VOT positivo. Notamos, entretanto, que nos dados de intervalo longo, não houve diferença estatisticamente significativa na duração do VOT, embora tenha sido mais longo diante de vogal baixa para os pontos alveolar e velar, ao passo que nos dados

de intervalo curto, foi a vogal alta que motivou os maiores índices de VOT em todos os pontos e a diferença foi estatisticamente significativa, ao menos na produção dos aprendizes.

Isso nos leva a concluir que, além de haver diferenças evidentes entre o comportamento das oclusivas com VOT positivo e negativo, também é preciso que atentemos para as diferenças mais sutis entre oclusivas com VOT positivo de intervalo curto e longo. Flagramos, portanto, três diferentes padrões e é preciso que, em pesquisas futuras, investiguemos com mais profundidade os fatores que determinam como se dá a produção dessas três categorias de vozeamento.

6 CONSIDERAÇÕES FINAIS

Voltemos agora aos questionamentos lançados na seção 3.1, procurando esclarecê-los à luz dos resultados obtidos neste estudo.

Pergunta 1: O nível de proficiência na língua inglesa interfere na produção das oclusivas por falantes de PB?

Tanto para as oclusivas surdas quanto para as sonoras, nossos resultados apontam para uma relação positiva entre o grau de proficiência do aprendiz e sua capacidade crescente em aproximar-se dos sons da L2 (ZIMMER, 2004; FRANÇA, 2011).

H1: Haverá diferença estatisticamente significativa entre a produção do grupo A e do grupo N com relação à duração do VOT nas oclusivas surdas.

Esta hipótese foi confirmada, havendo diferença estatisticamente significativa entre a produção do grupo A e do grupo N para todos os pontos de articulação. Como esperado, a duração de VOT do grupo N foi mais longa do que a do grupo A. Tais resultados corroboram os achados do piloto e pesquisas anteriores que atestam que falantes nativos de PB produzem as oclusivas surdas iniciais do inglês com níveis de aspiração inferiores aos de um falante nativo da língua (MAJOR, 1992; SANCIER; FOWLER, 1997; ROCCA, 2003; ZIMMER, 2004; COHEN, 2004; REIS; NOBRE-OLIVEIRA, 2007; ALVES, 2011; FRANÇA, 2011; SCHWARTZHAUPT, 2012).

H2: Não haverá diferença significativa entre a produção do grupo P e do grupo N com relação à duração do VOT nas oclusivas surdas.

Não houve diferença estatisticamente significativa entre os pontos alveolar e velar, mas houve para o ponto bilabial, o que faz com que esta hipótese seja apenas parcialmente confirmada. Em todos os casos, na produção do grupo N a duração relativa de VOT foi mais longa do que na produção do grupo P.

De toda forma, destacamos a proximidade existente entre a produção de nativos e falantes proficientes. Tendo em vista que falantes de PB em níveis mais avançados de aprendizagem aparentemente acabam por ter sucesso em coordenar os gestos envolvidos na produção da alveolar e da velar aspiradas, é possível supor que necessitem de instrução explícita somente para aprimorar a produção da bilabial e superar as dificuldades flagradas nesse processo de reorganização gestual.

H3: Haverá diferença significativa entre a produção do grupo A e do grupo P com relação à duração do VOT nas oclusivas surdas.

Também houve confirmação desta hipótese, havendo diferença estatisticamente significativa entre o grupo P e o grupo A em todos os pontos de articulação, sendo a duração relativa de VOT mais longa para o grupo P. Tal fato sugere o sucesso no processo de reorganização dos gestos envolvidos para o aumento da duração de VOT entre níveis iniciais de aprendizagem da língua e seu domínio mais amplo. Sob a perspectiva da Fonologia Gestual, é possível dar conta desse fato constatando que os falantes de PB aprendem, no decorrer do processo de aquisição, a coordenar o faseamento dos gestos articulatórios de desvozeamento laríngeo e constrição oral, responsáveis pela mudança de VOT, a fim de aproximar-se do padrão fonético-fonológico da L2.

Quanto às oclusivas sonoras, o que mais nos permitiu compreender as diferenças entre os grupos não foi propriamente a duração do VOT, mas a flutuação na realização dos dados. Tal conjuntura nos foi mais relevadora, pois observamos que parece haver uma relação positiva entre o nível de proficiência e a proporção de dados produzidos com VOT de intervalo curto. Inversamente, quanto mais dados produzidos com VOT pré-vozeado, menor parece ser a proficiência do sujeito. Além disso, os dados de intervalo curto também apontam para uma duração maior de VOT conforme o nível de proficiência aumenta. Por outro lado, não há nenhuma relação aparente entre a duração do VOT nas oclusivas pré-vozeadas e o nível de proficiência do sujeito.

Além disso, é possível pressupor que o processo de aquisição do padrão de vozeamento das oclusivas sonoras do inglês não passa necessariamente por uma evolução tão clara quanto a testemunhada para as oclusivas surdas. Isso pode se dever ao fato de que a pista da aspiração seja mais perceptualmente robusta e, por consequência, mais simples de ensinar e adquirir. Já a variação de produção característica das oclusivas sonoras aliada à maior dificuldade na percepção de diferentes graus de vozeamento, faz com que a instrução explícita desse aspecto seja dificultada ou mesmo impossibilitada. Por outro lado, não haveria, em princípio, motivação para que o falante de PB modificasse sua produção de a fim de adequá-la à apresentada pelos nativos, já que, de uma forma ou de outra, ele já realiza as sonoras seguindo os padrões da L2, embora as proporções entre realizações com VOT de intervalo curto e pré-vozeamento sejam diferentes.

Para que possamos chegar a conclusões como essa, faz-se necessário que nosso modelo de análise contemple o detalhe fonético fino. Nesse sentido, o Modelo de Assimilação Perceptual (*Perceptual Assimilation Model*, PAM), proposto por Catherine Best (1994, 1995)

fornece maior poder explicativo do que um modelo que tome o fonema como unidade distintiva, pelo fato de tomar como primitivo de análise o gesto articulatório e prever o faseamento entre os gestos na constituição dos sons da fala. Podemos, por exemplo, justificar as diferentes estratégias de vozeamento das consoantes sonoras por meio do faseamento dos diversos gestos envolvidos na articulação dessas consoantes.

Podemos ainda justificar a distância entre a produção do grupo N e do grupo P, por meio do pressuposto de que o falante de PB aprende a ignorar detalhes fonéticos que não são importantes na percepção dos sons da L2, voltando-se apenas a perceber aquilo que é relevante. Como ele tenderá a assimilar os sons da L2 de acordo com as pistas acústicas da L1, parece natural que o falante proficiente produza mais dados com VOT pré-vozeado, especialmente quando esse tipo de realização também faz parte do repertório de falantes nativos de inglês.. Assim, parece-nos que o falante proficiente de inglês está, com sucesso, sintonizando o sistema de percepção a fim de perceber detalhes fonéticos relevantes para a aquisição da L2 e ignorando detalhes irrelevantes.

Pergunta 2: O ponto de articulação interfere nos índices de VOT das oclusivas por nativos e falantes de PB?

Observamos clara interferência do ponto de articulação na duração do VOT para os três grupos estudados; no entanto, a premissa de que quanto mais posterior o ponto de articulação, mais longa é a duração de VOT nas oclusivas surdas (LISKER; ABRAMSON, 1964; CHO; LADEFOGED, 1999; YAVAS, 2008; STEIN, 2011; FRANÇA, 2011; SCHWARTZHAUPT, 2012) não foi completamente confirmada em nossa pesquisa.

H4: Haverá progressivo aumento na duração de VOT de [p], [t] e [k] e a diferença será significativa para todos os pontos, na produção dos grupos N, P e A.

Nossa quarta hipótese foi parcialmente confirmada, pois se concretizou apenas com relação à produção dos grupos P e A, para os quais houve aumento progressivo na duração de VOT e a diferença entre os pontos foi estatisticamente significativa do ponto bilabial para o alveolar e do alveolar para o velar em ambos os grupos. Todavia, o grupo N não apresentou diferença significativa na duração do VOT entre nenhum dos pontos.

Vale destacar ainda quanto às oclusivas sonoras com pré-vozeamento, a constatação de que quanto mais anterior o ponto de articulação, maior a duração do VOT, resultado que é justamente o oposto do encontrado para as oclusivas surdas. Tal resultado coincide com o encontrado por Lisker e Abramson (1964) para oclusivas do inglês em início de frase e

também com Melo et al. (2011) para o PB. Entretanto, a diferença entre os pontos não parece ser significativa em nenhum dos grupos, o que não pode ser comprovado com a realização de análise estatística devido à escassez de dados.

Pergunta 3: A altura da vogal seguinte interfere nos índices de VOT das oclusivas por nativos de inglês e falantes de PB?

Os resultados apontam para uma aparente influência da altura da vogal seguinte na duração do VOT de falantes nativos e falantes de PB, embora isso não se confirme totalmente quando da condução de testes estatísticos. Para as oclusivas surdas, foi observado um padrão único nos grupos de falantes de PB (grupos P e A), que, todavia, não se manteve no grupo N. Nos primeiros, o VOT das consoantes oclusivas foi sempre maior diante da vogal alta, ao passo que no segundo o VOT foi mais longo diante da vogal posterior para os pontos alveolar e velar. O mesmo resultado foi observado para as oclusivas sonoras, de onde se pode supor que a influência que altura da vogal seguinte exerce para as oclusivas é a mesma, independentemente de serem surdas ou sonoras.

Além disso, o ponto de articulação parece não alterar substancialmente a maneira como a altura da vogal influencia a oclusiva, já que, de forma geral, o VOT é mais alto na oclusiva diante de vogal alta e isso tem sido atestado em outros estudos (YAVAS; WILDERMUTH, 2006; FRANÇA, 2011). Todavia, se considerarmos somente a produção do grupo N, essa suposição precisará ser revista. O estudo de Schwartzhaupt (2012) também não obteve uniformidade em todos os dados, já que apesar de o VOT ser mais longo em oclusivas diante de vogal alta na maioria dos dados, no grupo dos falantes proficientes, a oclusiva alveolar teve VOT mais longo diante da vogal baixa.

H5: A duração de VOT das oclusivas surdas diante da vogal alta [i] será significativamente maior do que diante da vogal baixa [æ] na produção dos grupos N, P e A, mesmo ao considerar os pontos de articulação em separado.

Essa hipótese pode ser confirmada somente com relação à produção do grupo A, quando houve diferença significativa para todos os pontos de articulação e as oclusivas surdas exibiram VOT mais longo diante da vogal alta.

Pergunta 4: A posição ântero-posterior da vogal seguinte interfere nos índices de VOT das oclusivas por nativos de inglês e falantes de PB?

Nossos resultados sugerem que existe influência da posição ântero-posterior da vogal seguinte na duração do VOT das oclusivas para nativos de inglês e falantes de PB. Dessa vez, no entanto, o ponto de articulação parece alterar a maneira como se dá tal influência. Os

resultados encontrados coincidem em parte com aqueles verificados em Schwartzhaupt (2012), pois para ambos os estudos foi verificado VOT mais longo para oclusivas diante de vogal posterior nos pontos bilabial e alveolar, mas diante de anterior no ponto velar. O argumento proposto pelo autor e que se adequa à maior parte de nossos dados é baseado em uma premissa de Cho e Ladefoged (1999), segundo a qual quanto mais súbita a queda da pressão transglotal na produção da oclusiva, mais curto será o VOT. Daí decorre que quando temos uma sequência formada pela bilabial [p], cuja articulação se dá em um ponto anterior do trato, e a vogal [ʊ], cuja articulação se dá em um ponto posterior do trato, os articuladores se moverão mais lentamente para a produção dessa sequência e, conseqüentemente a queda na pressão transglotal será menos súbita, resultando em VOT mais longo se comparado com a sequência [pɪ]. Justifica-se, portanto, que nos pontos bilabial e alveolar a vogal posterior cause VOT mais longo, enquanto que no ponto velar, a vogal anterior provoque o mesmo efeito.

A resposta obtida para as sonoras pré-vozeadas não coincidiu com a das surdas, ao contrário do que observamos quanto à altura da vogal seguinte, quando houve coincidência nos efeitos observados tanto para surdas quanto para sonoras pré-vozeadas em todos os pontos de articulação. Podemos supor que o vozeamento aparentemente causou diferentes repercussões na maneira como a posição ântero-posterior da vogal seguinte influenciou o VOT.

H6: A duração de VOT das oclusivas surdas diante da vogal anterior [ɪ] será significativamente maior do que diante da vogal posterior [ʊ] na produção dos grupos N, P e A, mesmo ao considerar os grupos de articulação em separado.

Essa hipótese não pode ser completamente confirmada, valendo para 3 das 5 ocasiões em que se observou diferença significativa; com relação às oclusivas alveolares e velares do grupo A e com relação às velares do grupo P. Houve diferença significativa em mais duas instâncias, porém com VOT mais longo diante da vogal posterior: para as bilabiais do grupo A e do grupo P.

O fato de que não houve diferença significativa para o grupo N com relação à posição ântero-posterior da vogal, bem como quanto à altura da vogal seguinte e ponto de articulação nos leva a supor que índices mais longos de VOT tendem a mascarar os efeitos exercidos por essas variáveis. Tal suposição é reforçada pela constatação de que, relativamente ao grupo N, ao contrário do que esperávamos, houve relativa discrepância entre os dados de intervalo curto e os das oclusivas surdas. Apesar de ambos contarem com VOT positivo, a considerável diferença entre a duração do VOT e, conseqüentemente, entre os níveis de aspiração das duas

produções, pareceu afetar de forma diferente fatores como o ponto de articulação e a influência exercida pela altura da vogal seguinte. É necessário, no entanto, que se façam pesquisas com um maior número de dados a fim de que essas questões sejam esclarecidas.

Finalmente, cabe encerrar nossa pesquisa ponderando sobre os resultados obtidos, seus possíveis efeitos para o ensino da língua inglesa e as possibilidades que se abrem para futuros trabalhos.

Partimos da motivação inicial de observar mais detalhadamente uma particularidade de pronúncia presente no LFC que, segundo Jenkins (2000) é indispensável para a inteligibilidade entre não nativos, a saber, a aspiração que antecede as oclusivas [p], [k] e [t] quando ocorrem em posição inicial de sílaba tônica. Entretanto, pareceu-nos relevante estendermos nossa análise às oclusivas sonoras, com a expectativa de contribuir para novos vislumbres e uma compreensão mais aprofundada do fenômeno em foco.

Durante nossa trajetória, a maior limitação com a qual nos deparamos foi justamente com relação à análise das oclusivas sonoras, que tiveram sua realização dividida em ocorrências de duas naturezas. Ao mesmo tempo em que essa flutuação nos revelou algumas questões interessantes, também acabou ocasionando escassez de dados e, conseqüentemente, impossibilitou a execução de análises estatísticas. Diante disso, julgamos ser especialmente relevante que estudos futuros foquem a produção de consoantes oclusivas sonoras.

Além disso, consideramos oportuno que futuramente seja conduzida a análise concomitante da produção de monolíngues de PB falantes do dialeto curitibano, de modo a poder contrastar tais dados com aqueles produzidos por falantes bilíngues e verificar um possível condicionamento dialetal sobre esse evento acústico.

Em termos de contribuição para o ensino do inglês, esperamos ter propiciado ferramentas para que os professores da língua estejam melhor amparados na tarefa de levar os falantes brasileiros a refinarem sua pronúncia, adotando uma perspectiva de aquisição de L2 como um processo dinâmico, segundo o qual a produção do aprendiz não é vista em termos de presença ou ausência de um dado aspecto, mas levando em conta a aproximação que se faz da forma-alvo através da reorganização dos gestos articulatórios.

Esperamos, enfim, que nosso trabalho tenha contribuído para uma melhor compreensão sobre a aquisição das consoantes oclusivas por falantes de PB e que abra caminho para que novas pesquisas na área sejam realizadas a fim aprofundar a reflexão acerca da interação entre os sistemas linguísticos da L1 e da L2 à luz de modelos dinâmicos de análise fonológica.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMSON, A.S.; LISKER, L. *Voice Onset Time in stop consonants: acoustic analysis and synthesis*. In: 5th International Congress of Acoustics, Liège, 1965, p. 1-4.

_____. *Discriminability along the voicing continuum: cross-language tests*. In: Proceedings of the 6th International Congress of Phonetics Sciences, Prague, 1970, p. 569-573.

_____. *Voice timing in Korean stops*. In: Proceedings of the 7th International Congress of Phonetic Sciences, Montreal, 1971, p. 439-446.

ALVES, M. A. *Production of English and Portuguese voiceless stops by Brazilian EFL speakers*. Dissertação (Mestrado em Letras). Florianópolis: Universidade Federal de Santa Catarina, 2011.

ALVES, M.A.; DIAS, E.C.O. *Estudo da produção do VOT em plosivas não-vozeadas diante de vogal alta anterior e posterior no português brasileiro*. In: Anais do IX Encontro do Celsul, Palhoça, 2010.

ALVES, M.A.; SEARA, I.C.; PACHECO, F.S.; KLEIN, S.; SEARA, R. *On the voiceless aspirated stops in Brazilian Portuguese*. In: 8th International Conference, PROPOR, Computational Processing of Portuguese Language. Springer, 5190, 2008, p.248-251.

BEST, C. T. *The emergence of native-language phonological influence in infants: A perceptual assimilation model*. In: NUSSBAUM, H.; GOODMAN, J.; HOWARD, C. *The transition from speech to spoken words: The development of speech perception*. Cambridge, MA: MIT Press, p. 167–224, 1994.

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

_____; TYLER, M.D. *Nonnative and second-language speech perception: Commonalities and complementarities*. In MUNRO M. J.; BOHN O.-S. (eds.) *Second language speech learning: The role of language experience in speech perception and production* Amsterdam: John Benjamins, p. 13-34, 2007.

____; HALLÉ, P.A. *Perception of initial obstruent voicing is influenced by gestural organization*. In: *Journal of Phonetics*, v. 38, p. 109-126, 2010.

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*. In: *Calidoscópico (UNISINOS)*, v. 9, p. 28-40, 2011.

BOERSMA, P; WEENINK, D. *Praat: doing phonetics by computer*. Amsterdam: University of Amsterdam, 2012.

BOHN, O.S.; FLEGE, J. E. *Perceptual switching in Spanish/English bilinguals*. *Journal of Phonetics*. 21, p. 267-290, 1993.

BONATTO, M.T.R.L., *Vozes infantis: a caracterização do contraste de vozeamento das consoantes plosivas do português brasileiro na fala de crianças de 3 a 12 anos*. Tese (Doutorado em Linguística Aplicada e Estudos da Linguagem). São Paulo: Pontifícia Universidade Católica de São Paulo, 2007.

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

_____. *Articulatory gestures as phonological units*. In: *Phonology*, v. 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*, v.49, p. 155-180, 1992.

CARAMAZZA, A.; YENI-KOMSHIAN, G.H.; ZURIF, E.B.; CARBONE, E. *The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals*. In: *The Journal of the Acoustical Society of America*, v. 54, n. 2, p. 421-428, 1973.

- CHEN, L.; CHAO, K.; PENG, J. *VOT productions of word-initial stops in Mandarin and English: a cross-language study*. In: Proceedings of the 19th Conference on Computational Linguistics and Speech Processing, 2007, p. 303-317.
- CHO, T.; LADEFOGED, P. *Variation and universals in VOT: evidence from 18 languages*. *Journal of Phonetics*, v. 27, 1999, p. 207-227.
- CHOMSKY, N. *Syntactic structures*. The Hague: Mouton, 1957.
- _____. *Aspects of the theory of syntax*. The MIT Press, Cambridge: MA, 1965.
- _____; HALLE, M. *The sound pattern of English*. The MIT Press, 1968.
- COHEN, G.V. *The VOT dimension: a bi-direction experiment with English and Brazilian Portuguese stops*. (Mestrado em Letras). Florianópolis: Universidade Federal de Santa Catarina, 2004.
- CORDER, S. *The significance of learners' errors*. In: *International Review of Applied Linguistics*. v.5, p.161-170, 1967.
- CRYSTAL, D. *English as a global language*. 2nd Edition. Cambridge: Cambridge University Press, 2003.
- DAUER, R. M. *The Lingua Franca Core: a new model for pronunciation instruction?* *Tesol Quarterly*. Volume 39/3. September, 2005
- DERWIN, T.M; MUNRO, M.J. *Second language accent and pronunciation teaching: a research-based approach*. In: *Tesol Quarterly*. v. 39, n. 3, p. 379-397, September, 2005.
- DOCHERTY, G. J. *The timing of voicing in British English obstruents*. Berlin: Foris, 1992.
- EDWARDS, H. T. *Applied phonetics: the sounds of American English*. San Diego: Singular Publishing Group, 1992.
- ELLIS, R. *Second language acquisition*. Oxford: Oxford University Press, 1997.

ELLIS, N.C. *Constructions, chunking, and connectionism: the emergence of second language structure*. In: DOUGHTY, C; LONG, M. (Eds.) *The handbook of second language acquisition*. Oxford: Blackwell Publishing, p. 63-103, 2003.

ELMAN, J.; BATES, E.A.; JOHNSON, M.; KARMILOFF-SMITH, A.; PARISI, D.; PLUNKETT, K. *Rethinking innateness: A connectionist perspective on development*. MIT Press: London, 1996.

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

_____. *Factors affecting degree of perceived foreign accent in English sentences*. In: *The Journal of the Acoustical Society of America*. v. 54, n. 2, p.70-79, 1988.

FLEGE, J.E. *Age of learning affects the authenticity of voice-onset-time (VOT) in stop consonants produced in a second language*. *Journal of the Acoustical Society of America*, v. 89, p.395-411, 1991.

_____. *Second-language speech learning: Theory, findings, and problems*. In: *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, edited by W. Strange Timonium, MD: York Press, p. 233–272, 1995.

FLEGE, J. E.; EEFING, W. *Production and perception of English stops by native Spanish speakers*. In: *Journal of Phonetics* v.15, p. 67-83, 1987.

_____. *Imitation of a VOT continuum by native speakers of English and Spanish: evidence for phonetic category formation*. In: *The Journal of the Acoustical Society of America*. v. 83, n. 2, p. 729-740, 1988.

FLEGE, J. E.; HILLENBRAND, J. *Limits on phonetic accuracy in foreign language speech production*. In: *The Journal of Acoustical Society of America* v.76, n. 3, p. 708-721, September, 1984

FLEGE, J. E.; PORT, R. *Cross-language phonetic interference: Arabic to English*. In: *Language and Speech*, v. 24, part 2, p. 125-146, 1981.

FRANÇA, K. V. *A aquisição da aspiração das plosivas surdas do inglês por falantes de português brasileiro: implicações teóricas decorrentes de duas diferentes formas de descrição dos dados*. Dissertação (Mestrado em Letras). Pelotas: Universidade Católica de Pelotas, 2011.

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

GIMSON, A.C. *An introduction to the pronunciation of English*. 3.ed., London: Edward Arnold, 1980.

GÓSY, M. *The VOT of the Hungarian voiceless plosives in words and in spontaneous speech*. In: *International Journal of Speech Technology*, v. 4, p. 75-85, 2001.

HERNANDORENA, C.L.M. *Introdução à teoria fonológica*. In: BISOL, L. (org.) *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1999.

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

JENKINS, J. *Rethinking phonology in teacher education*. In: VIEWZ – Vienna English Working Papers. v.7, n.1, p. 40-46, June 1998.

_____. *The phonology of English as an international language*. Oxford: Oxford University Press, 2000.

KEATING, P.A.; MIKOS, M.J.; GANONG III, W.F. *A cross language study of range of voice onset time in the perception of initial stop voicing*. In: *Journal of Acoustical Society of America*. v.70, n.5, p. 1261-1271, November, 1981.

KENT, R.D.; READ, C. *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group, 1992.

KHATTAB, G. *VOT production in English and Arabic bilingual and monolingual children*. In: NELSON, D.; FOULKES (eds.) *Leeds Working Papers in Linguistics* 8, p.95-122, 2002.

KLEIN, S. *Estudo do VOT no português brasileiro*. Dissertação (Mestrado em Letras). Florianópolis: Universidade Federal de Santa Catarina, 1999.

LADEFOGED, P. *A Course in Phonetics*. 4. ed. Boston: Heinle & Heinle, 2001.

LADEFOGED, P.; MADDIESON, I. *The sounds of the world's languages*. Massachusetts: Blackwell Publishing, 1996.

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

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

LARSEN-FREEMAN, D.; LONG, M.H. *An introduction to second language acquisition research*. London: Longman, 1991.

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

LISKER, L.; ABRAMSON A. *A cross-language study of voicing in initial stops: acoustical measurements*. In: *Words*, v. 20, p.384-422, 1964.

_____. *Some effects of context on voice onset time in English stops*. In: *Language and Speech*, v.10, p. 1-28, 1967.

MASSINI-CAGLIARI, G. *Acento e ritmo*. São Paulo: Contexto, 1992.

MAJOR, R.C. *Stress and rhythm in Brazilian Portuguese*. In: *Language*, v. 61, n.2, p. 259-282, 1985.

MAJOR, R. C. *Losing English as a first language*. In: The Modern Language Journal, v. 76, p.190-208, 1992.

MELO, R. M. et al . *Caracterização acústica da sonoridade dos fones plosivos do português brasileiro*. Rev. CEFAC, São Paulo, 2011 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-18462011005000143&lng=en&nrm=iso>. access on 18 Jan. 2013. Epub Dec 22, 2011. <http://dx.doi.org/10.1590/S1516-18462011005000143>.

ÖGÜT F, KILIÇ M, ENGIN E, MIDILLI R. *Voice onset time for Turkish stop consonants*. In: Speech Commun, v. 48, p. 1094-1099, 2006.

OHALA, J. J. *Articulatory constraints on the cognitive representation of speech*. In: MYERS, T.; LAVER, J.; ANDERSON (eds.) The cognitive representation of speech, North-Holland Publishing Company, Amsterdam/New York/Oxford, p 111-122, 1981.

PORT, R.; ROTUNNO, R. *Relation between voice onset time and vowel duration*. In: The Journal of the Acoustical Society of America, v. 66, n. 3, p. 654-662, 1979.

REIS, M.; NOBRE-OLIVEIRA, D. *Effects of perceptual training on the identification and production of English voiceless plosives aspiration by Brazilian EFL learners*. In: New Sounds: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech, Florianópolis: 2007, p. 398-407.

RINGEN, C.; KULIKOV, V. *Voice Onset Time in Russian*. In: 16th Annual Mid-Continental Workshop on Phonology. Chicago, 2010.

ROCCA, P.D.A. *O desempenho de falantes bilíngues: evidências advindas da investigação do VOT de oclusivas surdas do inglês e do português*. In: D.E.L.T.A, 19, p. 303-328, 2003.

SAERENS, M., SERNICLAES, W.; BEECKMANS, R. *Acoustic versus contextual factors in stop voicing perception in spontaneous French*. In: Language and Speech, v. 32, n. 4, p. 291-314, 1989.

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

SCHWARTZHAUPT, B. M. *Factors influencing Voice Onset Time: analyzing Brazilian Portuguese, English and interlanguage data*. Monografia (Graduação em Letras). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2012.

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

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

SKINNER, B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts. 1957.

STEIN, C. C. *Adequação dos gestos articulatórios típicos das consoantes oclusivas no inglês ao português brasileiro*. In: VII Congresso Internacional da Abralín, 2011, Curitiba-PR. In: *Anais do VII Congresso Internacional da Abralín*, v. 1, Curitiba, 2011, p. 802-815.

TORIBIO, A. J., BULLOCK, B. E., BOTERO, C. G.; DAVIS, K.A. *Perseverative phonetic effects in bilingual code-switching*. In: GESS R. S.; RUBIN E. J. (eds.), *Theoretical and experimental approaches to Romance linguistics*, p. 291-306, Salt Lake City: John Benjamins, 2005.

VELOSO, J. *The role of consonantal duration and tenseness in the perception of voicing distinctions of Portuguese stops*. In: *Proceedings of the XIII International Congress of Phonetic Science*, Stockholm, 1995, p. 266-269.

XU, Y. *In defense of lab speech*. In: *Journal of Phonetics*, v. 38, p. 329-336, 2010.

YAVAS, M. *Factors Influencing the VOT of English Long Lag Stops and Interlanguage Phonology*. In: RAUBER, A. S.; WATKINS, M; BATISTA, B. O. (eds.) *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*. Universidade Federal de Santa Catarina. Florianópolis, 2008, p. 492-498.

_____; WILDERMUTH, R. The effects of place of articulation and vowel height in the acquisition of English aspirated stops by Spanish speakers. *IRAL* v. 44 n° 3, p. 251-263, 2006.

ZIMMER, M. C. *A desaspiração de plosivas surdas iniciais na transferência do conhecimento fonético do português brasileiro (L1) para o inglês (L2)*. In: Anais do 6º Encontro do Celsul. Florianópolis, 2004.

ZIMMER, M.C.; ALVES, U.K. *On the status of terminal devoicing as an interlanguage process among Brazilian learners of English*. In: Ilha do Desterro. v.55, p. 41-62, 2008.

_____; SILVEIRA, R. ALVES, U.K. *Pronunciation instruction for Brazilians: bringing theory and practice together*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2009.

ZLATIN, M. *Voicing contrast: perceptual and productive voice onset time characteristics of adults*. In: The Journal of the Acoustical Society of America, v. 56, n. 3, p. 981- 994, 1974.

8 APÊNDICES

APÊNDICE A

INSTRUÇÕES PARA O GRUPO A (APRENDIZES DE INGLÊS)

Serão mostradas a você 42 frases curtas em inglês, que deverão ser lidas uma após a outra, assim que aparecerem na tela do computador.

Haverá uma pequena pausa e, em seguida, você lerá as mesmas frases apenas em uma ordem diferente. No total, haverá cinco repetições.

Se você cometer qualquer erro, como gaguejar ou tossir, você deverá repetir apenas aquela frase no final da sessão.

Você poderá solicitar uma pausa, caso julgue necessário.

É importante que você mantenha sua boca bastante próxima ao microfone, caso contrário o sinal não será captado de maneira clara.

Por favor, desligue seu celular, pois mesmo em modo silencioso, ele pode produzir ruído de fundo e prejudicar a gravação.

Obrigada pela sua participação!

DADOS DO PARTICIPANTE

Nome: _____

Idade: _____

Local de nascimento: _____

Idade em que começou a aprender inglês: _____

Possui algum problema de fala ou audição? ()sim ()não

Já esteve em algum país de língua inglesa? ()sim () não Por quanto tempo? _____

Telefone: _____

E-mail: _____

Data da gravação: _____

APÊNDICE B**INSTRUÇÕES PARA O GRUPO P (FALANTES PROFICIENTES)**

Serão mostradas a você 42 frases curtas em inglês, que deverão ser lidas uma após a outra, assim que aparecerem na tela do computador.

Haverá uma pequena pausa e, em seguida, você lerá as mesmas frases apenas em uma ordem diferente. No total, haverá cinco repetições.

Se você cometer qualquer erro, como gaguejar ou tossir, você deverá repetir apenas aquela frase no final da sessão.

Você poderá solicitar uma pausa, caso julgue necessário.

É importante que você mantenha sua boca bastante próxima ao microfone, caso contrário o sinal não será captado de maneira clara.

Por favor, desligue seu celular, pois mesmo em modo silencioso, ele pode produzir ruído de fundo e prejudicar a gravação.

Obrigada pela sua participação!

DADOS DO PARTICIPANTE

Nome: _____

Idade: _____

Local de nascimento: _____

Idade em que começou a aprender inglês: _____

Tempo que leciona inglês: _____

Certificados internacionais: ()sim ()não Quais? _____

Possui algum problema de fala ou audição? ()sim ()não

Já esteve em algum país de língua inglesa? ()sim ()não Por quanto tempo? _____

Telefone: _____

E-mail: _____

Data da gravação: _____

APÊNDICE C

INSTRUCTIONS FOR GROUP N (NATIVE SPEAKERS)

You will be shown a group of 42 sentences which must be read one after the other, as soon as they appear on the screen.

Then, there will be a pause, and after that, you will read the same sentences in a different order. In total, there will be five repetitions.

If you commit a mistake, such as stuttering or coughing, you will be asked to repeat only that sentence in the end of the session.

You may ask for a pause during the recording session if you need it.

It's important to keep your mouth very close to the microphone; otherwise the signal may not be clear enough.

Please, turn off your cell phone because even when it is in silent mode it may produce background noise and spoil the recording.

Thanks for your participation!

PARTICIPANT DATA

Name: _____

Age: _____

E-mail address: _____

Place of birth: _____

Do you have any speech or hearing problems? ()yes ()no

How long have you been in Brazil? _____

How long have you been studying Portuguese? _____

Date of recording: _____

APÊNDICE D

Oclusivas surdas – duração absoluta

Consoante	[p]		[t]		[k]	
	Tokens	Média	Tokens	Média	Tokens	Média
Grupo N	210	50,94	210	58,03	210	56,50
Grupo P	210	23,43	210	41,29	210	50,62
Grupo A	210	12,58	210	23,83	210	34,74

APÊNDICE E

Oclusivas sonoras com pré-vozeamento – duração absoluta						
Consoante	[b]		[d]		[g]	
	Tokens	Média	Tokens	Média	Tokens	Média
Grupo N	88	-102,44	83	-89,93	89	-84,42
Grupo P	181	-127,58	166	-120,94	183	-119,73
Grupo A	202	-122,76	206	-105,05	208	-99,89

APÊNDICE F

Oclusivas sonoras com intervalo curto – duração absoluta						
Consoante	[b]		[d]		[g]	
	Tokens	Média	Tokens	Média	Tokens	Média
Grupo N	122	8,79	127	12,83	121	21,61
Grupo P	29	6,14	44	15,31	27	22,47
Grupo A	8	2,00	4	11,67	2	14,00

APÊNDICE G

Oclusivas surdas e a altura da vogal						
	N		P		A	
	Tokens	Média	Tokens	Média	Tokens	Média
[pæ]	30	54,87	30	21,57	30	7,50
[pi]	30	50,73	30	25,07	30	12,93
[tæ]	30	63,70	30	38,90	30	18,03
[ti]	30	67,63	30	50,67	30	29,23
[kæ]	30	59,83	30	49,97	30	28,37
[ki]	30	59,17	30	55,27	30	45,90

APÊNDICE H

Oclusivas sonoras pré-vozeadas e a altura da vogal						
	N		P		A	
	Tokens	Média	Tokens	Média	Tokens	Média
[bæ]	11	-99,22	22	-118,48	30	-113,17
[bi]	16	-94,72	27	-127,67	30	-125,27
[dæ]	11	-75,47	23	-118,81	28	-97,79
[di]	8	-105,42	26	-124,70	29	-117,73
[gæ]	13	-86,18	29	-106,53	30	-86,20
[gi]	13	-88,10	26	-125,91	30	-112,03

APÊNDICE I

Oclusivas surdas e a posição ântero-posterior da vogal						
	N		P		A	
	Tokens	Média	Tokens	Média	Tokens	Média
[pɔ]	30	49,90	30	36,73	30	18,13
[pi]	30	50,70	30	21,00	30	11,60
[tɔ]	30	54,37	30	41,73	30	22,87
[ti]	30	53,90	30	40,77	30	44,60
[kɔ]	30	53,33	30	48,50	30	34,33
[ki]	30	54,27	30	56,93	30	42,47

APÊNDICE J

Oclusivas sonoras com pré-vozeamento e a posição ântero-posterior da vogal						
	N		P		A	
	Tokens	Média	Tokens	Média	Tokens	Média
[bɔ]	15	-106,10	26	-127,57	30	-130,90
[bi]	13	-102,55	29	-128,38	30	-126,93
[dɔ]	13	-99,96	23	-122,28	30	-107,87
[di]	12	-84,00	22	-123,18	30	-106,87
[gɔ]	13	-68,81	24	-148,13	30	-104,00
[gi]	14	-104,05	26	-128,83	29	-112,56

APÊNDICE K

Oclusivas sonoras com VOT de intervalo curto e a posição ântero-posterior da vogal						
	N		P		A	
	Tokens	Média	Tokens	Média	Tokens	Média
[bɔ]	15	13,93	4	8,00	0	---
[bi]	17	9,44	1	0,00	0	---
[dɔ]	17	14,80	7	15,46	0	---
[di]	18	13,82	8	26,08	0	---
[gɔ]	17	22,05	6	23,60	0	---
[gi]	16	20,50	4	22,38	1	15,00