

Mariane Antero Alves

**ESTUDO DOS PARÂMETROS ACÚSTICOS RELACIONADOS  
À PRODUÇÃO DAS PLOSIVAS DO PORTUGUÊS BRASILEIRO  
NA FALA ADULTA: ANÁLISE ACÚSTICO-QUANTITATIVA**

Tese submetida ao Programa de Pós-Graduação em Linguística da Universidade Federal de Santa Catarina, como requisito à obtenção do título de Doutor em Linguística.

Orientadora: Profa. Dra. Izabel Christine Seara.

Florianópolis

2015

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Alves, Mariane Antero

Estudo dos parâmetros acústicos relacionados à produção de plosivas do português brasileiro na fala adulta : análise acústico-quantitativa / Mariane Antero Alves ; orientadora, Izabel Christine Seara - Florianópolis, SC, 2015.

253 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão. Programa de Pós Graduação em Linguística.

Inclui referências

1. Linguística. 2. Plosivas. 3. Português brasileiro. 4. Produção de fala. 5. Parâmetros acústicos. I. Seara, Izabel Christine. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Linguística. III. Título.

Mariane Antero Alves

**ESTUDO DOS PARÂMETROS ACÚSTICOS RELACIONADOS  
À PRODUÇÃO DAS PLOSIVAS DO PORTUGUÊS BRASILEIRO  
NA FALA ADULTA: ANÁLISE ACÚSTICO-QUANTITATIVA**

Esta Tese foi julgada adequada para a obtenção do título de Doutora em Linguística, e aprovada em sua forma final pelo Programa de Pós-Graduação em Linguística.

Florianópolis, 14 de setembro de 2015.



Prof. Heronides Maurílio de Melo Moura, Dr.  
Coordenador do Curso

**Banca Examinadora:**



Prof.ª Izabel Christine Seara, Dr.ª  
Orientadora

Universidade Federal de Santa Catarina



Prof.ª Adelaide H. Pescatori Silva, Dr.ª  
Universidade Federal do Paraná



Prof. Ubiratã Kickhofel Alves, Dr.  
Universidade Federal do Rio Grande do Sul



Prof.<sup>a</sup> Denise Kluge, Dr.<sup>a</sup>.  
Universidade Federal do Paraná



Prof.<sup>a</sup> Lilian Elisa Minikel Brod, Dr.<sup>a</sup>.  
Universidade do Estado do Mato Grosso



Prof.<sup>a</sup> Carla Cristofolini, Dr.<sup>a</sup>.  
Universidade Federal de Santa Catarina

**A todos aqueles que me ensinaram a voar.**



## AGRADECIMENTOS

Tantos a agradecer e tão poucas linhas pra escrever...

Agradeço primeiramente a Deus, que mais uma vez me deu força e paciência para concluir esta etapa tão esperada da minha vida.

Aos meus pais e minha irmã que, com incentivo, apoio e sacrifícios, me conduziram ao caminho do estudo que me trouxe até aqui.

Ao meu esposo, que com paciência, amor e carinho me confortou com palavras de incentivo a cada vez em que eu pensava em desistir.

À minha orientadora Dra. Izabel Christine Seara, que não só me orientou no doutorado, mas por toda a minha jornada acadêmica. A ela devo todo o meu amor pela Fonética. Foi através dessa grande pessoa, mestre e pesquisadora que pude ser apresentada ao mundo dos sons da fala, que pude me apaixonar por ele e que pude dedicar oito anos da minha vida a ele. Foi vendo o seu amor pela pesquisa e pela docência que tive inspiração para dar cada passo que dei na minha acadêmica: desde a iniciação científica, passando pelo mestrado e chegando até o doutorado. Seu exemplo de comprometimento com seus alunos e orientandos e de amor pela profissão irão continuar comigo por toda a minha vida.

Aos queridos amigos e familiares com quem muitas vezes tive de faltar por estar me dedicando a este trabalho. Meu muito obrigada por me ouvirem nos momentos de frustração, de derrota ou de impaciência e por me darem palavras de incentivo e conforto.

Aos queridos amigos e colegas do Fonapli. Primeiramente obrigada às professoras Dra. Izabel e Dra. Cristiane pelos ensinamentos e também por lutarem sempre para que o laboratório pudesse ser cada dia mais um lugar acolhedor para todos os alunos que por lá passam. Obrigada a todos os membros discentes do laboratório pelo conhecimento compartilhado em manhãs e tardes, regados a cafés e risadas deliciosos!

À minha amiga, parceira de caminhada no mestrado e doutorado e futura doutora em Linguística e comadre Eva Dias. Só nós sabemos o que passamos para chegarmos até aqui. Com fé, perseverança e companheirismo, vencemos juntas diversas barreiras. O seu apoio incondicional (para assuntos profissionais e pessoais) foi essencial para que eu seguisse em frente nas tantas vezes que pensei em desistir.

À Dra. Carla Cristofolini, pela oportunidade de realizar um dos meus estágios de docência em uma das suas disciplinas e por toda ajuda que me deu naquele e em outros momentos.

Aos meus colegas de jornada: colegas de classe, mestres, doutores e professores, por todo o conhecimento compartilhado.

À UFSC e todos os membros docentes e técnicos do CCE, em especial do DLLV e do DLLE. Foi essa Universidade que me acolheu há mais de 10 anos, que me deu a oportunidade de ser a profissional que sou hoje. Foram anos que vou lembrar sempre com muito carinho.

A CAPES pelo apoio financeiro, sem o qual eu não teria concluído esta pesquisa.

Enfim, a todas as pessoas que, de alguma forma me ajudaram a concluir a mais importante etapa da minha vida acadêmica.

Hoje, ainda almejamos saber por que estamos aqui e de onde viemos. O desejo profundo da humanidade pelo conhecimento é justificativa suficiente para nossa busca contínua. E nosso objetivo é nada menos do que uma completa descrição do universo onde vivemos.

(Stephen Hawking)



## RESUMO

Esta pesquisa tem como objetivo investigar a produção de plosivas vozeadas e não-vozeadas de uma variedade do português brasileiro, com o intuito de verificar os principais parâmetros acústico-articulatórios envolvidos na produção desses segmentos, estejam eles relacionados ao próprio segmento ou a contextos adjacentes. Os dados foram coletados em seções de leitura de frases com cinco informantes do sexo feminino, com idades entre 21-29 anos, oriundas da região de Criciúma, no sul de estado de Santa Catarina. Os parâmetros analisados, a partir de experimentos de produção, são: (i) duração da oclusão; (ii) duração do segmento; (iii) VOT; (iv) duração das vogais adjacentes; (v) características espectrais de *burst*; (vi) transição formântica; (vii) qualidade vocálica; (viii) tonicidade e (ix) posição na palavra e (x) especificidades acústicas. A duração da oclusão e da consoante se mostraram maiores para as plosivas não-vozeadas. O VOT das plosivas não-vozeadas inseriu-se na categoria de retardo curto e a plosiva velar [k] apresentou valores acima de 35 ms, que apontam para uma leve aspiração. O VOT das plosivas vozeadas enquadrou-se na categoria de pré-sonorização. Os parâmetros relacionados à frequência, *burst* e transição vocálica, apesar de apresentarem características já apontadas por outros estudos, não apresentaram resultados sistemáticos entre si. A transição de F1 não foi estatisticamente distinta entre plosivas com diferente vozeamento, apesar de as curvas apresentarem direção ascendente em todos os casos. Para F2, os valores de frequência de *onset* foram significativamente menores para as bilabiais [p] e [b] quando seguidas pela vogal [ɐ]. As demais plosivas não-vozeadas não apresentaram diferença entre si e, dentre as vozeadas, a velar [g] foi a única que apresentou menor valor significativo quando seguida por [ɹ]. Todas as curvas de transição de F2 se mostraram descendentes. A vogal [ɹ] atenuou significativamente o valor de frequência de *onset* de F2 para [k] e [g]. Encontramos resultados que apontam para o *locus* consonantal de F2 para as bilabiais em 1400 Hz e de 1800 Hz para as alveolares. Quanto à influência das plosivas na duração das vogais adjacentes, observamos que as vogais sempre apresentaram valores de duração maiores quando eram seguidas ou antecidas por plosivas vozeadas. A duração da consoante e do VOT se mostraram maiores quando a consoante era seguida pela vogal alta [ɹ] em comparação à vogal baixa [ɐ], especialmente para [p], [k], [d] e [g]. O contexto tônico propiciou o

aumento da duração da consoante de todas as plosivas analisadas. A posição inicial favoreceu o aumento da duração da consoante e do VOT das plosivas vozeadas. Encontramos especificidades acústicas em uma parte considerável dos dados (24%), fato que denota a gradiência da fala e que esta apresenta nuances que nem sempre estão de acordo com o esperado pela teoria. Os resultados aqui obtidos mostram que alguns parâmetros, especialmente aqueles relacionados a aspectos duracionais, dão conta de caracterizar as plosivas do português brasileiro no que tange a sua produção e podem ser indícios de que serão também relevantes para a percepção desses segmentos.

**Palavras-chave:** Plosivas. Português brasileiro. Produção de fala. Parâmetros acústicos.

## ABSTRACT

The present research is aimed at investigating the production of voiced and voiceless stops in a variety of Brazilian Portuguese, in order to verify the main articulatory and acoustic parameters involved in the production of such segments, being related to the segment itself or to adjacent contexts. Data was collected in sentence-reading sections with five female participants, with ages varying from 21-29 years of age, all of them from Criciúma, a city in the south region of Santa Catarina state in Brazil. The parameters that were analyzed in this investigation were: (i) duration of the segment; (ii) duration of closure interval; (iii) VOT; (iv) duration of adjacent vowels; (v) spectral characteristics of burst; (vi) formant transition; (vii) vowel quality; (viii) lexical tone; (ix) position within the word and (x) acoustic specificity. Both consonant and closure duration were longer for the voiceless stops. VOT of all voiceless stops were considered short lag but [k] presented values higher than 35 ms, which signalizes slight aspiration. VOT of voiced stops were considered as voicing lead. The parameters related to frequency, *i.e.*, burst and transition, were non-systematic to indicate differences among stops, although they have shown some characteristics already stated by other studies. F1 transition was not statistically different between stops with different kinds of voicing, although the transition curves were all in an ascending direction. In relation to F2, the *onset* values were significantly lower for [p] and [b] when followed by [ɐ]. The other voiceless stops did not show significant differences between them, and, among the voiced ones, [gʊ] was the only one that showed lower significant values. All the F2 transition curves were descending. The vowel [ʊ] significantly lowered the *onset* F2 values for [k] and [g]. The results point to the F2 consonant locus of bilabial stops at 1400 Hz and of alveolar stops at 1800 Hz. Regarding the influence of stops on vowel duration, we observed longer values when vowels were preceding or following voiced stops. The consonant duration and the VOT were longer when followed by the high vowel [ʊ] rather than the low vowel [ɐ], especially in the case of [p], [k], [d], [g]. Also, the consonant duration were enhanced when the stop was inserted in a stressed position. We found acoustic specificities among a considerable part of our data (24%), which shows that speech is a continuum and presents nuances that not always comply with what is stated by the literature. The results obtained by this research point out that some parameters are able

to characterize the stops in Brazilian Portuguese, especially those related to durational aspects. This fact can raise evidence that these parameters might be relevant to the perception of these stops.

**Keywords:** Stops. Brazilian Portuguese. Speech Production. Acoustic Parameters.

## LISTA DE FIGURAS

Figura 1 - Espectrograma da palavra “pega”, mostrando a plosiva não-vozeada [p] e a plosiva vozeada [g] .....	36
Figura 2 - Configuração de glote para (a) respiração/segmentos não-vozeados; (b) respiração profunda; (c) fonação/segmentos vozeados; (d) sussurro .....	37
Figura 3 - Pontos de articulação das plosivas do PB: [p] – bilabial; [t] – alveolar; e [k] – velar. ....	41
Figura 4 - Momentos de uma plosiva não-vozeada não-aspirada, descritas por Fant (1973), mostrando os momentos: transiente, fricção, aspiração e vozeamento .....	42
Figura 5 - Oscilograma e espectrograma de uma plosiva [k] aspirada. Em destaque, os momentos: 1-oclusão; 2-soltura/explosão; 3-aspiração; e 4-transição com a vogal seguinte .....	43
Figura 6 - Forma de onda e espectrograma da palavra "data". Em destaque: duração total das plosivas [d] e [t] (primeira linha) e duração da oclusão e VOT (segunda linha) .....	45
Figura 7 - Forma de onda e espectrograma da palavra “bigamo”. Em destaque: a interrupção no vozeamento na parte medial de [g] .....	50
Figura 8 – Forma de onda e espectrograma da palavra “polícia”. Em destaque: a presença de <i>edge vibrations</i> no momento de oclusão de [p] .....	51
Figura 9 – Forma de onda e espectrograma da palavra “cunha”. Em destaque: a presença de dupla explosão em [k] .....	53
Figura 10 – Forma de onda e espectrograma da palavra “caqui”. Em destaque: a presença de múltiplas explosões em [k] .....	54
Figura 11 - Forma de onda e espectrograma da palavra “aborta”. Em destaque: a ausência de vozeamento na parte final de [b] .....	55
Figura 12 – Forma de onda e espectrograma da palavra de língua inglesa <i>can</i> (verbo “poder”). Em destaque: a aspiração de [k] .....	61
Figura 13 – Configuração dos tubos ressoadores no trato oral durante a fase de fechamento e soltura da oclusão de uma plosiva .....	69
Figura 14 – Configuração de uma sílaba CV, com as transições. ....	70
Figura 15 – Zona de transição da consoante do PB [p] seguida pela vogal [e], com destaque pra transição formântica de F1 e F2, para a palavra “pegou” .....	71
Figura 16 – Transições de F1 e F2 das plosivas do inglês seguidas da vogal [a], resultado da investigação de Cooper <i>et al.</i> (1952) .....	75
Figura 17 – Transições formânticas entre plosivas vozeadas e vogais no estudo de Delattre <i>et al.</i> (1955) .....	76
Figura 18 – Duração da vogal [æ] do inglês no par mínimo [mæt] e [mæd] .....	79
Figura 19 – Tela do <i>software Praat</i> com a etiquetagem da palavra “gata” .....	102
Figura 20 - Forma de onda e espectrograma da palavra "data". Em destaque: duração total das plosivas [d] e [t] .....	104
Figura 21 - Forma de onda e espectrograma da palavra "data". Em destaque: duração da oclusão das plosivas [d] e [t] .....	105

Figura 22 – Gráfico de distribuição de frequências mostrando assimetria negativa (a) e positiva (b).....	107
Figura 23 – Comparativo de distribuição com curtose positiva e negativa.....	108
Figura 24 – Inspeção visual do pico de energia do <i>burst</i> para extração da FFT na produção de [k].....	109
Figura 25 – FFT do segmento [k] com a indicação do pico spectral de 3079 Hz (parte superior).....	109
Figura 26 – Comandos no <i>software Praat</i> para obtenção dos momentos espectrais.....	110
Figura 27 – Exemplo de etiquetagem de VOT em plosiva não-vozeada [k] no <i>software Praat</i> .....	111
Figura 28 – Exemplo de etiquetagem de plosiva vozeada no <i>software Praat</i> ..	112
Figura 29 - Abertura do trato oral em uma plosiva velar, focalizando o distanciamento entre os articuladores.....	129
Figura 30 - Forma de onda e espectrograma da palavra "quéchua". Em destaque: o VOT da plosiva [k] não-aspirada.....	152
Figura 31 - Forma de onda e espectrograma da palavra "quero". Em destaque: o VOT da plosiva [k] levemente aspirada.....	152
Figura 32 – Histograma mostrando os valores da variável no eixo x (abscissas) e a distribuição de frequências no eixo y (coordenadas).....	230
Figura 33 – Distribuição dos quartis.....	232
Figura 34 – Histograma da duração da consoante para as plosivas não-vozeadas (esquerda) e vozeadas (direita).....	242

## LISTA DE QUADROS

Quadro 1 – Exemplos de vocábulos utilizados no teste de produção, de acordo com os contextos em análise.....	96
Quadro 2 – Lista de palavras do <i>corpus</i> para a análise da duração da vogal que antecede ([a]) e que sucede as plosivas ([ɐ] e [u]) .....	98
Quadro 3 – Códigos utilizados na etiquetagem dos dados.....	102
Quadro 4 - Valores de $p$ utilizados para se reportar os resultados .....	114
Quadro 5 - Resultados para cada uma das hipóteses levantadas pela presente pesquisa .....	192
Quadro 6 – Resumo dos testes de correlação e suas respectivas variáveis .....	239
Quadro 7 – Resumo dos testes de diferenças e suas respectivas variáveis.....	239
Quadro 8 – Informações estatísticas referentes à P1 para a avaliação entre plosivas vozeadas e não-vozeadas .....	241
Quadro 9 – Informações estatísticas referentes à P2.1, para a duração absoluta das não-vozeadas .....	243
Quadro 10 - Informações estatísticas referentes à P2.2, para a duração absoluta das vozeadas .....	243
Quadro 11 – Informações estatísticas referentes à P3.....	244
Quadro 12 – Informações estatísticas referentes à P4.1, para as plosivas não-vozeadas .....	244
Quadro 13 – Informações estatísticas referentes à P4.2, para as plosivas vozeadas .....	244
Quadro 14 – Informações estatísticas referentes à P6 para a avaliação de pico espectral.....	245
Quadro 15 – Informações estatísticas referentes à P7, para a comparação de pico espectral das alveolares entre os contextos vocálicos adjacentes [ɐ] e [u].....	246
Quadro 16 – Informações estatísticas referentes à P8.....	247
Quadro 17 – Informações estatísticas referentes à P9.1 para a análise do VOT entre as plosivas não-vozeadas .....	247
Quadro 18 – Informações estatísticas referentes à P10, para a plosiva bilabial .....	248
Quadro 19 – Informações estatísticas referentes à P11, para as plosivas bilabiais .....	249
Quadro 20 – Informações estatísticas referentes à P12.1, para não-vozeadas.....	249
Quadro 21 – Informações estatísticas referentes à P13, para a duração da vogal [a] em frente a não-vozeadas x vozeadas.....	249
Quadro 22 – Informações estatísticas referentes a P15, para a duração da consoante [p] em frente as vogais [ɐ] e [u] .....	251
Quadro 23 – Informações estatísticas referentes a P16.1, para a duração da consoante [p] no contexto início e meio de palavra.....	251
Quadro 24 – Informações estatísticas referentes a P17.1, para a duração da consoante [p] no contexto tônico e átono.....	252



## LISTA DE TABELAS

Tabela 1 - Valores médios (ms) de duração absoluta das plosivas do PB.....	46
Tabela 2 - Resultados de Cristofolini (2013) para as especificidades acústicas, conforme a faixa etária dos participantes.....	52
Tabela 3 - Valores médios de VOT (ms) para as plosivas que apresentaram maior aspiração no inglês e no coreano, no estudo de Lisker e Abramson (1964).....	62
Tabela 4 – Categorias de plosivas não-vozeadas, de acordo com o VOT, propostas por Cho e Ladefoged (1999), com valores aproximados.....	65
Tabela 5 – Valores médios de VOT (ms) para as plosivas do PB, a partir dos resultados encontrados em diversos estudos.....	67
Tabela 6 – Resumo do banco de dados.....	100
Tabela 7 – Valores médios, desvio padrão e faixa de distribuição da taxa de elocução por participante (sílabas/segundo).....	116
Tabela 8 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com a comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas vozeadas e não-vozeadas.....	117
Tabela 9 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) entre os pares homorgânicos.....	118
Tabela 10 – Valores médios de duração absoluta (ms) das plosivas do PB nos estudos de Barbosa (1996) e de Ficker (2003).....	120
Tabela 11 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas não-vozeadas.....	121
Tabela 12 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas vozeadas.....	122
Tabela 13 – Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre plosivas vozeadas e não-vozeadas.....	123
Tabela 14 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre os pares homorgânicos.....	124

Tabela 15 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre as plosivas não-vozeadas .....	125
Tabela 16 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da consoante entre as plosivas vozeadas .....	126
Tabela 17 - Total de dados e de especificidades por plosivas.....	127
Tabela 18 – Total de dados com especificidade por plosiva, de acordo com o tipo apresentado.....	128
Tabela 19 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para o pico espectral, juntamente com os resultados dos testes obtidos na comparação de pico espectral entre as plosivas não-vozeadas .....	132
Tabela 20 – Valores de média de pico espectral (Hz), de desvio padrão e de faixa de distribuição de pico espectral (Hz), juntamente com os resultados obtidos na comparação de pico espectral entre as plosivas vozeadas .....	133
Tabela 21 – Valores de média do centroide (Hz), de desvio padrão e de faixa de distribuição do centroide (Hz), juntamente com os resultados dos testes obtidos na comparação do centroide entre as plosivas não-vozeadas.....	134
Tabela 22 – Valores de média do centroide (Hz), de desvio padrão e de faixa de distribuição do centroide (Hz), juntamente com os resultados dos testes obtidos na comparação do centroide entre as plosivas vozeadas.....	135
Tabela 23 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para a variância, juntamente com os resultados dos testes obtidos na comparação da variância entre as plosivas não-vozeadas.....	136
Tabela 24 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para a variância, juntamente com os resultados dos testes obtidos na comparação da variância entre as plosivas vozeadas.....	136
Tabela 25 – Valores de média, de desvio padrão e de faixa de distribuição para a assimetria, juntamente com os resultados dos testes obtidos na comparação da assimetria entre as plosivas não-vozeadas .....	137
Tabela 26 – Valores de média, de desvio padrão e de faixa de distribuição para a assimetria, juntamente com os resultados dos testes obtidos na comparação da assimetria entre as plosivas vozeadas .....	137
Tabela 27 – Valores de média de curtose, de desvio padrão e de faixa de distribuição de curtose, juntamente com os resultados dos testes obtidos na comparação da curtose entre as plosivas não-vozeadas.....	138
Tabela 28 – Valores de média de curtose, de desvio padrão e de faixa de distribuição de curtose, juntamente com os resultados dos testes obtidos na comparação da curtose entre as plosivas vozeadas.....	139
Tabela 29 – Valores médios e de desvio padrão para as características espectrais de <i>burst</i> , de acordo com a vogal seguinte a cada tipo de plosiva, juntamente com	

os resultados dos testes obtidos na comparação entre o contexto vocálico, para cada tipo de plosiva .....	141
Tabela 30 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas vozeadas e não-vozeadas .....	143
Tabela 31 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa de VOT entre os pares homorgânicos.....	144
Tabela 32 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas não-vozeadas .....	145
Tabela 33 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas vozeadas .....	148
Tabela 34 – Percentual de dados com aspiração dentre as plosivas não-vozeadas .....	150
Tabela 35 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, de acordo com a presença ou não de aspiração, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas sem e com aspiração.....	151
Tabela 36 – Resultados para a transição de F1 em Hz no primeiro e último momento de coleta, juntamente com o resultado dos testes obtidos na comparação de F1_1 entre os pares de plosivas homorgânicas .....	155
Tabela 37 – Resultados para a transição de F2 (Hz) no primeiro e último momento de coleta, juntamente com os resultados dos testes obtidos na comparação de F2_1 ( <i>onset</i> ) entre os pares de plosivas não-vozeadas.....	158
Tabela 38 – Resultados para a transição formântica de F2 (Hz) no primeiro e último momento de coleta, juntamente com os resultados dos testes obtidos na comparação de F2_1 entre os pares de plosivas vozeadas .....	162
Tabela 39 - Valores médios de duração absoluta (ms) e relativa da vogal (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração da vogal [a] no contexto VC, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa da vogal [a] quando seguida por plosivas vozeadas ou não-vozeadas.....	168
Tabela 40 – Valores médios de duração absoluta (ms) e relativa da vogal (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) da vogal [a] no contexto VC, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa da vogal entre os pares de plosivas homorgânicas.....	169

Tabela 41 - Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração das vogais [ɐ] e [ʊ] no contexto CV, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa dessas vogais quando seguidas por plosivas vozeadas ou não-vozeadas .....	171
Tabela 42 - Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração das vogais [ɐ] e [ʊ] no contexto CV, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa das vogais [ɐ] e [ʊ] entre os pares homorgânicos.....	172
Tabela 43 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração absoluta da consoante e do VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com o contexto vocálico seguinte.	173
Tabela 44 - Valores médios de duração da consoante e VOT (ms), de desvio padrão, de faixa de distribuição da duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e VOT para cada plosiva vozeada, de acordo com o contexto vocálico seguinte.....	174
Tabela 45 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com a posição na palavra.....	176
Tabela 46 – Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva vozeada, de acordo com a posição na palavra .....	177
Tabela 47 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com a tonicidade da sílaba.....	179
Tabela 48 - Valores médios de duração da consoante e VOT (ms), de desvio padrão, de faixa de distribuição da duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva vozeada, de acordo com a tonicidade da sílaba.....	180
Tabela 49 - Resumo dos resultados para cada plosiva por parâmetro avaliado .....	182
Tabela 50 - Resumo dos resultados para cada plosiva por ocorrência de especificidades acústicas .....	184

## LISTA DE ABREVIATURAS E SIGLAS

IA – inglês americano  
PB – português brasileiro  
LM – língua materna  
LE – língua estrangeira  
VOT – *voice onset time*  
PE – português europeu  
F1 – primeiro formante  
F2 – segundo formante  
F3 – terceiro formante  
F0 – frequência fundamental  
FFT – *Fast Fourier Transform*  
LPC – *Linear predictive coding*  
Hz – Hertz  
dB – decibéis  
ms – milissegundos  
ILE – inglês como língua estrangeira  
KS – teste de normalidade Kolmogorov-Smirnov  
SW – teste de normalidade Shapiro-Wilk  
VD – variável dependente  
VI – variável independente  
DP – desvio padrão  
 $\bar{X}$  – média  
≠ – diferença



## SUMÁRIO

SUMÁRIO .....	49
<b>1 INTRODUÇÃO .....</b>	<b>27</b>
1.1 OBJETIVOS.....	29
<b>1.1.1 Objetivo Geral.....</b>	<b>29</b>
<b>1.1.2 Objetivos Específicos.....</b>	<b>30</b>
1.2 JUSTIFICATIVA.....	31
1.3 ORGANIZAÇÃO DA TESE.....	31
<b>2 PRODUÇÃO DE SEGMENTOS PLOSIVOS.....</b>	<b>33</b>
2.1 TEORIA ACÚSTICA DA PRODUÇÃO DOS SONS PLOSIVOS.....	33
<b>2.1.1 Modo de articulação.....</b>	<b>34</b>
<b>2.1.2 Vozeamento.....</b>	<b>35</b>
<b>2.1.3 Ponto ou região de articulação.....</b>	<b>39</b>
2.2 PARÂMETROS ACÚSTICOS ENVOLVIDOS NA PRODUÇÃO DOS SONS PLOSIVOS.....	44
<b>2.2.1 Duração da consoante.....</b>	<b>45</b>
<b>2.2.2 Duração da oclusão.....</b>	<b>47</b>
<b>2.2.3 Especificidades acústicas.....</b>	<b>48</b>
<b>2.2.4 <i>Burst</i>.....</b>	<b>55</b>
<b>2.2.5 VOT e aspiração.....</b>	<b>58</b>
<b>2.2.6 Transição formântica.....</b>	<b>67</b>
2.2.6.1 Transição de F1.....	71
2.2.6.2 Transição de F2.....	75
<b>2.2.7 Duração e tipologia de vogais adjacentes.....</b>	<b>78</b>
<b>2.2.8 Posição na palavra.....</b>	<b>82</b>
<b>2.2.9 Tonicidade.....</b>	<b>83</b>
<b>2.2.10 Taxa de elocução.....</b>	<b>85</b>
2.3 RESUMO DOS ESTUDOS SOBRE A PRODUÇÃO DE PLOSIVAS.....	87
<b>2.3.1 As plosivas em geral.....</b>	<b>87</b>
<b>2.3.2 As plosivas no PB.....</b>	<b>89</b>
2.3 PERGUNTAS E HIPÓTESES DE PESQUISA.....	90

<b>3 METODOLOGIA.....</b>	<b>94</b>
3.1 PARTICIPANTES.....	94
3.2 <i>CORPUS</i> .....	95
3.3 PROCEDIMENTOS DE COLETA.....	100
3.4 ANÁLISE DE DADOS.....	101
<b>3.4.1 Etiquetagem.....</b>	<b>101</b>
<b>3.4.2 Análise Acústica.....</b>	<b>103</b>
3.4.2.1 Duração absoluta e relativa da consoante.....	103
3.4.2.2 Duração da oclusão da plosiva.....	104
3.4.2.3 Especificidades acústicas.....	105
3.4.2.4 Picos e momentos espectrais de <i>burst</i> .....	106
3.4.2.5 Duração total e relativa de VOT.....	110
3.4.2.6 Transição formântica.....	112
3.4.2.7 Duração da vogal precedente/seguite e qualidade vocálica.....	113
3.4.2.8 Taxa de elocução.....	113
<b>3.5 Estatística utilizada na tese.....</b>	<b>113</b>
<b>4 ANÁLISE DE DADOS E RESULTADOS.....</b>	<b>116</b>
4.1 TAXA DE ELOCUÇÃO.....	116
4.2 DURAÇÃO DA CONSOANTE.....	117
<b>4.2.1 Duração da consoante e vozeamento.....</b>	<b>117</b>
<b>4.2.2 Duração da consoante e ponto de articulação.....</b>	<b>119</b>
4.3 DURAÇÃO DA OCLUSÃO.....	123
<b>4.3.1 Duração da oclusão e vozeamento.....</b>	<b>123</b>
<b>4.3.2 Duração da oclusão e ponto de articulação.....</b>	<b>125</b>
4.4 ESPECIFICIDADES ACÚSTICAS.....	127
4.5 CARACTERÍSTICAS ESPECTRAIS DO <i>BURST</i> .....	132
<b>4.5.1 Pico espectral.....</b>	<b>132</b>
<b>4.5.2 Centroide.....</b>	<b>134</b>
<b>4.5.3 Variância.....</b>	<b>135</b>
<b>4.5.4 Assimetria.....</b>	<b>136</b>
<b>4.5.5 Curtose.....</b>	<b>138</b>
<b>4.5.6 Relação entre <i>burst</i> e contexto vocálico adjacente.....</b>	<b>140</b>
4.6 VOT E ASPIRAÇÃO.....	143

4.6.1 VOT e vozeamento.....	143
4.6.2 VOT e ponto de articulação.....	145
4.6.3 Aspiração.....	150
4.7 TRANSIÇÃO FORMÂNTICA.....	154
4.7.1 Transição de F1.....	154
4.7.2 Transição de F2.....	158
4.8 CONTEXTO VOCÁLICO ADJACENTE.....	167
4.8.1 Duração da vogal no contexto VC.....	168
4.8.2 Duração da vogal no contexto CV.....	170
4.8.3 Influência da vogal na duração da consoante e do VOT.....	173
4.9 POSIÇÃO NA PALAVRA.....	175
4.10 TONICIDADE.....	178
4.11 RESUMO DOS RESULTADOS.....	181
<b>5 CONCLUSÃO.....</b>	<b>186</b>
5.1 IMPLICAÇÕES PARA A PRODUÇÃO DE PLOSIVAS.....	187
5.2 LIMITAÇÕES.....	195
5.3 INDICAÇÕES PARA PESQUISAS FUTURAS.....	196
5.4 CONSIDERAÇÕES FINAIS.....	197
<b>APÊNDICE A – Questionário biográfico e Termo de Consentimento.....</b>	<b>208</b>
<b>APÊNDICE B – <i>Corpus</i> com as palavras contendo os segmentos-alvo para o teste de produção.....</b>	<b>210</b>
<b>APÊNDICE C – Script para a extração de dados relativos à transição vocálica.....</b>	<b>216</b>
<b>APÊNDICE D – A Estatística aplicada na validação de parâmetros acústicos em análise de fala.....</b>	<b>221</b>
A) “SENTA QUE LÁ VEM HISTÓRIA”.....	222
B) POR QUE ESTUDAR ESTATÍSTICA QUANDO ESCOLHI ESTUDAR LETRAS?.....	223
D) ESTATÍSTICA: pressupostos.....	225
<b>Estatística Descritiva: pressupostos teóricos.....</b>	<b>226</b>
<b>Estatística Inferencial: Testes Estatísticos.....</b>	<b>235</b>

Testes de correlação.....	235
Testes de diferença.....	235
<i>Testes Paramétricos</i> .....	235
<i>Testes não-paramétricos</i> .....	237
E) ESTATÍSTICA UTILIZADA NA TESE.....	240
<b>ANEXO I – Certificado de aprovação no Comitê de Ética .....</b>	<b>253</b>

## 1 INTRODUÇÃO

Desde meados do século passado, estudos acústicos vêm investigando os parâmetros acústicos que são responsáveis por definir as categorias de sons que compõem os inventários das mais diversas línguas (ABRAMSON; LISKER, 1965, 1973; COOPER; DELATTRE; LIBERMAN; BORST; GERSTMAN, 1952; LIBERMAN; DELATTRE; COOPER, 1958; LIBERMAN; HARRIS; HOFFMAN; GRIFFITH, 1957; LISKER; ABRAMSON, 1964a; STEVENS; KLATT, 1974, dentre outros). O que se tem observado, a partir desses estudos, é que cada parâmetro acústico está associado a uma determinada faixa de valores que irão defini-los e cada língua seleciona as suas respectivas faixas (de duração, por exemplo) que caracterizam seus segmentos de fala. É por essa razão que podem existir línguas que apresentem um mesmo segmento, mas a maneira como eles são foneticamente implementados pode ser distinta. Um exemplo dessas diferentes implementações ocorre com as plosivas não-vozeadas que, no inglês americano (doravante IA), podem ser produzidas com uma longa aspiração e, no português brasileiro (doravante PB), com leve ou nenhuma aspiração (ALVES, 2011).

Os segmentos plosivos são os únicos tipos de segmentos consonantais presentes em todas as línguas, conforme apontam Ladefoged e Maddieson (1996). Por essa razão, ao pesquisar a literatura que concerne esse tipo de consoante, observamos que existe uma ampla bibliografia, especialmente para a língua inglesa e, a grande maioria, de cunho acústico-articulatório, que contribuíram para a compreensão do funcionamento desses segmentos na produção e percepção de fala.

Os primeiros estudos de que tomamos conhecimento que investigaram os segmentos plosivos, em língua inglesa, datam do início da década de 50, desenvolvidos, especialmente, por um grupo de pesquisadores filiados ao *Haskins Laboratories*, nos Estados Unidos. Nesse grupo, estavam nomes como o de Alvin Liberman e Pierre Delattre e Franklin Cooper, que lançaram mão dos primeiros sintetizadores de fala para produzir dados para testes perceptuais para a investigação, por exemplo, dos parâmetros que poderiam caracterizar a distinção do ponto de articulação das plosivas, como a transição de F1 e F2 no momento de transição entre a plosiva e a vogal. Os resultados encontrados por esses pesquisadores, apesar de datarem de mais de 60 anos, são utilizados como referência até hoje por estudos sobre a transição formântica.

Filiados a esse Laboratório, também estavam Leigh Lisker e Arthur Abramson, conhecidos por serem os primeiros a identificar e caracterizar o VOT, um dos parâmetros acústicos mais utilizados na medição/caracterização das plosivas. Esses pesquisadores realizaram uma quantidade expressiva de estudos observando o comportamento do VOT, na língua inglesa e em outras línguas, e mostraram que as diferenças entre os segmentos não se dá de maneira dicotômica, como por exemplo, pela presença ou não de aspiração, mas sim dentro de um *continuum* de fala.

Não obstante, estudos com plosivas de outras línguas, como do holandês com Fischer-Jørgensen (1954), do tai, do marathi, do espanhol, com Lisker e Abramson (ABRAMSON; LISKER, 1965, 1973; LISKER; ABRAMSON, 1964a), de línguas de povos nativos (CHO; LADEFOGED, 1999) e muitos outros, trouxeram contribuições bastante valiosas para a caracterização acústica dos segmentos plosivos de maneira geral e também para a compreensão das possíveis diferenças existentes entre esses segmentos para cada uma dessas línguas.

No Brasil, os trabalhos relacionados ao VOT e aos segmentos plosivos começaram a surgir com maior intensidade somente a partir da década de 80 e 90, tendo maior expressividade nas pesquisas de Behlau (1986) e Klein (1999).

No entanto, ainda hoje, dispomos de poucos trabalhos sobre o tema para o PB, se compararmos à vasta literatura existente para outras línguas, especialmente para a língua inglesa. Muitas das pesquisas que avaliaram a produção desses segmentos no PB tiveram como foco a comparação entre grupos tais como: língua materna (LM) e língua estrangeira (LE), como em Cohen (2004); Bettoni-Techio (2005); Alves (2011); dentre outros; fala típica e com desvio, como em Ficker (2003); Berti e Marino (2011); Brito (2000); Britto (2010); Cemin (2014); dentre outros; diferentes idades na fala infantil como em Bonatto (2007); Bonatto e Madureira (2009); Cristofolini, (2008; 2013); dentre outros; e ainda fala adulta e fala infantil, como Melo, Mota, Mezzomo, Brazil, Lovato, Arzeno (2012).

Poucos são os estudos, com fala típica, adulta e monolíngue, que se propuseram a investigar puramente a caracterização dos parâmetros acústicos e possíveis contextos linguísticos de influência nas plosivas vozeadas e não-vozeadas do PB. Dentre esses se destacam Behlau (1986 *apud* FICKER, 2003); Klein (1999); Alves *et al.* (2008); Alves e Dias (2010) que se debruçaram, principalmente, sobre a análise do *voice onset time* (VOT).

Outro ponto a ser investigado por esta pesquisa seria a natureza gradiente de algumas características acústicas inerentes às plosivas, (CHO; LADEFOGED, 1999; McMurrau *et. al*, 2002; dentre outros). Estudos vêm demonstrando que há gradiência em propriedades acústicas da fala, tais como VOT, aspiração, vozeamento e outros, pois mostram que os segmentos nem sempre são produzidos de acordo com o que tradicionalmente é considerado como “típico” pela literatura clássica, apresentando o que Cristofolini (2008; 2013) chama de especificidades acústicas. Segmentos vozeados com desvozeamento; não-vozeados com presença de vozeamento durante produção (CRISTOFOLINI, 2008; 2013; SNOEREN, 2005, dentre outros) redução e desvozeamento de vogais (DIAS; SEARA, 2013; MENESES, 2012; dentre outros), são alguns exemplos de especificidades que podem ocorrer e apontam para um caráter gradiente presente em aspectos acústicos da fala.

Considerando o acima exposto, este estudo pretende investigar o comportamento dos parâmetros acústicos associados às plosivas do PB e dos contextos linguísticos envolvidos na sua produção. Nosso objetivo maior é trazer para discussão os principais parâmetros que caracterizam acusticamente as plosivas no PB, através de dados de falantes nativos (falar criciumense). Para tal fim, partiremos dos pressupostos abarcados pela Fonética Acústica para a descrição e caracterização dos segmentos plosivos do PB.

Esperamos, assim, contribuir para a pesquisa brasileira em Fonética e Fonologia, no que diz respeito à descrição de características acústicas relacionadas aos segmentos plosivos no PB, pois acreditamos que através da análise dos parâmetros envolvidos na produção de um determinado som, pode-se partir para a investigação perceptual dos mesmos.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Esta tese tem como objetivo geral desenvolver um estudo congregando os principais parâmetros acústicos observados na produção de fala dos segmentos plosivos do PB, descritos na literatura, explorando-os através de uma análise quantitativa de dados de fala adulta, para que seu funcionamento possa ser investigado com maior detalhamento. Visamos, assim, elaborar um trabalho que descreva com maior robustez as propriedades acústicas e também os contextos

linguísticos de influência envolvidos na produção dos segmentos plosivos vozeados e não-vozeados para o PB.

### 1.1.2 Objetivos Específicos

Como objetivos específicos, pretendemos avaliar, na produção da fala de indivíduos adultos, as propriedades acústico-articulatórias que caracterizam os segmentos plosivos do PB, de acordo com o seu ponto de articulação e vozeamento. Focaremos, especialmente, nos correlatos acústicos das diferentes estratégias articulatórias da produção de segmentos plosivos vozeados e não-vozeados no PB. Buscaremos também avaliar a influência de diferentes contextos linguísticos na realização desses segmentos.

Dentre os parâmetros acústicos e contextos linguísticos relacionados aos segmentos plosivos, analisaremos:

- (i) da oclusão;
- (ii) da consoante;
- (iii) do VOT (aspiração);
- (iv) relativa (considerando a duração do segmento em relação à duração da palavra em que o segmento está inserido);
- (v) da vogal que antecede a plosiva;
- (vi) da vogal que segue a plosivas;
- (vii) características espectrais de *burst*
- (viii) transição formântica
- (ix) influência da qualidade vocálica que segue o segmento plosivo;
- (x) influência da tonicidade da sílaba na qual a plosiva está inserida;
- (xi) influência da posição na palavra na qual a plosiva está inserida.

Além das análises acima, conduziremos uma inspeção qualitativa de espectrogramas para avaliar a possível ocorrência do que nomeamos especificidades acústicas, isto é, características presentes na produção de plosivas que não estão de acordo com o apontado pela literatura como “típicas” para esses segmentos (DENES, 1955; LISKER; ABRAMSON, 1964; STEVENS, 2000; BONATTO, 2007, CRISTOFOLINI, 2013; dentre outros). Utilizaremos a classificação elaborada por Cristofolini (2013) para relatar quais os tipos de especificidades ocorrem em nossos dados e em que grau (percentual) elas ocorrem para as plosivas.

## 1.2 JUSTIFICATIVA

A Fonética e Fonologia é uma área dentro da Linguística que se debruça sobre a fala. Para entendermos como os seres humanos são capazes de dar significado a elementos sonoros produzidos pelo aparato vocal – o que cabe à Fonologia –, necessitamos saber primeiramente como esses elementos (que doravante chamaremos de segmentos, sons ou fones) são realizados e quais nuances nessa realização são capazes de diferenciá-los – o que cabe à Fonética.

Assim, esta pesquisa pretende, de maneira detalhada, congregando em um único trabalho, a avaliação dos principais parâmetros acústicos utilizados na produção da fala de plosivas vozeadas e não-vozeadas no PB, tendo como população-alvo sujeitos adultos, com aquisição de fala típica. Ao conduzir um trabalho anterior comparativo sobre as plosivas não-vozeadas produzidas por falantes brasileiros de inglês americano (IA) como língua estrangeira (ALVES, 2011), notamos que poucas pesquisas no PB avaliaram a fala adulta típica (em falantes monolíngues) como principal grupo de análise, como foi o caso de Behlau (*apud* Ficker, 2003) e Klein (1999) para a avaliação do VOT. Esse tema, dentro dos estudos acústicos dos segmentos do PB, necessita de aprofundamento, uma vez que, até o presente momento, poucos são os estudos que investigaram todos esses parâmetros associados às plosivas, se comparados à grande quantidade de trabalhos já publicados para outras línguas.

Acreditamos que nossa pesquisa não tecerá contribuições somente para a Fonética e Fonologia, mas para outras áreas que também se utilizam da pesquisa com segmentos de fala, como a Linguística Forense, Linguística Aplicada à LE, Sociolinguística, Dialectologia, Fonoaudiologia, Processamento e Síntese de Fala, dentre outras, pois esperamos fornecer subsídios para a compreensão do funcionamento da produção de segmentos plosivos para o PB, que poderão servir de base para pesquisas futuras em quaisquer dessas áreas.

## 1.3 ORGANIZAÇÃO DA TESE

Esta tese está organizada em cinco capítulos. No Capítulo 2, fazemos uma descrição sobre a produção dos segmentos plosivos vozeados e não-vozeados, com ênfase na sua descrição acústico-articulatória, bem como realizamos um levantamento de pesquisas relacionadas aos parâmetros acústicos e contextos linguísticos que influenciam na produção desses segmentos. Apresentamos, ao final

deste capítulo, as questões norteadoras da pesquisa e suas respectivas hipóteses.

No Capítulo 3, discorremos sobre a metodologia, descrevendo os procedimentos utilizados na confecção do *corpus*, na coleta, etiquetagem e análise de dados.

Na sequência, no Capítulo 4, apresentamos a análise de dados e discussão dos resultados. Neste capítulo, responderemos as perguntas de pesquisa e verificamos a comprovação ou não das hipóteses levantadas.

Por fim, o Capítulo 5 traz as principais conclusões do estudo baseadas nos resultados, aponta os percalços encontrados ao longo desta pesquisa e também aponta questões para trabalhos futuros.

## 2 PRODUÇÃO DE SEGMENTOS PLOSIVOS

Este capítulo apresenta uma revisão acerca das principais teorias sobre a produção de fala dos segmentos plosivos.

Aborda, também, uma descrição sobre as teorias relacionadas aos parâmetros acústicos envolvidos na realização desses segmentos através do suporte de estudos anteriores que se debruçaram sobre cada um desses parâmetros.

### 2.1 TEORIA ACÚSTICA DA PRODUÇÃO DOS SONS PLOSIVOS

Para a produção dos sons da fala, é necessário o uso de determinados mecanismos articulatórios. São esses mecanismos que vão propiciar, por exemplo, a pressão necessária para que o fluxo de ar possa sair dos pulmões, passar pela glote e se encaminhar para a cavidade oral, ou ainda fazer os ajustes de glote necessários para que a fonação seja possível, no caso dos sons vozeados.

Um dos principais mecanismos para que a fala se inicie é a pressão causada pelo fluxo de ar que se forma na região conhecida como subglótica (aquela que se localiza abaixo da glote). É necessário que essa pressão seja maior do que a pressão do meio exterior para fazer com que haja um movimento do fluxo de ar dos pulmões em direção à cavidade supraglótica e, por fim, sua saída pelas vias orais e/ou nasais. Esse movimento de ar dos pulmões em direção ao meio exterior é conhecido como fluxo de ar “egressivo” e é o mais utilizado pelas línguas do mundo. A pressão exercida pelo fluxo de ar irá também determinar o quão intenso será um som, isto

é, se iremos ouvi-lo como um som mais forte e, portanto, mais intenso ou como mais fraco e menos intenso (CLARK; YALLOP, 1995).

A passagem do fluxo de ar pela cavidade supraglótica (acima da glote) vai ter papel fundamental no modo de produção ou articulação de um segmento, pois o trato vocal funcionará como um filtro, modulando o som produzido pela passagem do ar através das pregas vocais. O modo de articulação dependerá, fundamentalmente, de como os articuladores estão configurados dentro do trato, do grau de constrição entre eles e do caminho que o fluxo de ar fará para transpô-los até a saída da cavidade oral.

Podemos perceber, então, que os sons da fala são produzidos através de uma fonte que os gera (as pregas vocais) e um filtro que os molda (o trato vocal), atribuindo-lhes certas características que vão

diferenciá-los uns dos outros. Essa teoria é conhecida dentro dos estudos fonéticos como “Fonte-Filtro”, desenvolvida por Fant (1960), que declara que “(...) the speech wave is the response of the vocal tract filter system to one or more sound sources” (p.15)<sup>1</sup>. Essa teoria também é conhecida na literatura como Teoria Acústica da Fala e demonstra que uma determinada propriedade articulatória terá o seu correlato acústico. Segundo Berti (2006), essa teoria foi responsável por incontáveis avanços nas pesquisas de cunho acústico e conseguiu explicar, ao longo das décadas após a sua formulação, 90% dos casos de análises feitas pelos estudos que a utilizavam como arcabouço teórico.

Baseando-se nessa teoria, os sons consonantais são produzidos com obstrução total ou parcial à passagem do ar. As características que o trato oral (o filtro) e as pregas vocais (a fonte) irão atribuir a um segmento consonantal determinam o seu modo e ponto de articulação, bem como o seu vozeamento.

### 2.1.1 Modo de articulação

A maneira como o ar irá transpor o trato oral indicará o modo de articulação dos sons da fala. A primeira divisão entre os segmentos, de acordo com o seu modo de articulação, se dá entre vogais e consoantes. As primeiras, como apontado previamente, são produzidas sem qualquer impedimento ao fluxo de ar na cavidade oral. As consoantes, por sua vez, apresentam algum grau de constrição do trato durante a passagem do ar, podendo causar fricção ou até mesmo bloqueio completo ao fluxo. De acordo com o tipo de impedimento que causam, as consoantes podem ser divididas em plosivas, fricativas, laterais, nasais, vibrantes, africadas, aproximantes, tepe e plosivas. Nosso estudo se restringe à investigação do primeiro tipo de consoante.

As plosivas são, portanto, segmentos consonantais produzidos com uma oclusão total do trato, que impede, momentaneamente, a passagem do fluxo de ar na cavidade oral, seguido pela soltura súbita do ar que leva à geração de uma explosão (também conhecida como *burst*). Assim, é o completo fechamento ou oclusão do trato oral que define o modo de articulação plosivo ou oclusivo.

Portanto, essas consoantes podem ser chamadas de oclusivas ou plosivas. A denominação oclusiva se refere à oclusão do trato vocal quando ocorre o bloqueio à passagem da corrente de ar, sendo, portanto,

---

<sup>1</sup> “a onda sonora é a resposta do sistema de filtro do trato oral a uma ou mais fontes sonoras” (FANT, 1960, p. 15). Minha tradução.

relacionada a uma característica articulatória. Já a denominação plosiva se refere à geração da explosão (ou *burst*) quando ocorre a soltura do ar, sendo, portanto, relacionada a um correlato acústico (*burst*). Para a presente pesquisa, utilizaremos a denominação plosivas para diferenciá-las do termo oclusão, quando nos referirmos ao momento de silêncio ou bloqueio dessa consoante.

### 2.1.2 Vozeamento

O vozeamento é uma característica articulatória que está ligada à vibração das pregas vocais durante a produção dos segmentos. De modo geral, as vogais são sempre vozeadas, enquanto as consoantes podem ou não apresentar vozeamento.

É a vibração ou não das pregas vocais durante a passagem do ar pela glote que irá indicar o tipo de vozeamento, isto é, se a consoante será vozeada ou não-vozeada<sup>2</sup>. Segundo Veloso (1997), o correlato acústico da vibração das pregas é

“a presença, durante a articulação dos segmentos vozeados, de energia acústica de baixa intensidade e baixa frequência, visualizável, num espectrograma, sob a forma de pequenas estrias verticais junto à base do espectro (estas estrias recebem o nome de ‘barra de vozeamento’)” (p.60).

Lisker e Abramson (1964a) dividem os segmentos, primeiramente, entre aqueles chamados de vozeados, nos quais há a vibração de pregas vocais durante a sua produção e os não-vozeados, aqueles em que não há o movimento de vibração das pregas. Segundo esses pesquisadores, esses dois tipos de plosivas se distinguem acusticamente, na inspeção espectrográfica, pois

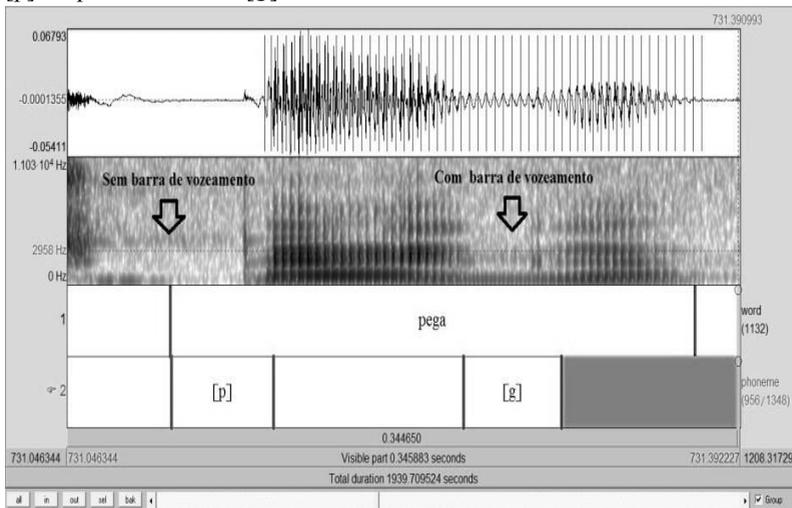
(...) for voiced stops the formantless segment corresponding to the closure interval is traversed by a small number of low-frequency harmonic components, while in the case of voiceless stops the closure interval is essentially blank (LISKER; ABRAMSON, 1964a, p.384).<sup>3</sup>

<sup>2</sup> Pode-se também encontrar na literatura o termo “surdo” para segmentos não-vozeados e “sonoro” para segmentos vozeados.

<sup>3</sup> (...) para plosivas vozeadas, o segmento sem formantes que corresponde ao intervalo de bloqueio é atravessado por um pequeno número de componentes harmônicos de baixa

Como vimos anteriormente, Veloso (1997) descreve essa energia de baixa frequência presente nas plosivas vozeadas como “barra de vozeamento”. Na Figura 1, podemos visualizar, a partir da forma de onda e do espectrograma, a diferença entre sons vozeados e não-vozeados, pela presença ou ausência da barra de vozeamento, respectivamente.

Figura 1 - Espectrograma da palavra “pega”, mostrando a plosiva não-vozeada [p] e a plosiva vozeada [g]



Fonte: Praat/Nossos dados

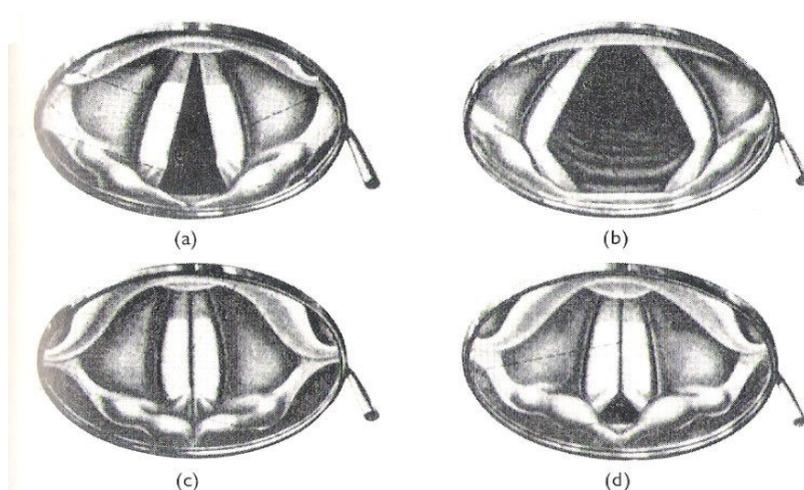
Pela Figura 1, podemos perceber algumas diferenças entre a primeira e a segunda região apontadas pelas setas. Primeiramente, visualizamos estriações verticais na forma de onda de [g] que representam a periodicidade relativa ao batimento de pregas presente em plosivas vozeadas (os pulsos glotais). A mesma periodicidade não pode ser vista na forma de onda de [p]. Com relação ao espectrograma, percebemos uma “faixa” horizontal escura na parte mais inferior da plosiva [g], a barra de vozeamento (VELOSO, 1997; LISKER; ABRAMSON, 1964a; JAKOBSON; FANT; HALLE, 1963; LIBERMAN, DELATTRE, GERSTMAN, COOPER 1956; LISKER; ABRAMSON, 1964b; LISKER, 1957).

frequência; enquanto, para as plosivas não-vozeadas, o intervalo de bloqueio é, essencialmente, vazio (LISKER; ABRAMSON, 1964, p. 384). Minha tradução.

O batimento de pregas vocais ocorre, segundo a teoria aerodinâmica mioelástica, desde que haja pressão de ar na região subglótica entre 4 e 8 cmH<sub>2</sub>O, e que ela seja diferente da pressão acima da glote (ABRAMSON, 1977). Essa pressão é causada pelo ar retido atrás da oclusão e aumenta até que seja suficiente para fazer com que as pregas se separem e, conseqüentemente, liberem a passagem desse ar pela glote. Assim que o ar é liberado com uma velocidade alta, a pressão abaixo das pregas começa a diminuir para cerca de 4 cmH<sub>2</sub>O, o que faz com que elas possam novamente se fechar e que o processo volte a acontecer (CLARK; YALLOP, 1995). Esse fenômeno é explicado pelo Princípio de Bernoulli, que afirma que quando a velocidade de um fluido aumenta enquanto ele passa por um tubo ou superfície a pressão diminui. Aplicando esse princípio à fonação, à medida que o ar retido atrás das pregas é liberado (em alta velocidade), a pressão na região da glote diminui, gerando uma força que empurra o fluxo de ar para fora da cavidade oral, fazendo com que as pregas vocais comecem a se movimentar.

No caso dos segmentos não-vozeados, não há batimento devido ao considerável afastamento entre as pregas durante a passagem de ar pela região da glote. A Figura 2 exemplifica a configuração das pregas para quatro momentos distintos, incluindo a fonação.

Figura 2 - Configuração de glote para (a) respiração/segmentos não-vozeados; (b) respiração profunda; (c) fonação/segmentos vozeados; (d) sussurro



A configuração de glote para o segmento não-vozeado pode ser visto na Figura 2 no exemplo (a), pois há considerável afastamento entre as pregas. O exemplo (c) mostra as pregas totalmente aproximadas, correspondendo ao período em que há maior pressão do ar retido nos pulmões e que pressionará as pregas a se afastarem e a vibrarem, representando a fonação.

O batimento de pregas também determina a frequência com que o ar sai da cavidade, conhecida como “frequência fundamental” (doravante F0) e que tem como correlato perceptual o *pitch*. O F0 é determinado pela própria anatomia das pregas, pelo comprimento e tensão exercidos pelos músculos laringeos sobre elas e também pela pressão causada pelo ar na região subglótica. Conforme a configuração desses músculos, o tempo do movimento de afastamento e aproximação das pregas vocais (o período fundamental - T) vai ser diferente e, conseqüentemente, a frequência fundamental, pois ela corresponde ao inverso desse período (1/T). Assim, cada indivíduo possui um F0 distinto. Os homens, como possuem pregas vocais mais espessas e mais longas do que as das mulheres<sup>4</sup>, levam, portanto, um período maior para afastarem e aproximarem novamente as pregas, resultando em um F0 menor. Em média, o F0 de homens é de 110 Hz. Já as mulheres possuem pregas menos espessas e mais curtas e fazem esse movimento de afastamento e aproximação mais rapidamente (em menos tempo), resultando em um valor de F0 mais elevado do que para os homens. Em média, o F0 de mulheres é de 220 Hz.

Stevens (2000) descreve, resumidamente, como ocorre o mecanismo de tensionamento e movimentação das pregas vocais.

The vocal and ventricular folds are supported at either end by a configuration of several cartilages. These supporting structures can be manipulated to adjust the position of the vocal folds so that they are approximated (or adducted) or are separated (or abducted), leaving a space between the vocal folds. The supporting structures can be displaced in a way that changes the vocal fold length, resulting in an increase or decrease in the tension of the folds, and a corresponding decrease or

---

<sup>4</sup> De acordo com Stevens (2000), as pregas vocais de homens apresentam, em média, 0,5mm de espessura e 1,5 cm de comprimento, enquanto a das mulheres apresenta, em média, 0,4mm de espessura e 1,0 cm de comprimento.

increase in the mass per unit length. (STEVENS, 2000, p.5)<sup>5</sup>

Segundo Johnson (2003), em uma plosiva vozeada típica, o ar que sai dos pulmões em direção ao trato oral deixando a cavidade subglótica é barrado no ponto da oclusão. Desse modo, não há mais ar suficiente na região subglótica para pressionar as pregas a baterem e, então, a vibração geralmente cessa ao final da oclusão. Para que o vozeamento seja mantido para as plosivas vozeadas, o movimento de batimento de pregas é auxiliado pelo mecanismo de distensão das paredes da faringe e abaixamento da laringe, fazendo com que a região entre a glote e o ponto de constrição seja aumentada e, assim, o tempo de vozeamento possa ser expandido. Essa manobra articulatória impede que a pressão causada pelo ar atrás da constrição aumente subitamente, o que pode prejudicar a manutenção do batimento de pregas. Além disso, a fase de oclusão é mais curta para as plosivas vozeadas do que para a sua contraparte não-vozeada, pois, nesse tempo mais curto, a vibração de pregas pode ser mantida.

### 2.1.3 Ponto ou região de articulação

A configuração do trato vocal durante a produção de um som pode lhe atribuir outras características além do modo de articulação e do vozeamento, que estão relacionadas ao local no qual os articuladores entram em contato para formar o bloqueio total ou parcial no trato. O ponto ou região de articulação divide as consoantes em:

- (i) bilabial: oclusão entre lábios inferiores e superiores.  
Exemplo: [p] em **p**ato e [b] **b**eijo no PB;
- (ii) interdental: (ponta da língua se posiciona entre os dentes).  
Ex.: [θ] e [ð] do inglês, como em *think* (pensar) e *these* (aqueles), respectivamente.

---

<sup>5</sup> “As pregas vocais e ventriculares são sustentadas em cada ponto por uma configuração de várias cartilagens. Essas estruturas de sustentação podem ser manipuladas para ajustar a posição das pregas para que elas fiquem próximas (aduzidas) ou afastadas (abduzidas), deixando um espaço entre elas. As estruturas de sustentação podem ser configuradas de forma que se possa mudar o comprimento das pregas, resultando em um aumento ou diminuição no seu tensionamento e em um aumento ou diminuição na massa [de ar] por unidade.” (STEVENS, 2000, p.5). Minha tradução.

- (iii) alveolar: ponta da língua toca ou se direciona aos alvéolos – parte rugosa atrás dos dentes superiores. Ex.: [t] em **todo** e [d] em **dedo**, no PB.
- (iv) palatal: corpo da língua toca o palato duro. Ex.: [ʎ] em **palha** e [ɲ] em **manhã**, no PB.
- (v) velar: dorso da língua toca o véu do palato ou palato mole. Ex.: [k] em **como** e [g] em **gato**, no PB.
- (vi) uvular: produzido pela vibração da úvula (conhecida popularmente como “campainha”). Ex.: [χ] do alemão, como em *dach* (telhado).
- (vii) glotal: som produzido na glote. Ex.: [h] do alemão, como em *hass* (ódio).

Como já vimos, o presente estudo investiga consoantes plosivas, vozeadas ou não-vozeadas, produzidas em três pontos de articulação distintos: bilabial, alveolar e velar, os três pontos de articulação usados na produção das plosivas no PB (SEARA, NUNES, LAZAROTTO-VOLCÃO, 2015).

A primeira delas é a região labial, na qual ocorre o contato entre os lábios inferiores (articulador ativo) e superiores (articulador passivo). Nessas plosivas, a língua não participa do movimento articulatório para a formação da oclusão e há um movimento de palato mole para impedir que o fluxo de ar penetre também a cavidade nasal (sem esse movimento, o som produzido seria a consoante nasal [m]). A vibração ou não das pregas vogais vai determinar se o segmento é não-vozeado, [p] ou vozeado [b].

A segunda se caracteriza por uma oclusão entre a ponta da língua e a região posterior dos dentes superiores ou dos alvéolos. Os lábios estão afastados e o palato mole também está obstruindo a passagem do ar para a cavidade nasal (do contrário, teríamos a produção da consoante [n]). A plosiva [t] não apresenta vozeamento, enquanto [d] é vozeado.

A última região é conhecida como velar por apresentar obstrução entre o dorso da língua e o palato mole (também conhecido como “véu do palato”), que se configura de tal forma a fechar a entrada da cavidade nasal, evitando a produção de um som nasal velar [ŋ]. Nessa região são produzidos os segmentos [k], não-vozeado e [g], vozeado. Através da Figura 3 se pode observar como o trato vocal se configura para a produção de cada um dos tipos de plosivas no PB.

Figura 3 - Pontos de articulação das plosivas do PB: [p] – bilabial; [t] – alveolar; e [k] – velar.



Fonte: Alves, 2011, p. 33 (adaptado de Moodle Ead/UFSC).

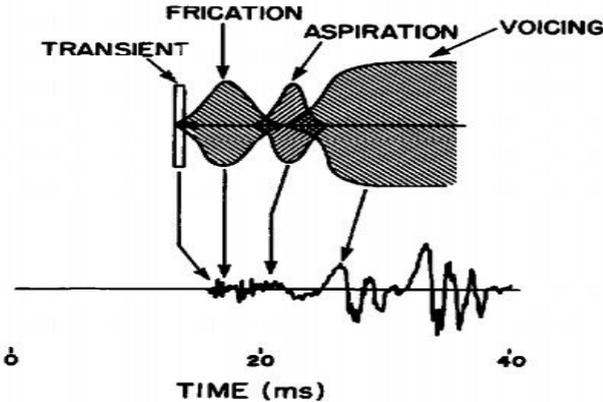
Cabe salientar que todas as configurações articulatórias apresentadas na Figura 3 correspondem a sons não-vozeados, por não apresentarem vibrações de pregas vocais. Assim, essa configuração de glote vai designar o terceiro aspecto que distingue os segmentos plosivos, o vozeamento.

Clark e Yallop (1995) identificam três momentos durante a produção de uma plosiva: (1) a oclusão, (2) a soltura e (3) a fase de transição para o segmento adjacente. A oclusão da plosiva ocorre devido ao contato entre os articuladores ativo e passivo, culminando em uma obstrução total dentro da cavidade oral que irá reter o fluxo de ar atrás dessa oclusão. Em seguida, há o afastamento dos articuladores, fazendo com que o ar saia com uma leve fricção gerando uma explosão por um curtíssimo período de tempo (momento denominado *burst*). Aos poucos, a cavidade vai se abrindo e a articulação da plosiva vai dando lugar à produção do segmento adjacente. Esse período é conhecido na literatura como “transição”.

Fant (1973), no entanto, decompõe a plosiva em cinco momentos (ou fases): (1) “oclusão”, momento do contato entre os articuladores que pode ou não apresentar batimento de pregas; (2) “transiente”, que seria a resposta do trato oral à pressão de soltura; (3) “segmento fricativo”, já destacado por Clark e Yallop (1995) como o momento em que ar é liberado, gerando fricção; (4) “segmento aspirado”, no qual um ruído originado na glote é produzido concomitantemente à produção da plosiva; e (5) a “transição”, momento final da plosiva e transição para o próximo segmento. Esses cinco momentos constituem, por exemplo, uma plosiva não-vozeada aspirada, como [t<sup>h</sup>]. As plosivas não

necessariamente apresentam todos esses cinco momentos, como é o caso das plosivas não-vozeadas do PB [p] e [t], que podem não apresentar a fase de aspiração<sup>6</sup>. Quatro dos cinco momentos descritos por Fant (1973) podem ser visualizados na Figura 4.

Figura 4 - Momentos de uma plosiva não-vozeada não-aspirada, descritas por Fant (1973), mostrando os momentos: transiente, fricção, aspiração e vozeamento



Fonte: Stevens (2000, p.348)

É importante salientar que, em ambas as descrições dos momentos de uma plosiva, propostas por Fant (1973) e Clark e Yallop (1995), estão presentes três momentos que são essenciais a todas as plosivas: fechamento, soltura e transição. Portanto, todas as plosivas necessariamente irão apresentar os três momentos descritos por Clark e Yallop (1995), mas nem todos irão apresentar os dois momentos “adicionais”, propostos por Fant (1973).

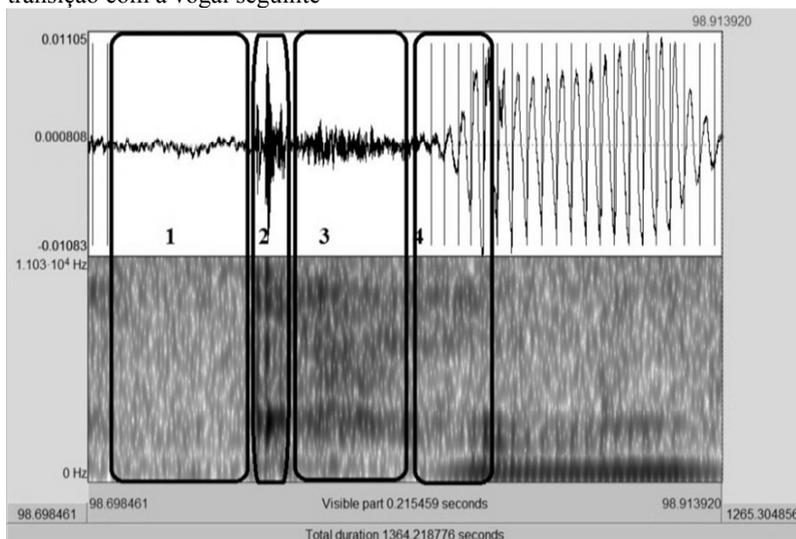
A aspiração é um ruído produzido na glote que pode ou não estar presente durante a fase de soltura de uma plosiva (STEVENS, 2000). Essa fase adicionada à produção da plosiva lhe confere o caráter de “aspirada”, mas não se configura como um parâmetro dicotômico. De acordo com Cho e Ladefoged (1999), as plosivas não-vozeadas das mais diversas línguas apresentam graus de aspiração, a depender da duração do VOT, que pode ou não ser utilizada como distintiva intra e interlínguas (LISKER, ABRAMSON, 1964). As plosivas não-vozeadas

<sup>6</sup> Fant (1973) fala ainda que [p] não irá apresentar momento de fricção por conta da rápida “deslabialização”, ou seja, porque o momento de afastamento entre os articuladores é muito curto para gerar fricção.

do espanhol, por exemplo, não apresentam essa fase e, portanto, não utilizam a aspiração como caráter distintivo, mas as plosivas não-vozeadas do coreano podem ser produzidas com leve ou forte aspiração e são distintas entre si (LISKER; ABRAMSON, 1964a).

Na Figura 5 podemos observar uma imagem espectrográfica que ilustra a produção de uma plosiva aspirada, nela se identificam os momentos de oclusão (1), explosão – em que há a presença do que Fant (1973) aponta como fricção – (2), aspiração (3) e transição (4).

Figura 5 - Oscilograma e espectrograma de uma plosiva [k] aspirada. Em destaque, os momentos: 1-oclusão; 2-soltura/explosão; 3-aspiração; e 4-transição com a vogal seguinte



Fonte: Praat/Nossos dados

Na Figura 5, podemos observar a produção de uma plosiva não-vozeada, pois não há periodicidade que indica batimento de pregas vocais durante a sua produção (sem barra de vozeamento nas regiões de baixa frequência). A periodicidade do sinal só pode ser observada na região que corresponde à vogal adjacente, onde é possível visualizar uma região escura na parte mais inferior do espectrograma ou ainda pelas barras verticais dispostas no oscilograma.

A região 1 corresponde ao momento em que os articuladores entram e se mantêm em contato, formando uma barreira à passagem do fluxo de ar egressivo. O momento 2, que apresenta um pico no espectrograma, corresponde ao período de *burst*, quando os

articuladores rapidamente se afastam, originando um ruído abrupto ou fricção causado pela liberação do contato entre os articuladores. O VOT, região compreendida entre a explosão e o início da vogal adjacente é apontado pela região 3, na qual pode ser observado certo ruído, característico da aspiração que, por sua vez, é gerado na região glotal. As plosivas não-aspiradas têm um VOT mais curto que o observado na Figura 5 e podem não apresentar nenhum grau de aspiração ( $VOT=0$ ) ou ainda um breve período de aspiração ( $0 < VOT < 35$  ms, como é demonstrado pela Figura 4, que exibe um brevíssimo momento de aspiração). Por fim, o momento 4 mostra o fim do segmento plosivo e o início da vogal que o segue, no qual há a transição formântica entre esses segmentos.

## 2.2 PARÂMETROS ACÚSTICOS ENVOLVIDOS NA PRODUÇÃO DOS SONS PLOSIVOS

Durante a produção de um segmento, diferentes movimentos articulatorios são acionados para que o segmento possa se concretizar. Esses movimentos terão eventos acústicos consequentes, que irão refletir diretamente na percepção desse segmento, uma vez que as pistas acústicas utilizadas na sua produção fornecerão subsídios para que o interlocutor possa identificá-lo como pertencendo a uma ou outra categoria. Além dos parâmetros que são intrínsecos<sup>7</sup> ao próprio segmento, os contextos adjacentes também são de grande importância na modificação das suas características e influenciarão as pistas acústicas que serão levadas em consideração na sua identificação.

Assim, se faz necessário o estudo de como se comportam os parâmetros acústicos e articulatorios inerentes às plosivas, bem como o comportamento desses parâmetros em relação aos seus contextos adjacentes. A seguir, iremos abordar, separadamente, esses parâmetros referentes às plosivas, que são tidos na literatura como alguns dos que podem exercer influência na produção desse tipo de segmento.

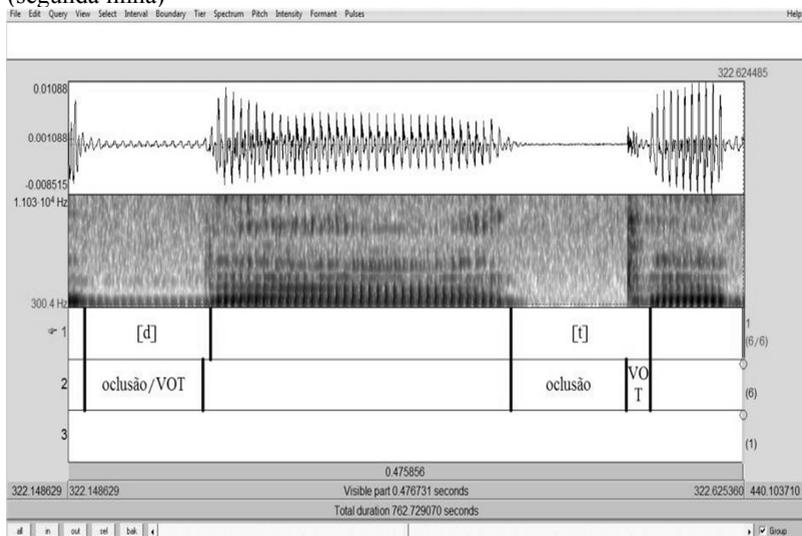
---

<sup>7</sup> Por parâmetros intrínsecos queremos denotar aqueles que fazem parte do próprio segmento, ou seja, aqueles presentes na sua produção e que são parâmetros decorrentes da sua articulação ou são os correlatos acústicos da articulação. Dessa forma, iremos contrapor parâmetros intrínsecos com parâmetros de contexto, que são aqueles que vão depender dos contextos adjacentes. Dentre aqueles avaliados por esta pesquisa destacamos como parâmetros intrínsecos: duração da consoante, da oclusão e do VOT, especificidades acústicas, *burst* e transição formântica; e como parâmetros de contexto: duração e tipologia das vogais adjacentes, tonicidade e posição na palavra.

### 2.2.1 Duração da consoante

Em relação à duração absoluta da consoante, as plosivas não-vozeadas tendem a apresentar maior duração do que a contraparte vozeada (KLATT, 1976; BARROCO *et al.*, 2007; BARBOSA, 1996; FICKER, 2003; CRISTOFOLINI, 2013). Para as consoantes plosivas não-vozeadas, a duração corresponde à duração do bloqueio + VOT; para as consoantes vozeadas, a duração da consoante corresponde ao seu VOT (que, por sua vez, corresponde à duração do bloqueio) e a toda a região compreendida entre o *burst* e o início da vogal seguinte. Uma das razões que pode explicar esse comportamento é a necessidade de uma duração mais curta para que o vozeamento possa ser mantido durante toda a produção das consoantes que apresentam vibração de pregas, como já relatado anteriormente. A Figura 6 apresenta a forma de onda e o espectrograma de uma plosiva vozeada e de uma não-vozeada, com ênfase na duração de cada uma dessas consoantes.

Figura 6 - Forma de onda e espectrograma da palavra "data". Em destaque: duração total das plosivas [d] e [t] (primeira linha) e duração da oclusão e VOT (segunda linha)



Fonte: Praat/Nossos dados

A palavra apresentada na Figura 6 faz parte do nosso banco de dados e mostra que a duração da consoante alveolar não-vozeada é de 98,22ms, enquanto que para [d] é de 88,72ms.

Para o PB, Barbosa (1996) observou, através da coleta de segmentos inseridos em logatomas em fala adulta, que as plosivas não-vozeadas apresentam valores de duração absoluta maiores se comparados à duração das vozeadas. Mesma tendência foi encontrada por Ficker (2003) que avaliou plosivas em posição inicial e tônica, em paroxítonas dissilábicas, coletadas com falantes adultos<sup>8</sup> através da leitura de frases veículos contendo a palavra com o segmento-alvo. Na comparação entre os pares homorgânicos, Ficker (2003) encontrou resultados estatisticamente relevantes para todas as comparações realizadas ( $p = 0,000$ ). A pesquisadora também investigou possíveis diferenças entre plosivas da mesma categoria e encontrou diferenças significativas somente entre [b] e [g] ( $p = 0,000$ ), sendo a duração da velar menor do que a da bilabial.

A Tabela 1 apresenta os valores médios de duração absoluta, encontrados por Barbosa (1996) e Ficker (2003)<sup>9</sup>.

Tabela 1 - Valores médios (ms) de duração absoluta das plosivas do PB

	[p]	[b]	[t]	[d]	[k]	[g]
<b>Barbosa (1996)</b>	120	86	113	71	121	67
<b>Ficker (2003)</b>	118	99	120	91	116	83

Observando a Tabela 1, percebemos que há divergências entre os valores de duração médios entre as plosivas não-vozeadas investigadas nos dois estudos com relação à plosiva que apresenta a menor e maior duração. No estudo de Barbosa (1996), [t] apresenta o menor valor, enquanto [k] o maior. Ficker (2003) encontra a tendência oposta. Nos dois estudos, o valor médio da bilabial é intermediário.

No entanto, esses dois estudos são convergentes com relação à medida de duração de plosivas vozeadas, sendo que as bilabiais exibiram os maiores valores e as velares os menores.

Ademais, Cristofolini (2013), ao avaliar plosivas em contexto tônico em paroxítonas dissilábicas ou trissilábicas na fala infantil, observa valores de duração relativa de plosivas mais elevados para [k] (28,2%), quando comparado com [p] (26,5%) e [t] (25,7%), seguindo a mesma ordem encontrada por Barbosa (1996). No que tange as plosivas

<sup>8</sup> Ficker (2003) avaliou a fala de uma informante adulta com fala típica e de uma informante em idade adulta com desvio fonológico. Os dados de Ficker (2003) que serão aqui apresentados se referem somente aos resultados obtidos para a informante com fala típica.

<sup>9</sup> Os sujeitos avaliados por Barbosa (1996) e Ficker (2003) eram adultos, oriundos do estado de São Paulo, com idades de 30 e 35 anos, respectivamente.

vozeadas, os valores de duração relativa para o estudo de Cristofolini (2013) se mostram maiores para a bilabial (24,2%), seguida pela alveolar (22,8%) e pela velar (21,3%), reiterando o encontrado por Barbosa e Ficker (2003) para as vozeadas. Por fim, os resultados de duração relativa de plosivas de Cristofolini (2013) apontam para a maior duração relativa das plosivas não-vozeadas.

### 2.2.2 Duração da oclusão<sup>10</sup>

A duração da oclusão corresponde à medida temporal entre o início de uma plosiva (o momento em que terminam os pulsos glotais da vogal anterior) até o momento da explosão ou *burst* (LISKER, 1957). Geralmente, esse parâmetro é melhor visualizado em plosivas intervocálicas, pois assim se consegue obter com maior exatidão o momento inicial da oclusão do trato. Segundo Kent e Read (1992, *apud* BARROCO; DOMINGUES; PIRES; LOUSADA; JESUS 2007), a faixa média de duração de oclusão fica entre 50 e 150 ms.

Segundo a literatura, as plosivas não-vozeadas apresentam maior duração de oclusão do que as vozeadas (BARROCO *et al.*, 2007; LADEFOGED; MADDIESON, 1996; LADEFOGED, 2001; LISKER, 1957, 1981; KLATT, 1975). Esse fato se mostra interessante ao se comparar palavras como *cap* (boné) e *cab* (táxi), na língua inglesa. Apesar de a vogal em *cab* ser maior do que em *cap*<sup>11</sup>, as duas palavras tendem a apresentar duração similar, pois a duração da oclusão de [p] acaba compensando essa diferença, por ser mais longa do que em [b]. Esta é uma condição para a produção das plosivas do inglês, que estabelece que plosivas não-vozeadas em final de sílaba tendem a ser mais longas do que a contraparte vozeada quando antecedidas pela mesma vogal (LADEFOGED; JOHNSON, 2011).

De acordo com Johnson (2003), a diferença entre a duração da oclusão de plosivas vozeadas e não-vozeadas pode estar relacionada ao fato de que as primeiras são produzidas com uma menor duração para possibilitar a manutenção do vozeamento durante toda a sua produção.

Lisker (1957) analisou espectrogramas das plosivas intervocálicas [p] e [b], inseridas em palavras da língua inglesa, e constatou que a duração da oclusão é maior para a plosiva não-vozeada. Para essas

<sup>10</sup> As palavras bloqueio, fechamento e oclusão são usadas aqui como sinônimos para denominar o mesmo momento na plosiva. Na presente pesquisa, optamos pela palavra oclusão.

<sup>11</sup> Vogais também tendem a apresentar maior duração quando sucedidas por consoantes vozeadas (LISKER, 1957; STEVENS; KLATT, 1974; LADEFOED; MADDIESON, 1996; LADEFOGED, 2001).

plosivas, duração da oclusão apresentou uma faixa de distribuição de 90-140ms, com valor médio de 120ms, enquanto, para as plosivas vozeadas, a faixa foi de 65-90ms, com média de 75ms.

O estudo de Barroco *et al.* (2007) com duas crianças falantes nativas de português europeu (PE) - uma delas com desvio fonológico - mostrou que a duração da oclusão é maior nas plosivas não-vozeadas, variando de 30 a 180ms, para a criança com aquisição típica da linguagem. Lousada, Martins e Jesus (2006) encontraram que a duração da oclusão de plosivas do PE também foi maior em consoantes não-vozeadas para dados de falantes adultos.

Para o PB, Berti e Marino (2011) investigaram o parâmetro duração da oclusão em relação a [t] e [k], obtendo valores absolutos e relativos à duração do segmento em si. Os resultados demonstraram que a duração da oclusão de [t] ocupa uma maior parte do segmento ( $\pm 60\%$ ) se comparada à mesma medida de [k] ( $\pm 40\%$ ).

Também para o PB, Melo *et al.* (2012)<sup>12</sup> observaram valores de duração de consoante mais elevados para as plosivas não-vozeadas. Na comparação entre plosivas homorgânicas, os valores médios obtidos foram: [p] = 141 ms e [b] = 96 ms ( $p < 0,001$ ); [t] = 133 ms e [d] = 88 ms ( $p < 0,001$ ); e [k] = 118 ms e [g] = 73 ms ( $p < 0,001$ ).

A partir dos resultados encontrados pelos estudos supracitados, podemos inferir que a duração da oclusão de plosivas é maior para as plosivas não-vozeadas, podendo ser uma pista na distinção entre o vozeamento de plosivas com o mesmo ponto de articulação. Além disso, ele pode se diferenciar também quanto ao ponto de articulação, uma vez que os estudos de Berti e Marino (2011) e Melo *et al.* (2012) apresentaram resultados distintos para os diferentes pontos de articulação de uma mesma categoria.

### 2.2.3 Especificidades acústicas

A descrição do vozeamento, do modo e do ponto de articulação que observamos acima é aquela descrita pela literatura como determinante de um som plosivo típico ou normal. No entanto, os segmentos nem sempre são produzidos como esperado. Na fala, muitos fenômenos podem ocorrer, fazendo com que segmentos possam sofrer alguma interferência de, por exemplo, uma velocidade de fala mais rápida (KLATT, 1976; ALVES; DIAS, 2010).

---

<sup>12</sup> A pesquisa de Melo *et al.* (2012) avaliou a fala de adultos como grupo de controle para a comparação com um grupo de crianças com desvio fonológico (população-alvo).

Denes (1955) afirma que muitos sons aperiódicos ou não-harmônicos<sup>13</sup> são percebidos como vozeados. Isso é claramente visível através da fala sussurrada quando, mesmo sem a presença de vozeamento, podemos produzir sons tidos como vozeados e que podem ser inteligivelmente percebidos por um interlocutor. Liberman *et al.* (1958) afirmam que essa identificação é possível porque outros parâmetros são também levados em consideração quando o ouvido humano vai identificar os segmentos da fala e que há nesses segmentos, características acústicas que podem diferenciá-los.

Muitos estudos vêm relatando a existência de dados que apresentam características não esperadas para a produção típica, como por exemplo, a presença de vozeamento em segmentos não-vozeados, desvozeamento parcial de segmentos vozeados ou, ainda, segmentos que parecem estar entre as duas categorias. Esse fenômeno ocorre tanto na fala infantil quanto na fala adulta<sup>14</sup> (CRISTOFOLINI, 2008; 2013; FORREST, WEISMER, MILENKOVIC, DOUGALL, 1988; BONATTO, 2007; MELO *et al.*, 2011, BONATTO; MADUREIRA (2009); dentre outros).

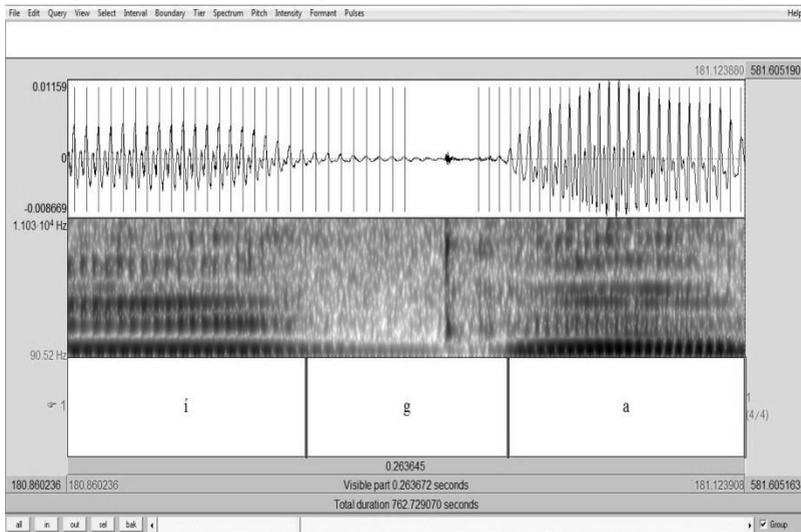
Clark e Yallop (1995) afirmam que o desvozeamento parcial é possível de ocorrer, pois há um “desequilíbrio aerodinâmico” promovido pela pressão abaixo e acima da glote que faz com que o vozeamento não consiga ser sustentado. Como exemplo de plosivas parcialmente desvozeadas, os autores citam aquelas em posição final de palavra no inglês, como em *rib* (costela), *rid* (livrar) e *pig* (porco). Um caso de interrupção no vozeamento pode ser visualizado através da Figura 7.

---

<sup>13</sup> Um som periódico é aquele em que há vibração de pregas. Quando um som desse tipo é produzido, o resultado é uma onda denominada complexa, formada pela soma de várias ondas periódicas. A frequência fundamental se refere ao número de vezes que essa onda complexa se repete durante um segundo. Ao decompor essa onda periódica (através de uma Transformada de Fourier), iremos obter as frequências componentes, sendo a primeira delas a frequência fundamental. As demais frequências serão múltiplas dessa frequência fundamental e são denominadas “harmônicos”. Assim, se a frequência fundamental for de 200Hz, o segundo harmônico será de 400Hz, o terceiro de 600Hz e assim por diante (LADEFOGED, 2001; JOHNSON, 2003). Dessa forma, um som não-periódico ou não-harmônico é aquele que não apresenta ondas periódicas durante a sua produção, por não apresentar vibração das pregas vocais.

<sup>14</sup> Segundo estudos, é mais comum encontrarmos especificidades na fala infantil do que na adulta, uma vez que crianças estão ainda em fase de refinamento articulatorio até chegar à idade adulta. Cristofolini (2013), ao analisar a fala infantil de crianças entre 6 e 12 anos, observou a ocorrência de muitas especificidades e que elas diminuíam à medida que a idade aumentava.

Figura 7 - Forma de onda e espectrograma da palavra “bígamo”. Em destaque: a interrupção no vozeamento na parte medial de [g]



Fonte: Praat/Nossos dados

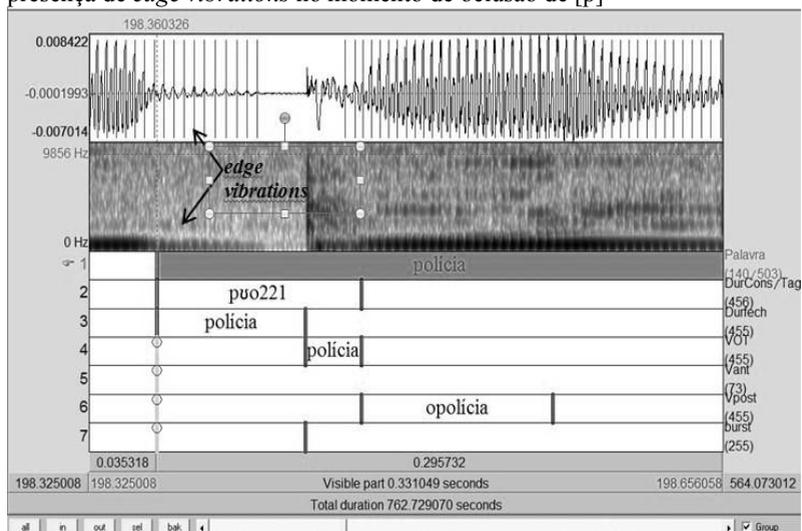
Já, no caso das plosivas não-vozeadas, é comum encontrar batimento de pregas na sua porção inicial, pois, apesar de haver uma interrupção do batimento, as pregas levam um tempo até cessar completamente o movimento. Ladefoged e Maddieson (1996) afirmam que esse fenômeno geralmente acontece com plosivas não-vozeadas em posição intervocálica, pois as pregas continuam em movimento até que haja abertura de glote suficiente para que elas possam parar de vibrar. Já Stevens (2000) aponta que a vibração glotal pode se estender de 10 a 20 ms durante o momento de oclusão da plosiva não-vozeada, mas com redução de amplitude comparada à vogal anterior. Essa vibração de baixa amplitude presente no período de oclusão das não-vozeadas foi observada por Lisker (1957) na língua inglesa, em dados de fala adulta. Em um estudo posterior, Lisker e Abramson (1964b) classificaram esse fenômeno como *edge vibrations*:

Spectrograms suggest that the laryngeal oscillations of the preceding voiced environment may simply continue for a while even after the glottis has begun to open for a voiceless stop; these vibrations are so low in intensity that any auditory effect they might have by themselves

seems to be masked out by the stop burst and the noise of turbulent air rushing through the glottis. (p.417)<sup>15</sup>

Através da Figura 8, podemos observar a presença de *edge vibrations* em uma plosiva não-vozeada bilabial na palavra “polícia”, extraído do nosso banco de dados.

Figura 8 – Forma de onda e espectrograma da palavra “polícia”. Em destaque: a presença de *edge vibrations* no momento de oclusão de [p]



Fonte: Praat/Nossos dados

Snoeren (2005), em uma pesquisa sobre assimilação na língua francesa, observa que existe diferentes faixas de categoria de assimilação e que nem todas são percebidas pelo ouvido humano para modificar o julgamento de uma plosiva não-vozeada como vozeada. O pesquisador aponta ainda que vozeamento de 0 a 20% seria considerado “desprezível” pelo ouvido humano, que vozeamento de 20 a 80% seria uma categoria intermediária e que acima disso seria um vozeamento completo.

<sup>15</sup> “Espectrogramas mostram que as oscilações da laringe de um contexto vozeado anterior podem continuar por certo tempo mesmo após a glote ter iniciado a abertura para uma plosiva não-vozeada; essas vibrações apresentam intensidade tão baixa que qualquer efeito audível que elas possam gerar pode ser mascarados pelo *burst* da plosiva e pelo ruído turbulento que passa pela glote.” (LSKER; ABRAMSON, 1964b, p.417). Minha tradução.

Cristofolini (2013) constatou, a partir da inspeção da forma de onda e do espectrograma dos dados coletados com crianças entre 6 e 12 anos e adultos, falantes nativos do PB, a presença desse e de outros tipos de dado com alguma especificidade ou ainda diferentes daqueles apresentados pela literatura, os quais a autora nomeou como “especificidades acústicas”. De um total de 3990 dados, a pesquisadora constatou a presença de 630 que continham algum tipo de especificidade, totalizando 16% dos dados.

Um fator interessante nesse estudo foi a comparação do percentual de produções atípicas conforme a idade dos participantes. Os resultados obtidos pela pesquisadora estão resumidos na Tabela 2, retirada do estudo de Cristofolini (2013).

Tabela 2 - Resultados de Cristofolini (2013) para as especificidades acústicas, conforme a faixa etária dos participantes

<i>faixa etária</i>	<i>6 anos</i>	<i>8 anos</i>	<i>10anos</i>	<i>12anos</i>	<i>adulto</i>
total de especificidades	214	157	112	96	51
% em relação ao total (3990) de dados totais	5,9%	3,9%	2,8%	2,4%	1,3%
% em relação ao total (630) de especificidades	34%	25%	18%	15%	8%

Fonte: Cristofolini, 2013, p.189

Podemos constatar, a partir dos resultados apresentados por Cristofolini (2013), que a presença de dados com alguma especificidade decresce à medida que a idade aumenta e a produção da fala vai se estabilizando, mas que não deixa de estar presente na fala adulta.

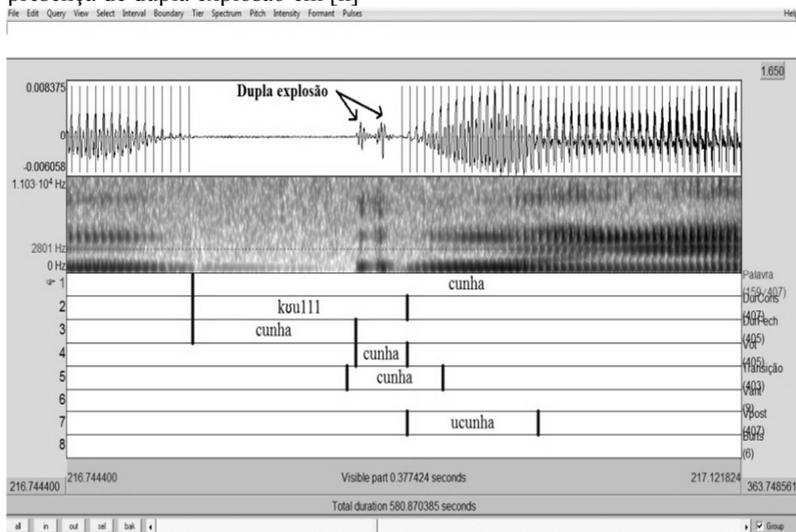
Ainda nesse estudo, a autora propõe a classificação dos tipos de especificidades encontradas por ela nesse e em um estudo anterior (CRISTOFOLINI, 2008). Utilizaremos a classificação proposta pela autora para a análise qualitativa do que propomos nesta tese chamar de “especificidades acústicas”, por serem produções que se afastam, em algum grau, do que é considerado como típico pela literatura. A classificação de cada um dos tipos de especificidades é apresentada abaixo e exemplificada com amostras encontradas em nosso banco de dados:

- (i) **Plosivas não-vozeadas totalmente vozeadas:** são aquelas em que há a presença de vozeamento (visualizada pela barra de vozeamento nas regiões de baixa frequência e pelas estriações

de pulsos glotais em espectrogramas) durante toda a duração da oclusão.

- (ii) **Plosivas não-vozeadas com vozeamento na parte inicial (*edge vibrations*):** como mencionado acima, são segmentos que apresentam vozeamento na parte inicial. A Figura 8 apresentou um exemplar desse tipo de dado. No estudo de Cristofolini (2013), esse dado correspondeu a 15% dos dados com especificidades e ocorreu, majoritariamente para [p].
- (iii) **Plosivas com duplas explosões:** são aquelas que apresentam dois picos de energia durante o *burst*, correspondendo a duplas explosões, conforme ilustrado pela Figura 9. Os resultados de Cristofolini (2013) apontaram 7% de dados com essa característica, a sua maioria em plosivas [k].

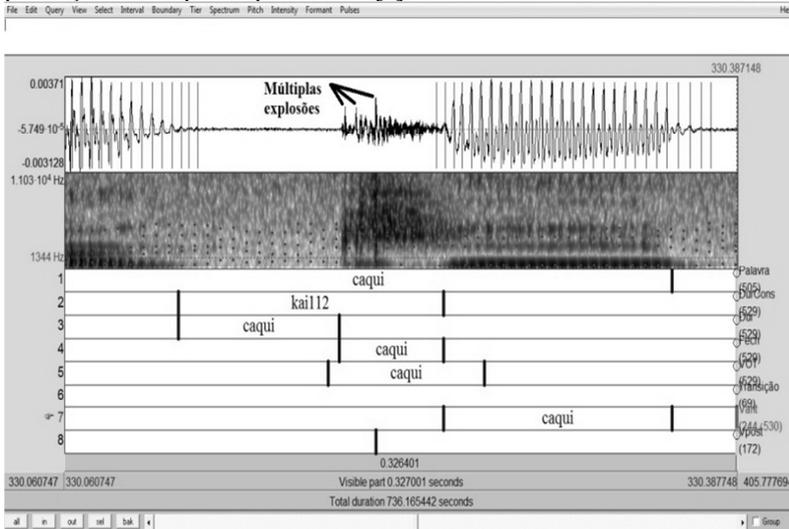
Figura 9 – Forma de onda e espectrograma da palavra “cunha”. Em destaque: a presença de dupla explosão em [k]



Fonte: Praat/Nossos dados

- (iv) **Plosivas com múltiplas explosões:** da mesma forma que a categoria anterior, essas plosivas apresentam três ou mais picos de explosão durante o *burst* (Figura 10). Cristofolini (2013) observou plosivas não-vozeadas com múltiplas explosões em 7% dos dados, especialmente em [k].

Figura 10 – Forma de onda e espectrograma da palavra “caqui”. Em destaque: a presença de múltiplas explosões em [k]

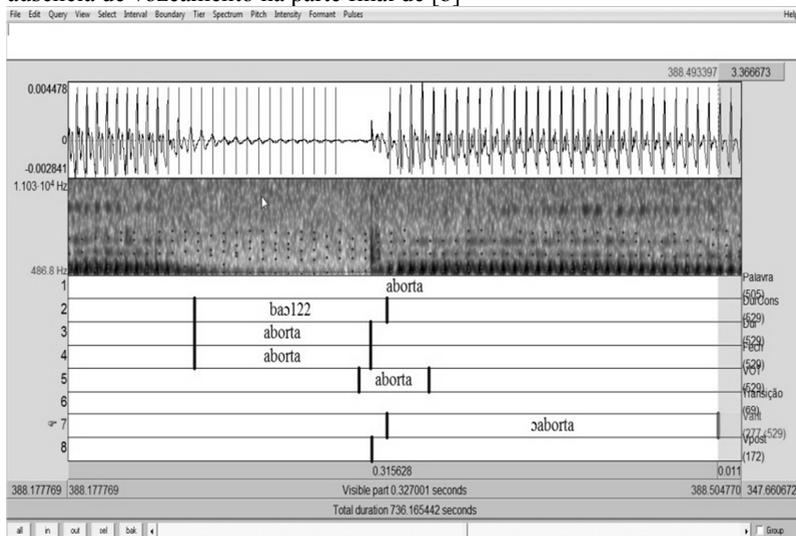


Fonte: Praat/Nossos dados

- (v) **Plosivas vozeadas com interrupção do vozeamento no início do segmento:** é o tipo de plosiva na qual o vozeamento é interrompido na porção inicial do segmento. Cristofolini (2013) encontrou esse tipo de dado em 0,16% do total (plosiva [g]).<sup>16</sup>
- (vi) **Plosivas vozeadas com interrupção do vozeamento no meio do segmento:** é o tipo de plosiva na qual o vozeamento é interrompido na porção intermediária do segmento. Em Cristofolini (2013), esses dados corresponderam a 1,75% do total e foi exclusivamente encontrado para a plosiva [g]. Podemos observar um exemplar na Figura 7.
- (xii) **Plosivas vozeadas com interrupção do vozeamento no final do segmento:** é o tipo de plosiva na qual o vozeamento é interrompido na porção final do segmento. Essa especificidade foi encontrada em 22% dos dados com especificidade de Cristofolini (2013), especialmente para [b]. Podemos observar um exemplar pela Figura 11.

<sup>16</sup> Não encontramos nenhum exemplar dessa especificidade em nossos dados.

Figura 11 - Forma de onda e espectrograma da palavra “aborta”. Em destaque: a ausência de vozeamento na parte final de [b]



Fonte: Praat/Nossos dados

- (xiii) **Plosivas vozeadas totalmente desvozeadas**: plosivas vozeadas que não apresentam vozeamento durante toda a sua produção.

### 2.2.4 *Burst*

Outra característica das plosivas é a dispersão de energia ao longo do espectro de frequência no momento da soltura da oclusão. Essa energia, no entanto, não se encontra aleatoriamente dispersa, mas sim concentrada em determinadas faixas de frequência para cada uma das plosivas (COOPER, DELATTRE, LIBERMAN, BORST, GERSTMAN 1952; LIBERMAN; DELATTRE; COOPER, 1952; LIBERMAN, DELATTRE, COOPER, GERSTMAN, 1954).

Lieberman *et al.* (1954) afirmam que o *burst* é o correlato acústico da explosão articulatória, isto é, da liberação do ar retido atrás da oclusão. A sua duração pode ser de 5 a 15 milissegundos (LIEBERMAN; BLUMSTEIN, 1988). Acusticamente, ele se configura como um ruído de curtíssima duração, que pode ser visualizado em um espectrograma pela súbita mudança no sinal de uma plosiva após o período de oclusão do trato. Essa explosão é causada pelo aumento da

pressão na cavidade, que faz com que o ar atrás da oclusão saia com grande velocidade.

O *burst* será moldado apenas pelos tubos ressoadores na cavidade formada na região anterior à oclusão da plosiva, pois as cavidades anterior e posterior à oclusão estão desacopladas, isto é, não formam um único tubo. Assim, as frequências ressoadoras do tubo posterior à oclusão acabam não influenciando na energia espectral do *burst*. Essas frequências só voltarão a atuar na medida em que ocorra a soltura da oclusão e que as cavidades anterior e posterior à oclusão voltem a se acoplar, modulando, assim, as frequências ressoadoras na transição da consoante para a vogal (JOHNSON, 2003).

As plosivas bilabiais não possuem cavidade anterior à oclusão dos lábios, fazendo com que a energia espectral do *burst* se dê em função das propriedades relativas à fonte sonora e não tenha nenhum tipo de modulação no trato oral. Por essa razão, o *burst* para essas consoantes tenderá a apresentar concentração de energia nas baixas frequências. As plosivas alveolares apresentam uma cavidade pequena na frente da oclusão, fazendo com que o espectro do *burst* se concentre nas altas frequências. Por terem uma cavidade maior do que as alveolares, as plosivas velares apresentam atenuação nas frequências concentradas, assim, nas zonas de médias frequências.

Halle, Hughes e Radley (1957) afirmam que o *burst* é um parâmetro com grande relevância na percepção de plosivas. Eles encontraram valores de espectro de *burst* para as plosivas do inglês que se dão da seguinte maneira: para [p] e [b], a energia do *burst* se concentra entre 500 e 1500 Hz; para [t] e [d], acima de 4000 Hz e para as velares [k] e [g] entre 1500 e 4000 Hz. A variação observada está relacionada à qualidade da vogal que as segue. Os pesquisadores ressaltam que, para as velares, o pico espectral do *burst* apresentou valores muito mais altos para as plosivas seguidas de vogais anteriores do que para aquelas seguidas de vogais posteriores. Os autores atribuem essa diferença à tendência de anteriorização da oclusão das plosivas que antecedem vogais anteriores. O F2 de vogais anteriores do inglês está entre 1200-1500 Hz e o pico de *burst* das plosivas que as antecediam se estabeleceu na região de 2000-4000Hz. No caso das plosivas velares que antecediam vogais posteriores (com F2 abaixo de 1200 Hz), elas apresentaram valores inferiores a 2000 Hz, mostrando que a diferença de frequência entre as vogais pode acarretar em uma atenuação no pico espectral do *burst* da plosiva que a antecede. Em suma, a relação que os

pesquisadores encontraram é que, quanto maior o F2 da vogal, mais elevado tende a ser o nível de concentração de energia.

Cooper *et al.* (1952) já haviam encontrado os mesmos resultados de Halle, Hughes e Radley (1957) com relação à concentração de energia de *burst* entre os três pontos de articulação das consoantes plosivas do inglês. Em todos os seus dados, as plosivas identificadas como alveolares encontravam-se em frequências mais altas, independentemente da qualidade da vogal adjacente. As velares foram percebidas quando se encontravam em regiões logo abaixo do nível de F2 da vogal seguinte e todos os outros casos foram identificados como [p].

Yavas (2011) e outros pesquisadores também apontam para as diferentes regiões de concentração de energia de *burst*, mostrando que as bilabiais são as que apresentam concentração em regiões mais baixas, enquanto as alveolares são as que apresentam concentração em regiões mais altas. Para as velares, a concentração de energia de *burst* parece ser dependente da qualidade da vogal que segue a plosiva (LIEBERMAN; BLUMSTEIN, 1988).

Forrest *et al.* (1988) investigaram dados de plosivas não-vozeadas do inglês americano a partir da fala de 10 informantes adultos que apresentavam fala típica. Para avaliar as características espectrais do *burst* de plosivas não-vozeadas, os autores utilizaram as seguintes medidas estatísticas: centroide ou média, variância, assimetria e curtose, o que chamaram de “momentos espectrais”. Segundo esses pesquisadores, essas medidas conseguem apresentar satisfatoriamente o *peakedness*, a concentração e a inclinação da distribuição de energia do espectro. Apesar disso, nem todas as medidas foram distintas para essas plosivas. Bilabial e alveolar se assemelham quanto à curtose, mas se diferenciam com relação à centroide e assimetria. Já velares e bilabiais são semelhantes quanto à centroide e à assimetria, mas não com relação à curtose. A única medida que não exerceu influência na distinção das plosivas, para o estudo de Forrest *et al.* (1988), foi a variância.

Berti e Marino (2011) avaliaram a fala de duas crianças nativas do PB, uma que apresentava desvio fonológico e outra que apresentava aquisição típica da linguagem. Como *corpus* da pesquisa, as autoras utilizaram as plosivas [t] e [k], inseridas em sílabas tônicas, e seguidas das vogais [a] e [u]. As pesquisadoras utilizaram as mesmas medições propostas por Forrest *et al.* (1988) para a investigação das características do *burst*. Os resultados dessas medições, a partir dos dados do sujeito com fala típica, mostraram que, para [ta], o pico espectral médio foi de 4530 Hz, enquanto, para [ka], foi de 2125 Hz, sendo essa diferença

significativa ( $p = 0,03$ ). Para as mesmas plosivas seguidas de [u], os valores médios encontrados foram de 5492 Hz e 5444 Hz, respectivamente, mas testes estatísticos não apontaram diferenças significativas entre elas. Para os outros momentos espectrais, as pesquisadoras só não encontraram diferenças significativas entre [t] e [k] para a variância, da mesma forma que Forrest *et al.* (1988) para o inglês.

Posteriormente, Cristofolini (2013) investigou a influência das características espectrais de *burst* entre crianças com diferentes faixas etárias e adultos. A autora encontrou valores diferenciados para cada um dos momentos espectrais, apesar de não apresentar valores referentes à relevância estatística entre cada plosiva, mas somente entre cada grupo de sujeitos avaliados.

Concluimos, a partir dos estudos analisados, que o espectro do *burst* é distinto para as plosivas e se configura como uma importante pista na distinção entre diferentes pontos de articulação das plosivas de uma mesma classe, pois cada uma delas irá apresentar diferentes concentrações de energia, na seguinte ordem: bilabiais < velares < alveolares.

### 2.2.5 VOT e aspiração

Um dos parâmetros mais utilizados na caracterização das plosivas é o *voice onset time* (VOT), descrito primeiramente por Lisker e Abramson (1964; 1965) e constantemente utilizado nos mais diversos estudos sobre esse tipo de segmento.

Como descrito anteriormente, a plosiva se caracteriza por apresentar um momento de oclusão, visualizado através de espectrograma como um período de silêncio, no qual os articuladores estão em contato, bloqueando o trato oral e impedindo o fluxo de ar, seguido de uma explosão causada pela liberação do ar contido atrás da oclusão. O VOT é a medida do tempo compreendido entre a soltura da oclusão de uma plosiva e o início do vozeamento do segmento adjacente. Como ele é um parâmetro temporal, é medido em milissegundos (ms).

Apesar de essa medida ser variável intra e interlínguas, em um estudo com 11 idiomas, Lisker e Abramson (1964) observaram que existem três categorias de VOT que dão conta de englobar as plosivas das línguas que investigaram. A primeira delas é a chamada de pré-sonorização (*voicing lead*), que abrange as plosivas que apresentam vozeamento antes da soltura da oclusão. Nessa categoria, inserem-se,

por exemplo, as plosivas vozeadas do espanhol e também do português brasileiro (KLEIN, 1999). Como o vozeamento se inicia antes da soltura, essa categoria é a única a qual são atribuídos valores negativos de VOT. Outra categoria é a de retardo curto (*short lag*), que engloba as plosivas que apresentam um pequeno momento entre a soltura e o início do vozeamento. Apesar de apresentar valores de VOT positivos, esse momento é considerado curto, com valores de 0-60ms. Por fim, a categoria de retardo longo (*long lag*), que compreende aquelas plosivas que apresentam um tempo considerável entre a soltura e o vozeamento. Os valores de VOT para essa categoria estão acima de 60ms.

Dentre as 11 línguas analisadas por Lisker e Abramson (1964), existem aquelas em que as plosivas se inserem em duas ou em três categorias, como o inglês, o espanhol e o coreano. Há até mesmo aquelas que apresentam quatro categorias, pois, além das três descritas, apresentam plosivas vozeadas (pré-sonorização) com aspiração (característica geralmente presente apenas nas plosivas de retardo longo).

Contudo, as línguas não apresentam plosivas com valores iguais de VOT. Sendo considerado um parâmetro contínuo, cada língua vai posicionar as suas plosivas dentro de uma determinada faixa de valores de VOT encaixados dentre esses três pontos de referência. Desse modo, essas faixas irão englobar cada categoria de plosiva presente em uma determinada língua (CLARK; YALLOP, 1995). Uma prova disso são estudos de VOT com variadas línguas (como o de Lisker e Abramson, 1964 e Cho e Ladefoged, 1999, por exemplo) que mostram que, apesar de apresentarem plosivas dentro de uma mesma categoria (e.g. plosiva alveolar não-vozeada não-aspirada), as faixas de VOT são distintas para essas línguas. É o caso de plosivas bilabiais não-vozeadas do inglês e do espanhol no estudo de Lisker e Abramson (1970). No caso do inglês, a faixa de VOT de [p] se estendeu de 20 a 90 ms e, portanto, entre retardo curto e longo; enquanto, para o espanhol, de 0 a 20 ms, faixa dentro do retardo curto apenas.

Uma constatação importante feita por Lisker e Abramson (1964) foi a influência do ponto de articulação das plosivas não-vozeadas sobre o VOT. Em seus dados, quanto mais posterior o ponto de oclusão, maior foi o valor atribuído ao VOT. Esse achado também se repetiu em outros estudos, para as mais diversas línguas (ALVES, 2011; BARROCO et al., 2007; CHO; LADEFOGED, 1999; COHEN, 2004; KLEIN, 1999; FISCHER-JORGENSEN, 1954; LOFQVIST, 1980; dentre outros).

Nem sempre as consoantes plosivas são diferenciadas pela presença ou não de vozeamento. Este é o caso dos dois conjuntos de

plosivas do inglês que, em início de palavra, são produzidas sem vozeamento (LADEFOGED, 2001; LISKER; ABRAMSON, 1964;1967). A característica fundamental para diferenciá-las nessa posição é a aspiração, definida por Ladefoged e Maddieson (1996) como

a period after the release of a stricture and before the start of regular voicing (or the start of another segment, or the completion of an utterance) in which the vocal folds are markedly further apart than they are in modally voiced sounds. (LADEFOGED; MADDIESON, 1996, p.70)<sup>17</sup>

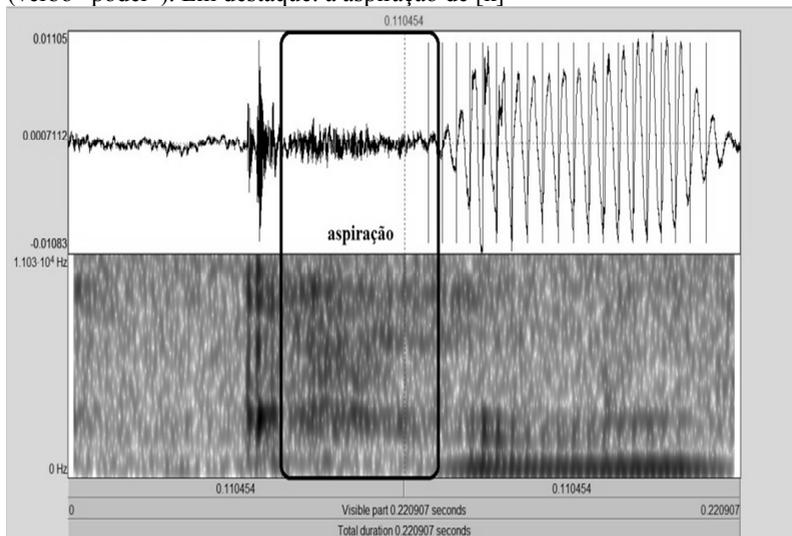
Para a língua inglesa, todas as plosivas em início de palavras não apresentam vozeamento, e a presença de aspiração as caracteriza como [p t k] e a sua ausência como [b d g] (LISKER; ABRAMSON, 1964b; LISKER, 1957). A aspiração também é a característica fundamental na diferenciação dos dois conjuntos de plosivas do dinamarquês, como explica Fischer-Jørgensen (1954). Segundo esse autor, as plosivas [p t k] dessa língua são fortemente aspiradas, enquanto [b d g] são não-aspiradas ou levemente aspiradas.

Ladefoged e Maddieson (1996) afirmam que a aspiração tem um caráter temporal, pois se estende desde a soltura da oclusão até o momento de início do vozeamento. Além disso, Abramson e Lisker (1965) a definem a partir de um aspecto fisiológico como uma “excitação turbulenta” durante a produção de um elemento não-vozeado. Sua principal manifestação acústica se dá pela presença de ruído entre as médias e altas frequências, entre o segundo e o terceiro formantes (LIBERMAN; DELATTRE; COOPER, 1958; LISKER; ABRAMSON, 1964b). A Figura 12 ilustra uma plosiva não-vozeada aspirada.

---

<sup>17</sup> ““(…) um período após a soltura da oclusão e antes do início do vozeamento regular (ou o início de outro segmento, ou o final de uma sentença) no qual as pregas vocais estão marcadamente mais afastadas do que elas estão nos sons vozeados.” (LADEFOGED; MADDIESON, 1996, p.70). Minha tradução.

Figura 12 – Forma de onda e espectrograma da palavra de língua inglesa *can* (verbo “poder”). Em destaque: a aspiração de [k]



Fonte: Praat/Nossos dados

A área em destaque na Figura 12 corresponde à aspiração na palavra de língua inglesa *can* (verbo “poder”) e tem duração de 54 ms. Na forma de onda, vê-se, na parte em destaque, que a onda apresenta um ruído similar a uma pequena fricção<sup>18</sup> e, no espectrograma, pode-se perceber esse ruído pela parte mais escura da área em destaque (especialmente nas médias e altas frequências), logo após a explosão.

Em termos articulatórios, plosivas não-vozeadas aspiradas ou não-aspiradas se diferenciam pelo tempo de abertura de glote. Nas plosivas não-vozeadas aspiradas, o movimento de abertura glotal se inicia ainda durante a oclusão, atingindo o máximo de abertura no momento em que ocorre a soltura da oclusão. Em contrapartida, nas plosivas não-vozeadas não-aspiradas, a abertura da glote acontece no momento de oclusão, mas o máximo de abertura ocorre na metade do tempo da soltura da oclusão e, portanto, tem menor extensão.

Apesar de ter papel fundamental na diferenciação de pares de plosivas para algumas línguas, a aspiração não parece ser uma característica dicotômica, ou seja, que está ou não presente na produção de determinado segmento. Fischer-Jørgensen (1954) afirma que, para as

18 Ladefoged (2001) sugere que a aspiração, por se parecer com uma fricção, deve ser representada numa transcrição estreita com um [h] sobrescrito seguindo a consoante.

plosivas [p t k] do dinamarquês, a aspiração é muito variável e é dependente de algumas características, como ritmo e acento.

Browman e Goldstein (1992) também ressaltam esse caráter variável da aspiração. Na teoria da Geometria de Traços, as plosivas aspiradas do inglês recebem o traço [+spread glottis] e as não aspiradas o traço [-spread glottis]. As autoras explicam que essa regra categórica não seria apropriada, pois, em muitos casos em que a plosiva é considerada não-aspirada, há certa abertura de glote durante a sua produção. Para elas,

presence or absence of aspiration in the output is generally not a discrete function of whether or not the glottis is spread, but rather is either a function of the timing of the glottal gesture with an associated oral gesture or a (gradient) function of the magnitude of the glottal gesture (BROWMAN; GOLDSTEIN, 1992, p.33)<sup>19</sup>

Assim, ao se comparar os dois conjuntos de plosivas do inglês em posição inicial, vemos que a diferença entre eles se dá pelo grau de abertura glotal que, por ser maior nas plosivas [p t k], vai propiciar a aspiração. Portanto, o grau de abertura glotal e o grau de aspiração são diretamente proporcionais (LADEFAGED, 2001).

No coreano, a aspiração assumirá caráter distintivo entre o conjunto [p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>], considerado não-vozeado levemente aspirado e o conjunto [p<sup>c</sup> t<sup>c</sup> k<sup>c</sup>], que é não-vozeado e fortemente aspirado (LISKER; ABRAMSON, 1964). Através da Tabela 3, podemos comparar os valores médios de VOT encontrados por Lisker e Abramson (1964) para as plosivas aspiradas do inglês e do coreano, produzidas em sentenças isoladas.

Tabela 3 - Valores médios de VOT (ms) para as plosivas que apresentaram maior aspiração no inglês e no coreano, no estudo de Lisker e Abramson (1964)

Plosivas	Inglês	Coreano
[p <sup>h</sup> ]	58	91
[t <sup>h</sup> ]	70	96
[k <sup>h</sup> ]	80	126

<sup>19</sup> “(...) presença ou ausência de aspiração no *output* caracterizada por a glote estar ou não aberta, mas sim está ligada ao *timing* do gesto glotal com um gesto oral associado ou, ainda, ligado (de forma gradiente) à magnitude do gesto glotal.” (BROWMAN; GOLDSTEIN, 1992, p.33). Minha tradução.

Podemos perceber, a partir dos valores acima, que cada língua selecionará diferentes valores para os seus respectivos grupos de plosivas aspiradas, fato que comprova que a aspiração não pode ser considerada um parâmetro discreto ou dicotômico.

No entanto, existem línguas que não utilizam a aspiração como parâmetro distintivo de plosivas. Esse é o caso do PB, que apresenta diferentes produções para as suas plosivas não-vozeadas, sendo possível encontrar produções sem ou com leve aspiração (ALVES, 2011; ALVES; SEARA; PACHECO; KLEIN; SEARA, 2008; KLEIN, 1999). Para diferenciar as duas produções, Alves (2011) utilizou a representação [p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>], seguindo a indicação do IPA e de Ladefoged (2001), a mesma que iremos adotar no presente estudo. Apesar da aspiração se apresentar, possivelmente, como um caso de variação entre as plosivas não-vozeadas do PB, ela pode servir como uma pista a mais na distinção entre plosivas vozeadas e não-vozeadas, pois as primeiras não apresentam aspiração para essa língua.

Löfqvist (1980) afirma que o mesmo ocorre com as plosivas não-vozeadas do sueco, que apresentam uma contraparte aspirada. Apesar de a diferença entre plosivas aspiradas e levemente aspiradas nessa língua não ser contrastiva, o autor afirma que ela auxilia na distinção entre o par não-vozeado e vozeado (este último não apresenta aspiração).

A aspiração e o VOT são dois parâmetros das plosivas que estão interligados. A aspiração, quando existe, acontece dentro do intervalo do VOT. A única diferença entre eles (no domínio temporal) é o momento da explosão, que será englobado pelo VOT, mas não pela aspiração. Outra característica similar aos dois parâmetros é que, assim como a aspiração, o VOT também é gradiente, apresentando um *continuum* de possibilidades (em termos de duração em milissegundos). Dentro deste *continuum*, irão existir determinadas faixas que abarcarão as plosivas vozeadas, plosivas não-vozeadas não-aspiradas, levemente aspiradas e aspiradas (LADEFOGED, 2001).

Cho e Ladefoged (1999) analisaram o VOT de plosivas não-vozeadas de 18 línguas e observaram as variações existentes dentro do *continuum* do VOT. Eles afirmam que essa variação é possível porque o VOT sofre influência de determinados aspectos relacionados ao trato. O primeiro deles é o tamanho da cavidade supraglótica formada atrás da oclusão. Para uma plosiva velar, por exemplo, o tamanho dessa cavidade é relativamente menor do que para as plosivas bilabiais e alveolares. Os pesquisadores afirmam que, quanto menor essa cavidade, maior será a pressão de ar retido nela, e isso fará com que leve mais tempo para que a

pressão supraglótica diminua. Assim, o VOT das plosivas velares tenderá a ser maior do que para as demais plosivas.

Da mesma forma, ao se analisar a cavidade formada na frente da oclusão, veremos que ela é relativamente maior para as plosivas velares do que para as alveolares<sup>20</sup>, o que, segundo os autores, irá dificultar a liberação do ar retido atrás da oclusão, pois há uma maior massa de ar na frente da oclusão para ser movimentada no caso das velares.

Outro fator que pode explicar a maior duração do VOT das velares é a massa relativa aos articuladores que formam a oclusão. O articulador ativo das plosivas velares é o corpo da língua, que possui maior massa se comparado à ponta da língua, articulador ativo das alveolares, por exemplo. Assim, leva um tempo maior para que o articulador de maior massa seja movimentado, contribuindo para que o tempo do VOT seja também maior.

Em seus dados, Cho e Ladefoged (1999) encontraram valores médios de VOT maiores para as plosivas velares, em todas as 18 línguas pesquisadas. Em relação às plosivas bilabiais e alveolares, esses autores não encontraram diferenças significativas entre as médias de VOT, fato que vai ao encontro de outros estudos com várias línguas, dentre elas o inglês, o PE e o PB (ALVES, 2011; ALVES *et al.*, 2008; COHEN, 2004; LISKER; ABRAMSON, 1964a, 1971; LOUSADA; MARTINS; JESUS, 2006).

Ainda, no mesmo estudo, Cho e Ladefoged (1999) demonstraram que cada língua apresentou diferentes faixas de VOT para cada categoria de plosiva existente em seu inventário fonológico. Diferentemente de Lisker e Abramson (1964), que apresentaram três categorias para as plosivas, dentre elas, duas para as plosivas não-vozeadas (conforme descrito anteriormente), Cho e Ladefoged (1999) apresentam seus dados de plosivas não-vozeadas distribuídos em quatro faixas, levando em consideração a aspiração. Assim, eles conseguiram demonstrar a variação existente entre plosivas com aspiração que, anteriormente, eram consideradas como pertencentes a uma única categoria, a de retardo longo. A Tabela 4 apresenta as quatro categorias de VOT indicadas pelos autores<sup>21</sup>.

<sup>20</sup> As plosivas bilabiais não apresentam cavidade na frente do bloqueio, pois tem como ponto de articulação a região dos lábios.

<sup>21</sup> No referido estudo, Cho e Ladefoged (1999) analisaram apenas plosivas não-vozeadas. Por essa razão, as quatro categorias apresentam apenas valores de VOT positivos. Plosivas vozeadas são, na maioria das línguas, não-aspiradas, com exceção de línguas como o Thai, que apresentam também plosivas vozeadas aspiradas (Lisker e Abramson, 1964). Apesar disso, nenhuma língua conhecida apresenta plosivas vozeadas aspiradas sem apresentar a contraparte não-aspirada.

Tabela 4 – Categorias de plosivas não-vozeadas, de acordo com o VOT, propostas por Cho e Ladefoged (1999), com valores aproximados.

<b>Categoria</b>	<b>VOT</b>
Plosivas sem aspiração	0-35 ms
Plosivas levemente aspiradas	36-55 ms
Plosivas aspiradas	56-95 ms
Plosivas fortemente aspiradas	96-150 ms

Essa categorização possibilita não só a distinção das categorias de plosivas dentro de cada língua, mas fornece também uma base plausível para a comparação entre as línguas. Apesar disso, em se tratando de diferenças entre plosivas em uma mesma língua, Cho e Ladefoged (1999) afirmam que não existem línguas que apresentam mais de três das categorias apresentadas na Tabela 4.

A partir do momento em que há comparação entre línguas ou se deseja detalhar uma língua com base em teorias fonéticas, não se pode deixar de lado a grande variabilidade que pode existir dentro do *continuum* de fala. A categorização mostrada na Tabela 4 se faz importante justamente para mostrar que, mesmo dentro dessa variabilidade, existem faixas de valores de VOT que vão englobar cada tipo de plosiva. Essa categorização é de grande relevância para a distinção de categorias fonéticas ([p t k] *versus* [p<sup>h</sup> t<sup>h</sup> k<sup>h</sup>]) e fonológicas (/b d g/ *versus* /p t k/ dentro de uma mesma língua, como para plosivas em início de palavra no inglês).

Para as plosivas do PB, destacamos alguns estudos que investigaram as questões relacionadas aos contextos que podem influenciar a produção de segmentos plosivos (ALVES, 2011; ALVES et al., 2008; BARBOSA, 1996; CRISTOFOLINI, 2008, 2013; FICKER, 2003; KLEIN, 1999). A maioria deles se concentrou em aferir medidas de VOT e sua relação com contextos como tonicidade, posição na palavra, qualidade vocálica e velocidade de fala.

Klein (1999) analisou plosivas não-vozeadas e vozeadas do PB em logotomas e palavras reais em sílabas do tipo CV, a partir da fala de quatro falantes nativos, dois do sexo masculino e dois do sexo feminino, todos adultos. Os resultados obtidos pela autora indicaram que: (1) as plosivas vozeadas se enquadraram na faixa de pré-sonorização, apresentando valores negativos de VOT, enquanto as não-vozeadas se enquadraram, na maioria dos casos, dentro da faixa de retardo curto; (2) as plosivas velares não-vozeadas foram as que apresentaram valores médios de VOT mais altos em relação aos mesmos valores encontrados

para as bilabiais e alveolares; (3) o valor médio de plosivas velares seguidas de [e] e [o] já se enquadra, segundo a divisão proposta por Cho e Ladefoged (1999), na categoria de plosivas levemente aspiradas; (4) as plosivas bilabiais e alveolares não-vozeadas apresentaram valores médios de VOT muito similares, não havendo diferença significativa entre eles; (5) dentre as plosivas vozeadas, a média encontrada para plosivas velares foi a menor e a média da bilabial foi a maior dentre os três tipos de plosivas.

Ficker (2003) investigou a produção e percepção de fala do PB, a partir de dados coletados com duas mulheres: uma com desenvolvimento típico da fala e outra com deficiência auditiva. Dentre suas análises de produção, a autora mediu a duração de VOT para plosivas vozeadas e não-vozeadas. No caso das não-vozeadas, todos os valores obtidos foram positivos, sendo o VOT das velares maior do que o das demais plosivas. Para as plosivas vozeadas, todos os valores foram negativos, sendo os maiores referentes às bilabiais. Análises estatísticas revelaram que há diferença significativa entre os pares de plosivas homorgânicas ( $p = 0,000$ ). Com relação ao ponto de articulação, diferenças significativas foram encontradas entre as três plosivas não-vozeadas ( $p \leq 0,046$ ) e entre as plosivas vozeadas ( $p \leq 0,002$ ).

Behlau (1986, *apud* FICKER, 2003) observou as mesmas relações para plosivas vozeadas e não-vozeadas do PB, apesar de encontrar valor positivo para a plosiva velar vozeada (0,17 ms).

Cohen (2004) analisou dados de falantes adultos, nativos de PB, também falantes de inglês como língua estrangeira, com o intuito de observar o comportamento do VOT na língua materna e na língua estrangeira. Os resultados para a língua materna mostraram que o VOT das plosivas não-vozeadas tende a aumentar à medida que a oclusão se torna mais posterior, indo ao encontro dos resultados de Behlau (1986, *apud* FICKER, 2003), Klein (1999) e Ficker (2003).

De maneira similar a Cohen (2004), Alves (2011) analisou o comportamento de alguns parâmetros de plosivas não-vozeadas em relação ao VOT, para dados de fala de três informantes adultos, falantes nativos de PB em nível avançado da língua inglesa, do sexo feminino. Em relação ao VOT, as plosivas não-vozeadas apresentaram a ordem  $[p] < [t] < [k]$ , corroborando os estudos anteriores. A autora também encontrou um considerável número de plosivas que apresentaram uma leve aspiração, especialmente para a plosiva velar, a qual apresentou 82% dos dados inseridos nessa categoria.

Cristofolini (2013), ao analisar dados de fala adulta, encontrou diferenças significativas entre os valores médios das plosivas não-

vozeadas ( $p \leq 0,001$ ). Nesse estudo, porém, a sequência de média de VOT  $[p] < [t] < [k]$  não foi encontrada, uma vez que o valor médio de  $[t]$  foi menor do que o de  $[p]$ . No caso das vozeadas, os valores médios de VOT seguiram a ordem  $[b] > [d] > [g]$ , resultado também encontrado para os demais estudos que investigaram essas plosivas no PB. A pesquisadora também observou que  $[k]$  foi a plosiva que apresentou maior incidência de dados com aspiração.

A Tabela 5 mostra os resultados de médias de VOT para vários estudos no PB.

Tabela 5 – Valores médios de VOT (ms) para as plosivas do PB, a partir dos resultados encontrados em diversos estudos

	[p]	[p <sup>h</sup> ]	[b]	[t]	[t <sup>h</sup> ]	[d]	[k]	[k <sup>h</sup> ]	[g]
Behlau (1986)	10	x	-36	15	x	-33	43	x	0,2
Klein (1999)	16	x	-92	17	x	-93	36	x	-78
Ficker (2003)	11	x	-90	19	x	-77	31	x	-66
Cohen (2004)	22	x	x	27	x	X	38	x	X
Alves (2011) <sup>22</sup>	19	43	x	24	44	X	27	57	X
Cristofolini (2013)	24	x	-110	23	x	-104	46	x	-100

É interessante notar também que os estudos de Behlau (1986), Klein (1999), Cohen (2004), Alves (2011) e Cristofolini (2013) apresentaram plosivas bilabiais e alveolares na categoria de não-aspiradas ( $VOT < 35$  ms) e velares dentro da categoria de levemente aspiradas ( $35 \text{ ms} < VOT < 55$  ms).

## 2.2.6 Transição formântica

Os formantes são as frequências de ressonância do trato oral, gerados a partir da modulação do fluxo de ar pelos variados tubos formados dentro da cavidade oral (FANT, 1960; JOHNSON, 2003).

<sup>22</sup> Alves (2011) apresentou médias diferenciadas para plosivas não-vozeadas não-aspiradas e levemente aspiradas.

Vários estudos na literatura apontam para a importância de informações contidas na zona de transição formântica entre a plosiva e a vogal que a segue. Algumas dessas pesquisas serão detalhadas a seguir.

Na visão de Cooper et al. (1952), as transições formânticas são

(...) rapid shifts in the frequency positions of the vowel formants where vowels and consonant join and are typically most marked for the second formant, although they are usually present in some degree for the other formants as well. (COOPER *et al.*, 1952, p.599)<sup>23</sup>

Em termos articulatórios, é na região de transição que ocorre a passagem do movimento de determinado segmento para o próximo. Como essa passagem é feita de maneira rápida, dentro de 20-40 milissegundos, as mudanças que ocorrem na configuração do trato entre um segmento e outro também tem curta duração, mas, mesmo assim, são capazes de causar certas alterações na configuração de ambos os segmentos (LIEBERMAN; BLUMSTEIN, 1988). Podemos perceber isso, por exemplo, no caso da consoante plosiva [k], que quando produzida diante de [i], se torna mais anterior. Isso se deve ao fato de que o dorso da língua está numa região posterior do trato para a produção da plosiva e precisa rapidamente passar dessa posição para uma região anterior (para a produção de [i]). Assim, o fechamento de [k], que é bem posterior quando esse segmento está em frente a uma vogal posterior como [u], torna-se mais anterior para facilitar a produção de [ki]. As informações contidas na zona de transição vão, dessa forma, ser de grande importância para se determinar qual segmento está sendo produzido diante de uma determinada vogal.

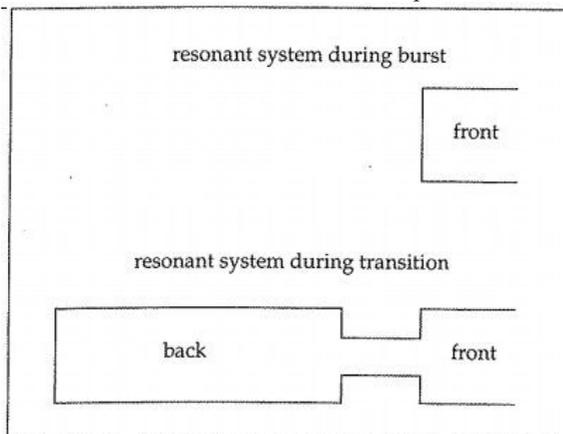
As modificações articulatórias dentro do trato oral irão ocasionar mudanças nos tubos ressoadores que modularão o sinal de fala. De acordo com Johnson (2003), no momento da oclusão, os tubos ressoadores da cavidade na frente do fechamento são os que moldam o ruído no momento da explosão. Assim, o *burst* sofrerá influência somente das cavidades anteriores ao fechamento. Na medida em que a abertura da cavidade ocorre com a soltura da oclusão, as cavidades anterior e posterior ao fechamento vão se “acoplando” novamente, e

---

<sup>23</sup> “(...) transições rápidas nas faixas de frequências dos formantes das vogais onde as consoantes e as vogais se encontram e são mais tipicamente marcadas pelo segundo formante, apesar de que também podem estar presentes para os outros formantes.” (COOPER *et al.*, 1952, p.599). Minha tradução.

ambas passam a influenciar os movimentos de formantes na fase de transição. A Figura 13 ilustra a configuração dos tubos ressoadores durante o momento de fechamento e de soltura da plosiva.

Figura 13 – Configuração dos tubos ressoadores no trato oral durante a fase de fechamento e soltura da oclusão de uma plosiva



Fonte: Johnson, 2003, p. 142

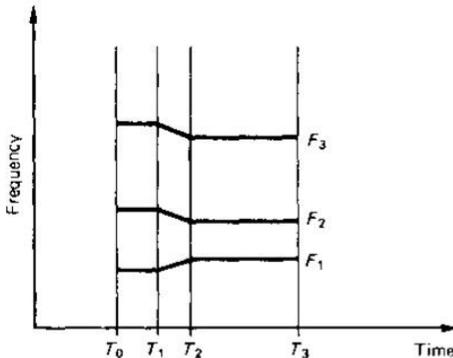
Na zona de transição, localizam-se as frequências de transição, que tendem também a apontar para uma determinada região para cada tipo de plosiva, região que é conhecida como *locus* consonantal. Johnson (2003) explica por que o *locus* consonantal não é um ponto realmente aparente em espectrogramas, mas apenas uma região para onde apontam os formantes:

The fact that the locus is not actually reached is due to the changes in the acoustic coupling of the front and back cavities during stop production. (...). At the instant of stop release, the front and back cavities are uncoupled, so the resonant frequencies of the back cavity do not have an impact on speech spectrum; but as the closure is opened further, the front and back cavities become coupled. This explains the fact that the release burst of [d] has higher-frequency resonances than does the release burst of [g], while [g] has a

higher F2 locus than [d]. (JOHNSON, 2003, p.143).<sup>24</sup>

Alguns teóricos afirmam que as bilabiais apresentarão concentração de energia de F2 nas baixas-frequências, em torno de 700 Hz (COOPER *et al.*, 1952; LIBERMAN, 1974; JOHNSON, 2003). Outros afirmam que a transição de F2 de bilabiais pode ser um pouco mais elevada, em torno de 1000 Hz (LEHISTE; PETERNSON, 1961; KEWLEY-PORT, 1982; FICKER, 2003). No entanto, todos eles parecem concordar que as alveolares, apesar de apresentarem *burst*<sup>25</sup> mais concentrado nas altas-frequências, com a influência das ressonâncias da cavidade posterior ao fechamento, irão apresentar considerável atenuação no *locus* consonantal (em torno de 1800 Hz), se comparadas à mesma região das velares (que varia o *locus* dependendo do tipo de vogal que a segue). A Figura 14 demonstra, esquematicamente, como se configura uma sílaba CV contendo uma plosiva e uma vogal.

Figura 14 – Configuração de uma sílaba CV, com as transições.



Fonte: Clark e Yallop, 1995, p. 283.

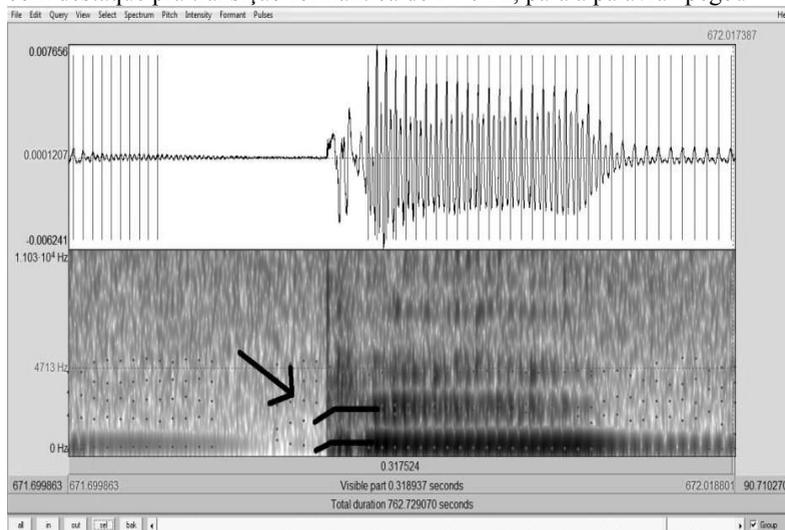
<sup>24</sup> “O fato de que o ponto de *locus* não é atingido se deve às mudanças no acoplamento acústico das cavidades que estão na frente e atrás do bloqueio durante a produção de uma plosiva. (...) No instante da soltura da oclusão, essas duas cavidades estão desacopladas, fazendo com que as ressonâncias da cavidade posterior não influenciem o espectro de fala. Mas, a medida que acontece a soltura da oclusão, as cavidades anterior e posterior se acoplam. Isso explica o fato de o *burst* em [d] apresentar ressonâncias em frequências mais altas do que em [g], enquanto que [g] apresenta um *locus* de F2 mais alto do que [d].” (JOHNSON, 2003, p.143). Minha tradução.

<sup>25</sup> Como visto anteriormente, o *burst* é influenciado apenas pelos tubos ressoadores da cavidade anterior ao fechamento da plosiva (HALLE; HUGHES; RADLEY, 1957; JOHNSON, 2003).

No esquema de Clark e Yallop (1995),  $T_0$ - $T_1$  representa a zona de explosão cujas frequências são influenciadas, segundo os autores, pelo tipo de constrição e fonte de excitação usados na produção da consoante. A região  $T_1$ - $T_2$  corresponde à zona de transição, onde os formantes começam a se modificar a partir da soltura da consoante até a vogal que segue a plosiva.  $T_2$ - $T_3$  representa a região onde os formantes da vogal são estáveis.

A Figura 15 ilustra a configuração de um espectrograma, com ênfase para a região de transição e para as alterações nos formantes de uma vogal [e] antecedida por uma plosiva [p].

Figura 15 – Zona de transição da consoante do PB [p] seguida pela vogal [e], com destaque pra transição formântica de F1 e F2, para a palavra “pegou”



Fonte: Praat/Nossos dados

Na Figura 15, pode-se observar que a transição, tanto de F1 quanto de F2, é ascendente, pois a plosiva [p] apresenta F1 e F2 nas baixas frequências, enquanto a vogal [e] apresenta F1 em torno de 500 Hz e F2 em torno de 1900 Hz (SEARA, 2000).

### 2.2.6.1 Transição de F1

Começaremos nossa investigação sobre a transição formântica com o primeiro formante (F1). De acordo com Cooper *et al.* (1952), o primeiro formante (F1) da transição tende a ser uma pista importante na

percepção do vozeamento das duas categorias de plosivas (isto é, vozeada ou não vozeada), enquanto o segundo formante (F2) auxiliará na distinção entre os pontos de articulação (bilabial, alveolar e velar).

Várias são as pesquisas que afirmam que as transições de F1 do tipo ascendentes correspondem a dados de plosivas vozeadas e que transições fracas ou ausentes são pistas para plosivas não-vozeadas (COOPER *et al.*, 1952; DELATTRE; LIBERMAN; COOPER, 1955; LIBERMAN; DELATTRE; COOPER, 1958; STEVENS; KLATT, 1974; FICKER, 2003). Segundo esses autores, as frequências de *onset* F1 são mais baixas para as plosivas vozeadas do que para as não-vozeadas, resultado assim, em curvas de transição mais planas ou “fracas” para plosivas não-vozeadas.

Delattre *et al.* (1955) buscaram encontrar um *locus* de F1 que identificasse as plosivas. Eles afirmam que todas as plosivas vozeadas são identificadas com maior clareza quando o F1 da transição está na posição mais baixa possível (entre 0-240 Hz), mas que alterações na frequência de F1 não provocaram alterações na identificação dessas consoantes. Os pesquisadores concluíram, então, que transição de F1 não está relacionada com ponto de articulação, mas sim ao vozeamento, corroborando os achados de Cooper *et al.* (1952).

Outra pista importante na distinção de plosivas também passou a ser estudada, especialmente na década de 50, por pesquisadores do *Haskins Laboratories*: o *cutback* de F1, ou seja, o atraso na excitação dos formantes da região de F1 na região de transição. Esse atraso ocorre porque, com um VOT mais longo e conseqüente atraso no início do vozeamento, o F1 em plosivas não-vozeadas só apresentará energia na parte mais final da transição e não na parte inicial, como para as plosivas vozeadas (fato posteriormente observado por Benkí, 2001). Assim, os valores de F1 na região de transição das não-vozeadas tendem a ser mais altos do que para as plosivas vozeadas (PEREIRA; MADUREIRA, 2011).

Para investigar essa pista, Liberman *et al.* (1958) conduziram três experimentos. No primeiro, os estímulos de [b], [d] e [g] seguidos pelas vogais [e], [æ], [ɔ] foram gerados sinteticamente. O intuito desse experimento foi excluir, de maneira gradual, a parte inicial da transição de F1 de 10 em 10 ms, cortando trechos de 10 até 50ms, além da barra de vozeamento. Os estímulos gerados foram julgados por 28 ouvintes nativos de inglês americanos. De acordo com os resultados, à medida que aumentava a duração do corte (o *cutback* de 0 até 50ms), mais diminuía a identificação da plosiva como vozeada. A transição da

plosiva [d] para [t] foi a que requereu maior valor de *cutback* e a transição de [b] para [p] a que requereu menor valor. Outro dado interessante encontrado foi que, para esse experimento, o tipo de vogal não gerou diferenças na identificação dos estímulos, a não ser nas plosivas alveolares, para as quais o valor de *cutback* necessário para a transição ocorreu na sequência [e, æ, ε].

Através do segundo experimento, os pesquisadores avaliaram a influência do *delay* (atraso) no tempo de início de F1 e o grau de ascendência de F1 na distinção entre [d] e [t] diante da vogal [o]. Através de testes perceptuais com 27 informantes, constataram que ascendência na transição de F1 é uma pista para a identificação de plosivas vozeadas e que atraso no início da transição do primeiro formante é indicativo de plosiva não-vozeada.

Outros autores também observaram a influência do *cutback* de F1 na identificação do vozeamento de plosivas. Stevens e Klatt (1974) fizeram um estudo perceptual com estímulos sintéticos para observar a influência da transição formântica e *cutback* de F1 na identificação da distinção não-vozeado/vozeado em plosivas inseridas em sequências CV. Da mesma forma que os autores citados anteriormente, Stevens e Klatt (1974) afirmam que a transição de F1 influencia nessa distinção no sentido de que a região de transição é muito bem definida para as plosivas vozeadas, enquanto é praticamente inexistente para as não-vozeadas. A partir de seus resultados, os pesquisadores puderam observar que o ouvinte usa algumas pistas acústicas em tal distinção. A primeira delas seria o VOT, sendo que consoantes que tivessem duração de VOT maior do que 20ms eram percebidas como sendo não-vozeadas, fato que reitera estudos anteriores. A segunda pista é a rápida mudança de espectro causada pelas transições formânticas. Se o estímulo apresentasse essas mudanças antes do início do vozeamento, o estímulo era percebido também como não-vozeado. Porém, se essas mudanças ocorressem depois do início do vozeamento, os estímulos passavam a ser percebidos como vozeados. Houve casos de estímulos que apresentavam um VOT maior do que 20ms e mudanças de espectro depois do vozeamento e os ouvintes apresentaram julgamentos conflitantes, mostrando que as duas pistas são realmente fatores que interferem no julgamento do vozeamento.

Lisker (1975) conduziu um experimento com diferentes padrões de transição de F1 sintetizados em sílabas CV e conduziu testes perceptuais para observar a influência desse parâmetro na identificação de [k] *versus* [g] para a língua inglesa. Os resultados mostraram que o

parâmetro que melhor identificou [g] foi aquele que apresentou a mais baixa frequência de F1, mesmo sem a transição ascendente, característica apontada como fundamental por outros estudos, como vimos acima. Mesmo quando esse padrão foi contrastado com outro que apresentava a transição de F1 ascendente, com valores de frequência um pouco mais altos, esse resultado se manteve. Segundo o autor, esse resultado demonstra que “(...) it is not so much F1 frequency shift as simply F1 onset frequency that favors /g/” (LISKER, 1975, p. 1549)<sup>26</sup>.

Benkí (2001) avaliou a influência da transição na identificação de vozeamento nas plosivas no inglês. Ele sintetizou dados de plosivas do inglês com variações na transição de F1 de quatro tipos: frequência alta (450Hz) + duração curta (30ms); frequência alta (450Hz) + duração longa (60ms); frequência baixa (180Hz) + duração curta (30ms); frequência baixa (180Hz) + duração longa (60ms). Testes de percepção foram aplicados com quatro falantes nativos de inglês americano. De acordo com os resultados, a combinação de frequência e duração de F1 que mais identificaram plosivas não-vozeadas foi a alta/curta e a baixa/longa foi a que menos apresentou julgamentos para esse tipo de plosiva. Desse modo, os resultados de Benkí (2001) corroboram os achados dos estudos citados acima, pois demonstram que a transição de F1 que melhor aponta para plosivas não-vozeadas é aquela que apresenta redução no aspecto temporal e em frequências mais altas.

No PB, Ficker (2003) avaliou esse parâmetro através do *onset* de F1 e constatou que as plosivas não-vozeadas do PB apresentam valores de frequência mais altos do que as vozeadas, seguindo a tendência encontrada para outras línguas. Resultados de testes estatísticos mostraram diferenças significativas entre os pares homorgânicos ( $p = 0,000$ ).

Em resumo, podemos concluir que a transição de F1 é, de forma geral, um parâmetro que está presente na transição entre uma plosiva e a vogal que a segue. A transição de F1 de uma consoante vozeada para uma vogal é, de maneira geral, bem formada e com ascendência (pois inicia em regiões de baixa frequência), enquanto para as plosivas não-vozeadas a transição de F1 inicia-se, em geral, na parte final do segmento e apresenta uma ascendência menos elevada, pois apresenta valores de frequência mais altos. Esse parâmetro não parece ser dependente do contexto vocálico adjacente, uma vez que a transição de F1 é ascendente, tanto para plosiva vozeada quanto para a não-vozeada,

---

<sup>26</sup> “(...) não é a mudança na frequência de F1, mas simplesmente o início da frequência de F1 que favorece /g/.” (LISKER, 1974, p. 1549). Minha tradução.

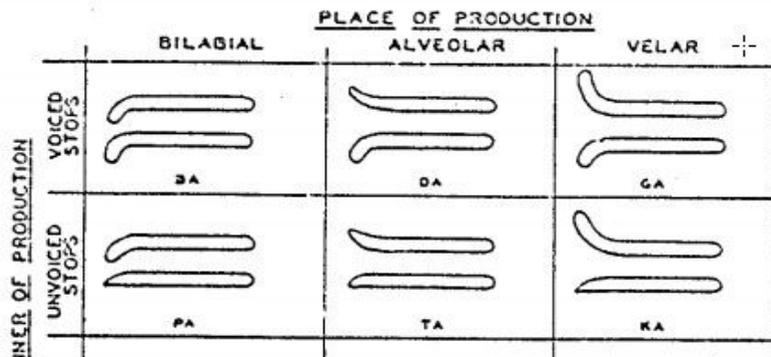
independentemente da qualidade da vogal seguinte, como se pode observar pelas transições na Figura 16 e na Figura 17 (YAVAS, 2011).

### 2.2.6.2 Transição de F2

Outro parâmetro relacionado à região de transição da plosiva é a frequência de F2. Vários estudos apontam para a influência da transição de F2 da consoante para a vogal na distinção entre diferentes pontos de articulação (CLARK; YALLOP, 1995; COOPER et al., 1952; LIBERMAN; DELATTRE; COOPER, 1952; LIBERMAN et al., 1954).

No estudo de Cooper *et al.* (1952), a transição de F2 foi um dos parâmetros sob investigação. Para isso, eles sintetizaram sílabas CV do inglês americano, gerando 11 transições diferentes para cada uma das sete vogais, num total de 77 estímulos, que foram julgados em testes de percepção. Em geral, as plosivas bilabiais foram identificadas em transições ascendentes, enquanto as alveolares e velares em transições descendentes. A Figura 16 apresenta um gráfico com as transições para plosivas vozeadas e não-vozeadas, seguidas da vogal [a], conforme os resultados desses pesquisadores.

Figura 16 – Transições de F1 e F2 das plosivas do inglês seguidas da vogal [a], resultado da investigação de Cooper *et al.* (1952)



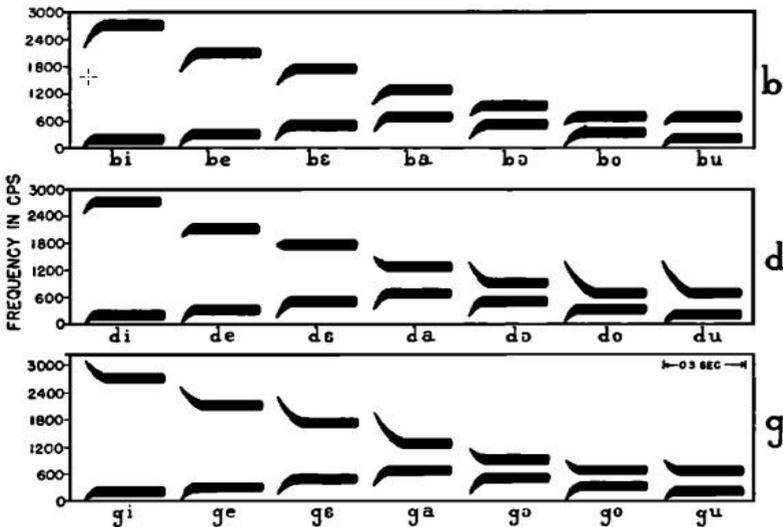
Fonte: Adaptada de Cooper *et al.*, 1954, p.604.

Os pesquisadores também mencionam que as transições apresentam diferentes configurações, a depender da qualidade da vogal que segue a plosiva, resultado esse que foi corroborado por outros estudos (LIBERMAN; DELATTRE; COOPER, 1952; LIBERMAN *et al.*, 1954).

Dois estudos realizados por Liberman *et al.* (1952; 1954) repetiram os achados de Cooper *et al.* (1952) no que concerne as plosivas bilabiais do inglês, que em todos esses estudos, apresentaram transição ascendente em direção à região estável da vogal. Com relação às outras plosivas, os estudos não se mostraram conclusivos no que tange o *locus* consonantal.

Em mais um estudo que buscava observar a influência das transições de plosiva para vogal, Delattre, Liberman e Cooper (1955) produziram estímulos com transições variáveis para cada tipo de vogal, de acordo com a Figura 17.

Figura 17 – Transições formânticas entre plosivas vozeadas e vogais no estudo de Delattre *et al.* (1955)



Fonte: Delattre *et al.*, 1955, p. 770

Por meio de alguns experimentos, os pesquisadores observaram que o início das transições de F2 não são pontos no espectro, mas sim regiões para as quais elas apontam. Se observarmos, por exemplo, as transições de [d] na Figura 17, iremos notar que elas tendem a apontar para uma região em torno dos 1800 Hz. No caso de [g], no entanto, os padrões de transição apontam para uma região de alta-frequência quando a plosiva é seguida por vogais anteriores e para uma região um pouco mais baixa quando seguida por vogais posteriores.

Fant (1973) analisou as transições de F2 durante os 10-30ms depois da soltura da oclusão que, segundo ele, é a região que pode conter informação sobre o ponto de articulação. As plosivas foram seguidas por diferentes vogais da língua sueca. O pesquisador observou que as alveolares apresentam F2 mais alto do que bilabiais, com valores em torno de 1800 Hz. No caso das velares, elas apresentaram valores divergentes com relação às vogais anteriores e posteriores, sendo mais altos quando eram seguidos pelas primeiras.

Posteriormente, outros autores continuaram a defender a visão de que as pistas acústicas como *burst* e transição formântica de F2 são influenciados pelo contexto vocálico e são pistas importantes na identificação dos segmentos plosivos (JOHNSON, 2003; ZAHORIAN; NOSSAIR; COLEMAN, 1987), mas também outros mostraram que, para seus dados, a qualidade vocálica não foi importante na distinção de plosivas (BLUMSTEIN; STEVENS, 1979, 1980; KEWLEY-PORT, 1982).

Para o PB, Berti e Marino (2011) observaram dados de fala de duas crianças, uma com aquisição típica e outra com desvio fonológico, a partir da produção de [t] e [k], seguidos de [a] e [u]. As pesquisadoras constataram que as transições são diferentes para cada plosiva, nos diferentes contextos vocálicos, apesar de não encontrarem diferenças significativas em alguns deles. Nesse estudo, os valores de transição de F2 de [t] se concentraram em torno de 1900 Hz, independentemente do tipo de vogal que a seguiu, enquanto o valor de F2 de [k] foi oscilante, dependendo do contexto vocálico: [ka] = 1769 Hz e [ku]=1653.

Ficker (2003) também observou que houve diferença nas transições das plosivas do PB. Os resultados mostraram transições similares para os pares homorgânicos, sendo ascendente para as bilabiais, descendente para as alveolares e bastante descendente para as velares. As diferenças se mostraram significativas para os três pontos de articulação ( $p= 0,000$  para as não-vozeadas e  $p< 0,002$  para as vozeadas). Com relação às frequências da região de início da transição de F2 da plosiva para a vogal tônica [a], a pesquisadora encontrou de 1285 Hz para [p] e 1370 Hz para [b], de 1530 Hz para [t] e 1575 Hz para [d] e 1725 Hz para [k] e 1824 Hz para [g]. Testes estatísticos apontaram para diferenças significativas entre os três pontos ( $p = 0,000$ ). Os resultados desse estudo estão de acordo com o encontrado para outras línguas.

### 2.2.7 Duração e tipologia de vogais adjacentes

Um dos contextos apontados pela literatura como os que podem influenciar a produção das plosivas é a duração das vogais adjacentes.

Com relação à duração da vogal que antecede a plosiva, Klatt (1976) afirma que

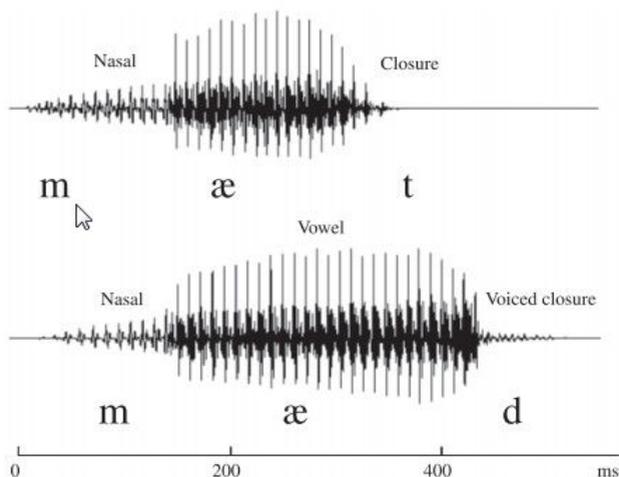
There is a tendency in many languages for vowels to be slightly shorter if followed by a voiceless consonant than if followed by a voiced consonant (Delattre, 1962) (...). The universal influence of the voicing feature of a postvocalic consonant on the duration of visible acoustic energy in the preceding vowel is probably a result of the natural tendency to make a slightly early glottal opening gesture for a postvocalic voiceless consonant in order to insure that no low-frequency voicing cue is generated during the obstruent. (KLATT, 1976, p.1214)<sup>27</sup>

Outros estudos também mostraram que a duração da vogal antecedente independe da qualidade vocálica e é mais curta quando esta precede uma consoante não-vozeada (KLATT, 1973; LADEFOGED; JOHNSON, 2011; LADEFOGED; MADDIESON, 1996; LISKER, 1957; STEVENS; KLATT, 1974). Na Figura 18, pode-se observar a duração da vogal [æ] do inglês no par mínimo [mæ̃t] e [mæ̃d], cuja duração é maior quando a vogal é seguida pelo segmento vozeado.

---

<sup>27</sup> “Existe uma tendência em muitas línguas de vogais serem levemente mais curtas quando são seguidas por consoantes não-vozeadas do que por consoantes vozeadas (Delattre, 1962) (...). A influência universal do vozeamento de uma consoante pós-vocálica na duração da energia acústica da vogal anterior é, provavelmente, resultado da tendência natural de realização de um pequeno e antecipado gesto de abertura glotal na produção da consoante pós-vocálica não-vozeada, para garantir que nenhum resquício de vozeamento nas regiões de baixa frequência seja gerado durante a sua produção.” (KLATT, 1976, p.1214). Minha tradução.

Figura 18 – Duração da vogal [æ] do inglês no par mínimo [mæt] e [mæd]



Fonte: Ladefoged e Johnson, 2011, p.60

As formas de onda acima representam palavras pronunciadas isoladamente. A duração da vogal em [mæd] (louco) é, segundo Ladefoged e Johnson (2011), quase o dobro se comparada à mesma vogal em [mæt] (tapete). Apesar disso, os autores comentam que a duração das duas palavras tende a ser a mesma, uma vez que a diferença na duração da vogal de cada palavra é compensada pela diferença na duração da consoante. Assim, para a língua inglesa, consoantes plosivas não-vozeadas em final de palavra tendem a ser mais longas do que a sua contraparte vozeada na mesma posição (ABRAMSON, 1977; DENES, 1955).

Denes (1955) observa que, na produção do par (*the*) *use* – o uso – e (*to*) *use* – usar – que se diferenciam, respectivamente, por [s] e [z] na língua inglesa, a duração da vogal que antecede a fricativa foi maior para a vogal diante da consoante vozeada e a duração da consoante fricativa foi maior para a não-vozeada. Isso ratifica o “efeito” de compensação apontado por Ladefoged e Johnson (2011).

Lisker (1974) apresenta quatro razões para a diferença na duração da vogal com relação ao contexto consonantal seguinte:

- atribui essa diferença a uma regra de “expansão constante de energia para a sílaba”, que indica que vogais longas e

consoantes plosivas não-vozeadas demandam mais energia articulatória e, por isso, não seriam produzidas juntas.

- afirma que vogais tendem a ser alongadas diante de consoantes vozeadas devido ao ajuste de laringe necessário à manutenção do vozeamento ao longo da produção da sequência VC<sub>vozeada</sub>.
- salienta que as vogais tendem a ser encurtadas diante de consoantes plosivas não-vozeadas porque essas consoantes são consideradas como “fortis” (*versus lenis*), e assim requerem uma antecipação da oclusão articulatória.
- afirma ainda que vogais tendem a ser mais curtas diante de não-vozeadas porque o gesto de oclusão para plosivas “fortis” tende a ser completado mais rapidamente, fazendo com que a vogal tenha duração menor.

Barroco *et al.* (2007) observaram que, em posição medial, as vogais que precedem as plosivas do PE não apresentam diferenças em duração. Porém, quando essas consoantes estão em posição final, a duração da vogal que precede à plosiva não-vozeada é menor do que para a mesma vogal precedendo uma plosiva vozeada. Com relação à vogal que segue a plosiva, os pesquisadores notaram que, quando a plosiva está em posição inicial, a duração da vogal seguinte é semelhante para vozeadas e não-vozeadas, com valores entre 50-200 ms. Em posição medial, a duração da vogal que segue a plosiva também é semelhante para os dois tipos de plosivas, com valores entre 0-100 ms, com exceção de [b].

Em contraposição a esses resultados, Lousada, Martins e Jesus (2006) também para o PB verificaram que, tanto para a vogal precedente quanto para a vogal seguinte à plosiva, os valores de duração das vogais tenderam a ser sempre menores em contextos de plosivas não-vozeadas, seguindo o padrão encontrado em estudos anteriores, para outras línguas.

Com relação ao PB, Brito (2000, *apud* FICKER, 2003) avaliou a fala de crianças que apresentavam aquisição típica e observou que a duração da vogal que precede as plosivas vozeadas foi cerca de 22% maior do que para as não-vozeadas. As primeiras apresentaram duração média de 119 ms, enquanto as últimas apresentaram 97 ms de média.

Ficker (2003) encontrou resultados semelhantes para os dados de fala da informante que apresentava fala típica, no que concerne a vogal que antecede a plosiva (contexto VC). Em valores absolutos, a pesquisadora observou que a duração média da vogal [a] foi de 90 ms

antes de [t] e 92 ms antes de [d] e, para as velares, de 95 ms antes de [k] e 98 ms antes de [g]. Para as bilabiais, no entanto, a autora encontrou padrão inverso, pois a média de duração da vogal [a] foi de 89 ms diante de [p] e de 85 ms diante de [b]. No entanto, não encontramos informações referentes à relevância estatística da diferença entre a duração da vogal nos diferentes contextos (se diante de vozeada ou não-vozeada).

Britto (2010) analisou a produção de plosivas por sujeitos com e sem desvio fonológico. Dentre outras análises, a pesquisadora observou a diferença na duração de vogais orais que antecedem plosivas vozeadas e não-vozeadas. Os resultados mostraram que, para todas as sete vogais orais do PB, a duração foi maior quando esses segmentos se encontravam diante de plosivas vozeadas, para ambos os grupos de sujeitos analisados.

A duração da vogal que segue a plosiva também pode ser influenciada pela consoante anterior. Fant (1973) encontrou um valor médio de duração da vogal em sílabas CvozV que excedia em 30ms a mesma duração para a vogal em sílabas CnvozV.

Os resultados de Ficker (2003) para a duração da vogal [a] tônica que segue as plosivas (contexto CV) mostraram-se menores quando esta segue uma consoante não-vozeada, conforme podemos constatar pelos valores encontrados: [pa] =135 ms e [ba] =142 ms, [ta] =133 ms e [da] =150 ms, [ka] =131 ms e [ga] =156 ms, os quais apresentam diferença significativa entre si ( $p < 0,001$ ). Esse resultado segue uma mesma tendência de outras línguas (FANT, 1973; LADEFOGED, 2001).

Com relação à qualidade vocálica, alguns estudos mostraram que a duração do VOT das plosivas também pode variar dependendo da altura de vogal que as segue.

Port e Rotunno (1979) analisaram plosivas aspiradas do inglês diante de vogais, a partir de dados de uma lista de palavras produzidas por falantes nativos. Eles puderam observar que o VOT é maior para plosivas seguidas de vogais altas [i] e [u] do que para as mesmas consoantes seguidas de [a]. Testes estatísticos mostraram que a diferença encontrada foi significativa ( $p < 0,001$ ).

No PB, Klein (1999) observou o comportamento do VOT de plosivas vozeadas e não-vozeadas diante de vogais [e], [o] e [a]. Os resultados para as palavras inseridas em frases-veículo apontaram que, para as bilabiais vozeadas, houve diferenças significativas entre o VOT de [bo] x [ba] ( $p = 0,011$ ) e diante de [be] x [bo] ( $p = 0,003$ ), tendo como valores médios de VOT -106,88 ms para [ba], -108,80 ms para

[be] e -99,25 ms para [bo]. Com relação à plosiva alveolar vozeada, houve diferenças significativas entre [da] e [de] ( $p = 0,001$ ) e entre [da] e [do] ( $p = 0,000$ ) e os valores médios de VOT foram maiores para as plosivas seguidas por vogais médias. Os resultados para as plosivas velares vozeadas também mostraram diferenças significativas entre os mesmos pares encontrados para as alveolares ([ga] x [ge]  $p = 0,000$ ; [ga] x [go]  $p = 0,000$ ) e os valores médios também foram maiores para as vogais médias. Os valores médios de VOT para plosivas não-vozeadas foram sempre maiores quando se encontravam diante de vogais médias. Para as bilabiais, os valores médios de VOT encontrados foram de [pa]=11,34 ms; [pe]=13,98 ms e [po]=27,77 ms ( $p < 0,001$  para todas as comparações). Em relação às alveolares, os valores médios de VOT foram [ta]=14,03 ms; [te]=18,64 ms e [to]=17,67, com diferença significativa apenas entre a vogal baixa e as demais. Por fim, com relação à velar, os valores de VOT encontrados foram de [ka]=33,84 ms, de [ke]=40,67 ms e de [ko]=39,04 ms. Apenas a comparação entre [ka] e [ke] se mostrou significativa.

Alves (2011) investigou a diferença entre grupos de vogais consideradas altas, médias e baixas e encontrou padrão semelhante ao estudo de Klein (1999) para as plosivas não-vozeadas do PB. Nessa pesquisa, as médias de VOT geralmente apresentaram valores maiores quando diante de vogais altas ( $p < 0,001$ ). Esse estudo não encontrou diferenças significativas na duração de VOT entre o contexto vogal média e baixa.

### 2.2.8 Posição na palavra

A influência do contexto posição na palavra se mostrou bastante importante para as plosivas da língua inglesa em posição inicial, uma vez que, nesse contexto, as vozeadas e não-vozeadas não apresentam vibração de pregas vocais, sendo diferenciadas apenas pela aspiração, que está presente no grupo [p, t, k]. Em contexto intervocálico, a diferença entre as plosivas do inglês passa a ser novamente considerada pela vibração ou não das pregas vocais e, então, as vozeadas passam a ser classificadas dentro da categoria de pré-sonorização (LISKER; ABRAMSON, 1964c).

Klatt (1976) chama a atenção para o fato de que estudos anteriores apontam para a diferença na duração da consoante, dependendo da sua posição em relação à palavra. Ele relata que as consoantes tendem a ser de 10 a 30 ms mais curtas em posição final de palavra se comparadas com as mesmas consoantes em posição inicial.

As consoantes inseridas em meio de palavra também têm tendência de serem mais curtas do que aquelas em posição inicial.

Klein (1999) analisou as plosivas do PB em contexto inicial, medial e final e seus resultados mostraram que, de forma geral, pode haver alterações significativas nos valores de VOT a depender da posição que a plosiva ocupa na palavra. A pesquisadora encontrou valores significativamente maiores de VOT para as plosivas bilabiais em posição medial e em posição final para as alveolares e velares

Alves (2011) investigou o papel da posição na palavra no PB e constatou diferenças significativas apenas para plosivas velares não-aspiradas inseridas em posição inicial e medial ( $p = 0,013$ ), sendo o VOT maior em contexto inicial (29,04 ms) do que em medial (24,93 ms). Para todos os outros contextos, não foram encontradas diferenças significativas, o que pode mostrar que o contexto posição na palavra pode ou não ter influência na produção das plosivas do PB.

### **2.2.9 Tonicidade**

Muitos estudos com o inglês americano apontaram para a influência da tonicidade da sílaba na qual a plosiva está inserida sobre o seu VOT (GIMSON, 1980; KLATT, 1976; LISKER; ABRAMSON, 1967).

Lisker e Abramson (1964c) postularam que as plosivas não-vozeadas do inglês são influenciadas pela tonicidade quando estão inseridas em posição não-inicial, pois elas tendem a ser aspiradas quando estão na sílaba tônica e não-aspiradas em sílabas átonas. Baseados nos seus resultados, eles afirmam também que a média de VOT nas plosivas não-vozeadas em sílabas tônicas tende a apresentar maior duração do que aquelas inseridas em contexto átono, da ordem de: [p] em contextoônico com VOT médio de 35 ms e em átono de 30 ms; [t] em contextoônico com VOT médio de 48 ms e em contexto átono de 40 ms; [k] em contextoônico com VOT médio de 55 ms e em átono de 45 ms. Outro dado importante foi em relação à comparação entre as médias de VOT para as plosivas homorgânicas (dentro da categoria de retardo curto) em contexto átono eônico. Para todas as comparações entre os pares homorgânicos, a diferença entre as médias de VOT foi maior quando as plosivas estavam em contextoônico. Com relação às plosivas [b d g], os autores afirmam que elas não apresentaram diferenças sistemáticas com relação às médias de VOT em relação à tonicidade, tanto quando apresentavam retardo curto quanto pré-sonorização .

Em um estudo posterior, Lisker e Abramson (1967) investigaram mais profundamente o papel de alguns contextos no VOT das plosivas do inglês e chegaram à conclusão de que a tonicidade tende a influenciar esses segmentos, apesar de ser de forma limitada. Os resultados mostraram que, para as plosivas não-vozeadas, o VOT médio tende a ser maior para aquelas que estão inseridas em contexto tônico. No caso das vozeadas que apresentaram retardo curto, houve tendência contrária nas médias de VOT. Já, para as vozeadas com pré-sonorização, a bilabial e a velar apresentaram maiores médias de VOT quando em contexto tônico, enquanto a alveolar apresentou maior média em contexto átono.

Outros autores também comentam sobre a relação entre VOT e tonicidade. Gimson (1980) postula que a aspiração em plosivas inseridas em posição átona tende a ser mais fraca do que em plosivas em posição tônica.

Por outro lado, Klatt (1976), baseando-se em estudos de outros pesquisadores, comenta que a tonicidade lexical tende a influenciar na duração dos segmentos consonantais, no sentido de que, quando eles estão inseridos em contexto pretônico, eles tendem a ser levemente mais longos do que em qualquer outro contexto de tonicidade.

Em relação ao PB, estudos também vêm demonstrando a influência da tonicidade no VOT das plosivas. Klein (1999) observou que, para todas as plosivas vozeadas, a posição tônica foi a que apresentou maiores valores médios de VOT. A comparação da posição tônica com os demais contextos analisados (a saber, pré-tônico e pós-tônico) apresentou diferenças significativas. Com relação às plosivas não-vozeadas, o contexto tonicidade não apresentou resultados semelhantes entre elas. As plosivas bilabiais e alveolares apresentaram maiores valores de média de VOT no contexto pós-tônico. Diferenças significativas foram encontradas na comparação entre o contexto pós-tônico e os demais, para ambas as plosivas ( $p < 0,001$ ). No caso da velar, o contexto que apresentou maior valor médio foi o tônico e os testes estatísticos apontaram para diferenças significativas entre esse contexto e os demais ( $p < 0,001$ ).

Alves *et al.* (2008), através da investigação de dados de fala semi-espontânea, constataram que a tonicidade da sílaba na qual a plosiva não-vozeada estava inserida não apresentou influência sistemática sobre as médias de VOT, no caso das não-aspiradas. Os resultados para as aspiradas mostraram a mesma tendência encontrada por Klein (1999): valores médios maiores para  $[p^h]$  e  $[t^h]$  em contexto pós-tônico e em  $[k^h]$  no contexto tônico.

Os resultados de Alves (2011) com relação às plosivas não-vozeadas não-aspiradas e aspiradas mostrou que a influência da tonicidade não foi sistemática. Assim como Klein (1999), as médias de VOT foram maiores em posição pós-tônica para as plosivas bilabiais e alveolares não-aspiradas, com diferenças significativas para ambas ( $p < 0,05$ ). No caso das velares, a posição que apresentou maiores valores de VOT foi a tônica ( $p < 0,001$ ). Em relação às plosivas aspiradas, a pesquisadora não encontrou diferenças significativas entre nenhum contexto de tonicidade, para nenhum tipo de plosiva.

Apesar dos estudos, na sua maioria, apontarem para a influência da tonicidade sobre o VOT, acreditamos que, assim como apontado por Klatt (1976), a tonicidade também pode influenciar a duração total da consoante plosiva no PB, apesar de não encontrarmos pesquisas avaliando essa questão. Essa questão será investigada nesta tese.

## 2.2.10 Taxa de elocução

A taxa de elocução é um fator relevante na produção dos segmentos. Quando ela é diminuída, durações extras podem ser distribuídas entre alongamentos e pausas, assim como a aceleração da velocidade pode reduzir a duração ou ainda apagar os segmentos consonantais e vocálicos (KLATT, 1976). Desse modo, a análise que leve em consideração aspectos de duração na fala deve atentar para esse parâmetro, pois diferenças em termos de duração entre as produções de falantes podem ser influenciadas por disparidades entre a taxa de elocução intersujeitos.

Changes in speaking rate exert a complex influence on the durational patters of a sentence. When speakers slow down, a good fraction of the extra duration goes into pauses (Goldman-Eisler, 1968). On the other hand, increases in speaking rate are accompanied by phonological and phonetic simplifications as well as differential shortening of vowels and consonants” (KLATT, 1976, p. 1210).<sup>28</sup>

---

<sup>28</sup> “Mudanças na velocidade de fala exercem uma influência complexa nos padrões de duração de uma sentença. Quando a velocidade de fala é diminuída, uma boa parte de duração extra é inserida em pausas (Goldman-Eisler, 1968). Por outro lado, um aumento na velocidade de fala é acompanhado por simplificações fonéticas e fonológicas da mesma forma que uma diminuição diferenciada em vogais e consoantes” (KLATT, 1976, p.1210). Minha tradução.

Em geral, a taxa de elocução tende a ser calculada a partir da medida da taxa de palavras pronunciadas por minuto ou ainda de sílabas por segundo. Klatt (1976) afirma que a velocidade normal de fala pode variar entre 150 a 250 palavras por minuto ou ainda de 4 a 7 sílabas por segundo (essa mesma taxa é considerada normal por Huggins, 1964).

Uma forma de se controlar as diferenças de taxa de elocução é extrair a duração relativa, seja do segmento em relação à palavra na qual está inserido ou mesmo de parte do segmento em relação a ele mesmo (BERTI; MARINO, 2011).

Cristofolini (2013) encontrou as seguintes porcentagens médias da duração da plosiva em relação à palavra, na fala adulta: 26,5% para [p]; 27,5% para [t]; 28,2% para [k]; 24,2% para [b]; 22,8% para [d] e 21,3% para [g]. Esses dados revelam um fato interessante, uma vez que mostram que, percentualmente, plosivas não-vozeadas tendem a ocupar uma porção maior da palavra se comparadas às plosivas vozeadas, respeitando a tendência geral encontrada na literatura. Outro fator observável a partir desses valores é que, no caso das plosivas não-vozeadas, [k] ocupa a maior taxa percentual da palavra, seguido de [t] e [p], corroborando vários estudos sobre o ponto de articulação, como vimos anteriormente. O mesmo acontece com as vozeadas, mas, no sentido inverso, a bilabial apresenta o maior valor percentual, seguida da alveolar e da velar. A pesquisadora também observou a duração relativa do VOT com respeito à palavra e encontra os seguintes valores percentuais médios na fala adulta: 5,8% para [p]; 4,5% para [t]; 10,1% para [k]. Esses dados mostram que o VOT das plosivas velares ocupa uma porção maior das palavras em que estão inseridas.

Em suma, os estudos acima apontam para a existência de parâmetros acústicos utilizados na produção das plosivas do PB e também de outras línguas que dão conta de descrever a forma como elas são produzidas pelos falantes dessas línguas. Não obstante, esses estudos também mostram que esses parâmetros podem ser utilizados para caracterizar os diferentes tipos de segmentos plosivos e que, para alguns casos, podem ser utilizados na sua identificação perceptual. Cada um dos parâmetros acima descritos serão investigados na presente pesquisa.

## 2.3 RESUMO DOS ESTUDOS SOBRE A PRODUÇÃO DE PLOSIVAS

### 2.3.1 As plosivas em geral

Com base nos estudos relatados acima, podemos inferir que:<sup>29</sup>

- A duração do fechamento é maior para as plosivas não-vozeadas do que para as vozeadas. Esse parâmetro é influenciado pela taxa de elocução, sendo que, quanto maior a velocidade de fala, menor é a duração do fechamento.
- Quanto mais posterior o ponto de articulação, maior tendem a ser os valores de VOT das plosivas não-vozeadas. Assim, os valores de VOT se estabelecem na ordem [p] < [t] < [k]. Para as plosivas vozeadas tendem a apresentar padrão inverso: [b]>[d]>[g]
- Quanto maior a massa do articulador envolvido no fechamento da plosiva não-vozeada, maior o VOT. Esse fator contribui para obtenção de valores de VOT maiores em plosivas velares. Esse fator também pode explicar porque as plosivas bilabiais e alveolares não apresentam valores de VOT com diferenças significativas, uma vez que a massa relativa ao lábio inferior (articulador ativo das bilabiais) e à ponta da língua (articulador ativo das alveolares) é praticamente a mesma.
- Presença da barra de vozeamento, isto é, de concentração de energia na região de F0 no espectrograma, indica vibração das pregas durante a produção de um segmento. Contudo, essa pista não é fundamental na caracterização das plosivas. Sempre que ela está presente, a plosiva é avaliada como vozeada, mas quando ela está ausente, a plosiva ainda será caracterizada como vozeada caso apresente duração de fechamento curta ou ainda quando ela não apresentar aspiração.
- Com relação à transição formântica, a região de F1 das plosivas vozeadas tende a ser bem visível no espectro de fala e com ascendência, enquanto que para as não-vozeadas, apresenta pouca elevação na transição para a

---

<sup>29</sup> Para informações detalhadas e referências, ver Capítulo 2.

vogal. Portanto, F1 está relacionado com o vozeamento das plosivas. Já, F2 está relacionado com o ponto de articulação. Assim, a transição de bilábias, alveolares e velares apresentará valores de frequência distintos, seguindo a ordem [bilabial<alveolar<velar].

- Uma vogal precedendo uma plosiva não-vozeada tende a ser mais curta do que a mesma vogal diante de plosivas vozeadas. O mesmo padrão ocorre para as vogais que seguem as plosivas.
- Há divergências sobre a influência da posição da consoante em relação à palavra. Aqueles autores que afirmam que há influência postulam que as consoantes tendem a ser mais longas em posição inicial, especialmente para a língua inglesa.
- Existem diferentes posicionamentos com relação à influência da tonicidade na produção das plosivas. Alguns autores atestam a influência da tonicidade da sílaba na qual está inserida a plosiva como um fator que tende a alterar os valores médios de VOT de plosivas, os quais tendem a ser maiores quando esses segmentos estão inseridos em sílaba tônica. Outros apontam para valores médios maiores em posição pretônica.
- A taxa de elocução altera a duração do VOT das plosivas, pois, ao se aumentar a velocidade de fala, diminui-se a duração do VOT.
- A duração total dos segmentos plosivos tende a ser maior para os segmentos não-vozeados.

Cabe lembrar que todos os parâmetros aqui apontados auxiliam na caracterização de uma ou outra plosiva. Apesar de muitos estudos com fala sintetizada mostrarem que a presença de um único parâmetro pode ser suficiente na identificação das plosivas, na fala natural, eles irão atuar de forma conjunta durante a produção do segmento e dentro de um contínuo de fala (COOPER *et al.*, 1952; LIBERMAN *et al.*, 1954; ALBANO, 2001).

### 2.3.2 As plosivas no PB

Considerando os estudos sobre as plosivas do PB, destacam-se os seguintes pontos:<sup>30</sup>

- a duração do fechamento das plosivas do PB segue tendência apontada na literatura, ou seja, é maior para as plosivas não-vozeadas.
- as medidas das características espectrais de *burst* mostraram que o pico espectral médio das alveolares é maior do que o das velares.
- em relação ao VOT, as plosivas velares são aquelas com valores médios mais elevados, havendo diferenças significativas entre o VOT relativo aos três pontos de articulação. É possível a realização de plosivas levemente aspiradas [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>].
- a transição de F2 das consoantes plosivas do PB parecem apresentar a mesma ordem nos valores de frequência que os estudos para outras línguas.
- a duração, tanto da vogal anterior quanto da posterior é, em geral, maior diante de plosivas vozeadas do que de plosivas não-vozeadas.
- O contexto tonicidade parece não apresentar influência sistemática sobre as plosivas não-vozeadas do PB. Contudo, há tendência de as médias de VOT serem maiores no contexto pós-tônico para as plosivas [p] e [t] e no contextoônico para [k]. Todas as vozeadas em contextoônico apresentam valores médios maiores de VOT se comparadas aos demais contextos de tonicidade.
- a duração relativa da plosiva em relação à palavra na qual ela está inserida é maior nas plosivas não-vozeadas do que nas vozeadas, mostrando que as primeiras tendem a ser mais longas do que as segundas. Dentre as não-vozeadas, a taxa percentual apresentou a seguinte ordem: [k] > [t] > [p] e, para as vozeadas, [b] > [d] > [g].

---

<sup>30</sup> Para informações detalhadas e referências, ver Capítulo 2.

## 2.3 PERGUNTAS E HIPÓTESES DE PESQUISA

Baseados nos relatos dos estudos apresentados anteriormente, nossa pesquisa será guiada pelas seguintes perguntas e hipóteses:

**Pergunta 1:** Há diferença entre a duração das plosivas vozeadas e não-vozeadas?

**Hipótese 1:** Haverá diferença entre as consoantes plosivas vozeadas e não-vozeadas, sendo que as últimas apresentarão valores de duração de consoante maiores do que as primeiras. Essa tendência se manterá na comparação entre os pares de plosivas homorgânicas.

**Pergunta 2.1:** Há influência do ponto de articulação na duração da consoante entre as plosivas não-vozeadas?

**Hipótese 2.1:** Não haverá diferença entre os três pontos de articulação com relação à duração das plosivas não-vozeadas.

**Pergunta 2.2:** Há influência do ponto de articulação na duração da consoante entre as plosivas vozeadas?

**Hipótese 2.2:** O ponto de articulação influenciará a duração das consoantes vozeadas, pois, na medida em que o ponto vai se tornando mais posterior, diminui-se a duração da consoante.

**Pergunta 3:** Há diferenças entre a duração da oclusão de plosivas vozeadas e não-vozeadas?

**Hipótese 3:** A duração da oclusão em plosivas não-vozeadas será maior do que o mesmo momento em plosivas vozeadas, mesmo na comparação entre plosivas homorgânicas.

**Pergunta 4.1:** Há influência do ponto de articulação na duração da oclusão entre as plosivas não-vozeadas?

**Hipótese 4.1:** Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.

**Pergunta 4.2:** Há influência do ponto de articulação na duração da oclusão entre as plosivas vozeadas?

**Hipótese 4.2:** Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.

**Pergunta 5:** Os dados de falantes adultos também apresentarão especificidades acústicas, como as apresentadas em dados de estudos com crianças falantes nativas do PB?

**Hipótese 5:** Os dados de falantes adultos apresentarão especificidades acústicas, sendo o tipo mais comum a presença de *edge vibrations* nas plosivas não-vozeadas e interrupções de vozeamento nas vozeadas.

**Pergunta 6:** Os valores das características espectrais do *burst* diferem entre os pontos de articulação?

**Hipótese 6:** Os valores de características espectrais de *burst* se diferenciarão com relação ao ponto de articulação.

**Pergunta 7:** As medidas de *burst* podem sofrer alteração de acordo com o tipo de vogal que segue a plosiva?

**Hipótese 7:** Existirá diferenças no *burst* das plosivas seguidas por diferentes vogais para o PB, sendo que plosivas seguidas por vogais posteriores podem apresentar maiores valores de frequência do que aquelas seguidas por vogais anteriores.

**Pergunta 8:** Há diferenças entre o VOT de plosivas vozeadas e não-vozeadas?

**Hipótese 8:** O VOT das plosivas não-vozeadas apresentará valores positivos e as plosivas não-vozeadas poderão ser incluídas na categoria de não-aspiradas ou levemente aspiradas, além da categoria de retardo curto. O VOT de plosivas vozeadas do PB apresentará valores negativos, indicando pré-sonorização.

**Pergunta 9.1:** Há diferenças entre os valores de VOT entre as plosivas não-vozeadas?

**Hipótese 9.1:** O valor da plosiva velar não-vozeada se diferenciará das demais plosivas, pois apresentará maiores valores de VOT, enquanto os valores de VOT poderão ser similares para as plosivas bilabiais e alveolares. As plosivas velares serão aquelas com mais chance de apresentar valores de VOT dentro da categoria de leve aspiração.

**Pergunta 9.2:** Há diferenças entre os valores de VOT entre as plosivas vozeadas?

**Hipótese 9.2:** Os valores de VOT para plosivas vozeadas seguirão a sequência [b] > [d] > [g], apresentarão valores negativos (por possuírem pré-sonorização) e diferentes entre si.

**Pergunta 10:** Há diferenças entre o VOT das plosivas que apresentaram aspiração e o VOT das que não apresentaram?

**Hipótese 10:** Existirá diferença significativa entre o VOT das plosivas aspiradas e não-aspiradas.

**Pergunta 11:** Há diferenças em F1 entre os pares de plosivas homorgânicos?

**Hipótese 11:** A região de F1 para plosivas vozeadas apresentará valores mais baixos do que para as não-vozeadas. Espera-se encontrar transições mais acentuadas para as plosivas vozeadas.

**Pergunta 12.1:** Há diferenças em F2 entre as plosivas não-vozeadas?

**Hipótese 12.1:** Se encontrará diferenças significativas entre os valores de transição de F2 para os diferentes pontos de articulação. Ademais, haverá influência da vogal que segue a plosiva na transição de F2, sobretudo no caso das plosivas velares.

**Pergunta 12.2:** Há diferenças em F2 entre as plosivas vozeadas?

**Hipótese 12.2:** Se encontrará diferenças significativas entre os valores de transição de F2 para os diferentes pontos de articulação. Ademais, haverá influência da vogal que segue a plosiva na transição de F2, sobretudo no caso das plosivas velares.

**Pergunta 13:** Há diferença na duração das vogais que antecedem as plosivas não-vozeadas e plosivas vozeadas?

**Hipótese 13:** A duração das vogais que antecedem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.

**Pergunta 14:** Há diferença na duração das vogais que seguem as plosivas não-vozeadas e plosivas vozeadas?

**Hipótese 14:** A duração das vogais que seguem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.

**Pergunta 15:** Existe diferença na duração da consoante e de seu VOT dependendo da altura da vogal que a segue?

**Hipótese 15:** A duração da consoante e de seu VOT serão maiores quando a vogal adjacente é uma vogal alta (em detrimento da vogal baixa).

**Pergunta 16.1:** O contexto posição na palavra exerce influência sobre a duração total da consoante e VOT das plosivas não-vozeadas do PB?

**Hipótese 16.1:** A posição na palavra poderá influenciar a duração total da consoante e VOT de plosivas não-vozeadas.

**Pergunta 16.2:** O contexto posição na palavra exerce influência sobre a duração total da consoante e VOT das plosivas vozeadas do PB?

**Hipótese 16.2:** A posição na palavra poderá influenciar a duração total da consoante e VOT de plosivas vozeadas.

**Pergunta 17.1:** O contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas não-vozeadas do PB?

**Hipótese 17.1:** Se o PB seguir a mesma tendência para outras línguas (KLATT, 1976), as consoantes em posição átona terão maior duração do que em posição tônica. No que tange o VOT, o contexto átono propiciará maiores valores de VOT no caso de bilabiais e alveolares e o tônico propiciará a maior duração do VOT de consoantes velares.

**Pergunta 17.2:** O contexto de tonicidade exerce influência sobre o VOT das plosivas do PB?

**Hipótese 17.2:** Se o PB seguir a mesma tendência para outras línguas (KLATT, 1976), as consoantes em posição átona terão maior duração do que em posição tônica. Além disso, as plosivas vozeadas inseridas em sílabas tônicas apresentarão valores maiores de VOT.

### 3 METODOLOGIA

Para que possamos investigar a influência dos parâmetros na produção das plosivas do PB (relatados na Seção 2.2), iremos conduzir uma pesquisa a partir de dados coletados com falantes nativos dessa língua, que produziram sentenças contendo as palavras com os segmentos-alvo.

As subseções a seguir descrevem, em detalhes, como foi delineada a metodologia empregada para tal fim.

#### 3.1 PARTICIPANTES

Os participantes selecionados são cinco informantes do sexo feminino, falantes nativas do PB. Todas as informantes tinham entre 21 e 29 anos na data da coleta (média de 26 anos) e nasceram e passaram o período de aquisição da linguagem na região de Criciúma, Santa Catarina. Optamos por coletar dados com informantes da mesma região para evitar possíveis diferenças dialetais, um fator que pode influenciar quaisquer resultados que possamos encontrar.

Com relação ao local de coleta, a cidade de Criciúma está localizada ao sul do estado de Santa Catarina e, assim como as outras cidades da mesma região, foi colonizada no início do século XIX principalmente por imigrantes italianos (TRAINOTTI; MARGOTTI, 2008). Ela também recebeu imigrantes oriundos de outras regiões da Europa como Alemanha, Áustria e Polônia, e outros vindos do Oriente Médio mais recentemente (meados do século XX), como Turquia, Síria e Líbia. O seu dialeto, portanto, sofre alguma influência da língua italiana, pois alguns dos habitantes mais antigos são descendentes dos primeiros italianos e ainda falam italiano ou, pelo menos, utilizam algumas poucas palavras em seu linguajar. No caso dos falantes mais novos, como as participantes desta pesquisa, poucos são aqueles que guardam resquícios da língua italiana no seu falar.

Podemos citar como algumas características presentes no falar dessa região<sup>31</sup>:

- (a) uso do pronome “tu” em detrimento de “você”;
- (b) palatização de /t/ e /d/ diante da vogal alta, *e.g.*, leite [ 'lejtʃi] e disco [ 'dʒiskʊ] ou ainda a variação [ 'lejtʃi];

---

<sup>31</sup> Até o momento não encontramos nenhuma pesquisa que faça uma descrição do dialeto da cidade de Criciúma. Essas características são impressões observadas pela pesquisadora, que nasceu e passou toda a fase de aquisição da linguagem nesta mesma região.

- (c) retroflexão facultativa do /r/ em coda silábica;
- (d) quando não retroflexo, o /R/ em coda silábica é produzido como [r].

Todas as participantes do nosso estudo apresentam as características relatadas acima em sua fala. Por isso, não incluímos em nosso *corpus* plosivas alveolares seguidas de vogal alta anterior. A influência de uma língua estrangeira também foi um fator controlado. Todas as informantes tiveram contato com alguma língua estrangeira (inglês ou espanhol) durante o período escolar, mas nenhuma delas declarou que fala fluentemente algum desses idiomas ou que fala qualquer outra segunda língua (L2) ou língua estrangeira (LE)<sup>32</sup>.

A opção por falantes do mesmo sexo ocorreu para que também pudéssemos excluir esse fator extralinguístico da coleta. A escolha por falantes femininos ocorreu também em função da disponibilidade de informantes.

Todas as informantes relataram não ter problemas auditivos ou na produção da fala e nem com relação à leitura, fato que foi atestado durante o preenchimento do questionário biográfico (Apêndice A), conduzido com a presença da investigadora. A participação no teste se deu de forma voluntária e todas assinaram o termo de consentimento livre e esclarecido.

### 3.2 CORPUS

O *corpus* foi formado a partir de uma lista de palavras contendo os segmentos-alvo, para a leitura em frases veículo do tipo “Digo \_\_\_\_\_ baixinho” e “Digo \_\_\_\_\_ pra ela”.

Os contextos vocálicos precedente e seguinte, a tonicidade da sílaba, a posição do segmento em relação ao vocábulo foram controlados para que pudessem ser investigados. A lista completa contendo as palavras utilizadas no experimento está localizada no Apêndice B. O Quadro 1 apresenta alguns exemplos de palavras usadas no experimento. No total, a lista de palavras a ser gravada foi constituída de 260 palavras que continham cada uma um ou mais segmentos plosivos a serem analisados. Nosso *corpus* continha, assim, em sua forma original, 283 segmentos-alvo.

---

<sup>32</sup> Inserimos aqui a distinção entre L2 e LE, pois a primeira é aquela que pode ser falada no lar, entre os membros da família e a LE é aquela aprendida na escola ou em cursos de idioma.

Quadro 1 – Exemplos de vocábulos utilizados no teste de produção, de acordo com os contextos em análise

Plosiva	Sílabas no vocábulo	Posição na palavra	Vogal posterior <sup>33</sup>	Tonicidade	
				Tônica	Átona
[p]	Dissílabo	Medial	[a]	capaz	capa
			[e]	ipê	-
			[ɛ]	tripé	-
			[i]	-	jipe
			[o]	capô	-
			[ɔ]	cipó	-
			[u]	-	topo

Nosso *corpus* se restringiu com relação aos seguintes aspectos:

- As palavras que preencheriam o contexto [\_\_i] foram descartadas para as plosivas [t] e [d], devido à possível palatização de [t] diante de [i] no PB (SEARA; NUNES; LAZZAROTTO-VOLCÃO, 2015);
- O quadrilátero vocálico do PB em posição de final de palavra é reduzido de sete vogais para três vogais orais, a saber [ɐ], [i] e [u]. Isso acontece devido à neutralização dos fonemas /e/ e /o/<sup>34</sup> nessa posição, para alguns dialetos do PB, incluído o dialeto dos nossos participantes (SEARA; NUNES; LAZZAROTTO-VOLCÃO, 2015).
- De acordo com as regras fonotáticas do PB, os segmentos plosivos não podem ocupar coda silábica, com exceção de palavras como “advogado” e “apto”, nas quais pode se observar, geralmente, a inserção de uma vogal epentética, obtendo-se [ˈapitʊ]. Esses

<sup>33</sup> Cabe lembrar que a qualidade vocálica está relacionada, de acordo com a literatura apresentada, somente com a vogal seguinte. Portanto, controlamos qualidade vocálica somente em relação a essa vogal. No que tange à vogal anterior, coletamos informação de sua duração em pares mínimos do tipo “cata” e “cada” para observarmos se houve diferença entre a duração da vogal [a] precedendo consoantes vozeadas e não-vozeadas.

<sup>34</sup> Pode-se dizer que o fonema /a/ também sofre neutralização em final de palavra, uma vez que a qualidade vocálica é alterada de [a] para [ɐ].

casos foram desconsiderados em nosso *corpus* (SEARA; NUNES; LAZZAROTTO-VOLÇÃO, 2015).

- Quando a plosiva está inserida na última sílaba da palavra, a vogal final pode ser ensurdecida ou ainda ter considerável redução, acarretando em um aumento do VOT dessa plosiva por uma possível sobreposição de gestos articulatórios (ALVES; DIAS, 2010; MENESES; 2012). Optamos por excluir esse tipo de dado.
- Encontramos poucos monossílabos que são palavras “reais” da língua, a sua maioria com a plosiva [p]. Para que pudéssemos avaliar esse contexto, precisaríamos incluir no *corpus* logatomas monossilábicos. Assim, resolvemos também desconsiderar esse contexto.
- Vogais [ɛ] e [ɔ], no dialeto investigado, são sempre pronunciadas em contexto tônico. Assim, pronúncias como [tɛˈsɪdʊ] ou [mɔˈrãɡʊ] são incomuns no dialeto dos participantes.
- A sequência -que, -qui, -gue, -gui pode ou não apresentar a pronúncia da semivogal [w], a depender da palavra. Esse fato acabou dificultando o levantamento de dados para alguns contextos [ke] [ki] e [ge] [gi], que acabaram não preenchidos. Os dados que foram pronunciados com [w] foram excluídos.
- Ao desenvolver o *corpus*, inserimos algumas plosivas que antecediam ou seguiam nasais e ditongos. Ao iniciarmos a análise de dados, nos deparamos com possíveis alterações na realização dessas plosivas quando comparadas aos contextos [plosiva]não-nasal ou não-nasal[plosiva] e ainda [plosiva]monotongo ou [monotongo]plosiva. Em virtude dessas alterações, decidimos por excluir esse tipo de dado, uma vez que, em termos estatísticos, eles não seriam comparáveis aos demais (por apresentaram baixa ocorrência). No total, foram excluídos 21 dados.

Para a avaliação da vogal precedente, optamos por avaliar apenas um único contexto vocálico, pois o tipo de vogal parece influenciar as plosivas apenas em contexto posterior. Assim, o contexto vocálico escolhido foi a vogal tônica [a], por ser um dos mais recorrentes no PB. Para a análise da vogal seguinte, que pode influenciar alguns dos

parâmetros avaliados nesta pesquisa, optamos por investigar a vogal final [ɐ] e [ʊ]. O Quadro 2 apresenta o *corpus* elegido para essa análise.<sup>35</sup>

Quadro 2 – Lista de palavras do *corpus* para a análise da duração da vogal que antecede ([a]) e que sucede as plosivas ([ɐ] e [ʊ])

	[p]	[b]	[t]	[d]	[k]	[g]
Pares mínimos	napa	naba	cata	cada	vaca	vaga
	capo	cabo	nata	nada	paca	paga
	rapo	rabo	nato	nado	saca	saga
Pares análogos	capa	baba	data	fada	faca	baga
	tapa	gaba	gato	fado	saco	lago
	sapo	nabo	mato	lado	caco	pago

Outro ponto importante relacionado ao *corpus* foi a opção pelo contexto de fala controlado através de frases-veículo. Acreditamos que a opção por esse tipo de dado nos permite obter a maior parte dos contextos almejados. Se houvéssimos optado pela fala espontânea, não poderíamos garantir que todos esses contextos seriam efetivamente produzidos pelas informantes. Além disso, como relatado por Berti e Marino (2011), o uso de frases-veículo facilita a manutenção da curva entoacional para as palavras que contém o segmento-alvo, evitando-se o que é conhecido entre os pesquisadores como “efeito-lista”. Ademais, Umeda (1977) afirma que o ritmo em frases-veículo tende a ser regular, evitando, assim, qualquer influência prosódica na produção dos segmentos-alvo, o que poderia acarretar em alguma alteração como, por exemplo, a duração das vogais.

Em terceiro lugar, na nossa experiência com fala semi-espontânea (ALVES, 2011), notamos que as tendências encontradas para dados de fala controlados em estudos anteriores (KLEIN, 1999; LISKER; ABRAMSON, 1964a, dentre outros) tendem a ser replicados na fala não-controlada, especialmente no que concerne o VOT (KLOKER, 1975).

Ainda sobre esse aspecto, Klatt (1976) aponta que

Speech material can be obtained under a number of conditions which range from the reading of lists of nonsense syllables to

<sup>35</sup> Devido ao grande volume de dados, todas as análises que observaram contexto vocálico (a saber: *burst*, transição vocálica e qualidade e duração da vogal) foram feitas utilizando-se somente os dados referentes à produção do *corpus* apresentado no Quadro 2.

normal conversational speech. It is by no means certain that rules derived from nonsense syllable studies have anything to say about spontaneous speaking habits, but comparisons between measurements of, e.g., Oller (1973) and Lehiste (1975) on nonsense syllable sequences, Klatt (1975) and Umeda (1975) on a read discourse, and Kloker (1975) on spontaneous speech suggest that the similarities are greater than the differences. (KLATT, 1976, p.1209).<sup>36</sup>

Além disso, a etiquetagem deve ser extremamente cuidadosa e os limites de segmentos ou palavras devem ser encontrados. Essa tarefa é facilitada na fala controlada. Halle, Hughes e Radley (1957) apontam para essa problemática em seu estudo:

Our first problem, how to define a transition, illustrates well the difficulties that are encountered in the study of natural speech, but can easily be avoided when one has control over the stimulus. (HALLE; HUGHES; RADLEY, 1957, p.112).<sup>37</sup>

Todos os aspectos apontados anteriormente são enfrentados pelo pesquisador que deseja trabalhar com fala não-controlada, mas que, de forma alguma, inviabilizam o trabalho.

Desse modo, respaldados pela teoria e pela empiria, decidimos partir da análise de dados de fala controlada do PB, observando um número considerável de parâmetros que podem influenciar as plosivas, objetivando um detalhamento do seu funcionamento na produção desses segmentos no PB.

---

<sup>36</sup> “Dados de fala podem sob várias condições, que vão desde a leitura de listas de sílabas até a fala extraída de uma conversação. Não é certo que as regras que derivam de estudos com sílabas têm algo a dizer sobre a fala espontânea, mas a comparação entre medidas de, por exemplo, sequências de sílabas em Oller (1973) e Lehiste (1975), de leitura em Klatt (1975) e Umeda (1975), e de fala espontânea em Kloker (1975) sugerem que as similaridades são maiores do que as diferenças.” (KLATT, 1996, p.1209). Minha tradução.

<sup>37</sup> “Nosso primeiro problema, de como definir a transição, ilustra bem as dificuldades que são encontradas em estudos com fala natural, mas que podem ser facilmente evitadas quando se tem controle sobre os estímulos.” (HALLE; HUGHES; RADLEY, 1957, p.112). Minha tradução.

### 3.3 PROCEDIMENTOS DE COLETA

Para a coleta de dados de fala de plosivas, foram conduzidas seções individuais de leitura com as participantes. Realizamos uma seção de leitura com 20 minutos de duração em média, com intervalos de 5 minutos, para que os participantes pudessem descansar, evitando o “efeito-lista” que poderia ocorrer, caso mantivéssemos a gravação sem intervalos. A lista de palavras randomizada foi lida duas vezes por cada informante, com 260 vocábulos por informante, totalizando 283 segmentos para análise (algumas palavras continham mais de um segmento a ser analisado). Como desses 283 segmentos-alvo, 21 foram excluídos (como vimos na seção anterior), totalizamos 262 segmentos para análise. Assim, esperávamos obter em nosso banco de dados um total de 2620 segmentos (2 repetições x 5 informantes x 262 segmentos-alvo). No entanto, após a etiquetagem dos dados, nosso banco de dados contabilizou 2227, pois 393 dados foram excluídos devido a fatores como inserção de pausa, dificuldade na visualização de algum elemento da plosiva, dentre outros. A Tabela 6 resume a quantidade de dados válidos encontrados para cada um dos tipos de dados discriminados na Tabela 6.

Tabela 6 – Resumo do banco de dados

<b>Tipo</b>	<b>Núm. de dados esperados</b>	<b>Número de dados válidos obtidos</b>
<i>Corpus</i>	2260	1939 (86%)
Plosivas + vogal nasal (excluídas)	-140	x
Plosivas + ditongo (excluídas)	-70	x
<i>Corpus</i> para a análises com vogal	360	288 (80%)
Total de segmentos-alvo	2620	2227 (85%)

Cabe ainda ressaltar que, dentre os dados, 1170 correspondem a plosivas não-vozeadas e 1057 a plosivas vozeadas.

Antes da realização da gravação, cada informante respondeu um questionário biográfico com a pesquisadora e assinou o termo de autorização de coleta de dados. Esses dois documentos encontram-se no Apêndice A. As informantes também foram informadas de que a pesquisa desenvolvida nesta tese encontra-se ligada a um projeto maior intitulado “O detalhe fonético: análise acústica exploratória de dados de

fala”, sob autoria da Profa. Dra. Izabel Christine Seara, orientadora desta tese, que foi aprovado pelo Comitê de Ética da UFSC sob o número 2057, FR:434924 (Anexo I).

Na primeira leitura, as informantes eram orientadas a inserir as palavras na frase “Digo \_\_\_\_\_baixinho”. Na segunda leitura, a frase foi modificada para “Digo \_\_\_\_\_pra ela”. Optamos por modificar a frase-veículo entre as leituras, pois pareceu-nos menos cansativo trocar o contexto seguinte à palavra-alvo, diversificando a leitura. Após a realização dos experimentos, todas as cinco participantes relataram que acharam menos cansativo a leitura da segunda frase.

Foram utilizados para a coleta um microfone dinâmico supercardióide *Le Son* modelo SM 50 Plus acoplado a um gravador digital profissional de 2 canais M-Audio *Microtrack* modelo 24/96. O microfone foi mantido a uma distância de 10 cm da boca do informante e os dados foram coletados em ambiente silencioso. A taxa de amostragem selecionada para a gravação foi de 44050Hz<sup>38</sup>. Os participantes foram instruídos a manterem um ritmo normal de fala durante a leitura das frases.

As palavras-alvo apareciam isoladamente na tela de um computador durante a leitura, uma de cada vez, sendo que, no início do experimento, as participantes eram instruídas a inseri-las nas frases-veículo. Lembramos que a ordem das palavras foi aleatorizada para cada uma das informantes.

### 3.4 ANÁLISE DE DADOS

#### 3.4.1 Etiquetagem

Após a gravação, os dados foram analisados através do *software* de análise acústica *Praat* (BOERSMA; WEENINK, 2013). Inspeções visuais dos oscilogramas e espectrogramas foram realizadas pela pesquisadora. Os dados foram etiquetados em oito camadas ou *tiers*. Na primeira camada, marcamos a palavra em que a plosiva estava inserida. Na segunda, etiquetamos a consoante com os códigos específicos para auxiliar na identificação dos contextos em que essa consoante estava inserida. Na terceira, selecionamos o momento de oclusão e, na quarta, o VOT. Na quinta camada, marcamos o momento de transição para a extração de valores de F1 e F2. Na sexta e sétima camadas, marcamos a vogal anterior e a seguinte à plosiva analisada. Na última camada,

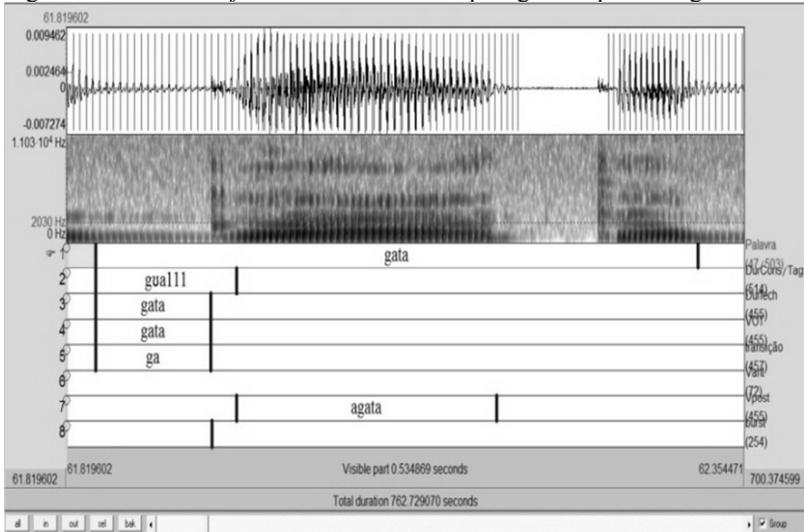
---

<sup>38</sup> Menor taxa disponível no gravador utilizado.

selecionamos o ponto de maior pico de energia do *burst* para a extração do valor do pico e dos momentos espectrais.

A Figura 19 ilustra a etiquetagem dos dados numa tela do *software Praat*.

Figura 19 – Tela do *software Praat* com a etiquetagem da palavra “gata”



Fonte: Praat/Nossos dados

Etiquetas foram utilizadas para classificar os parâmetros e contextos analisados, de acordo com os códigos apresentados no Quadro 3.

Quadro 3 – Códigos utilizados na etiquetagem dos dados

Posição do código	Descrição	Tag
1	Tipo de consoante	p, t, k, b, d, g
2	Tipo de vogal anterior	i, I, e, ε, a, ε, o, u, u
3	Tipo de vogal posterior	i, I, e, ε, a, ε, o, o, u, u
4	Tonicidade	1-tônico; 2-átono
5	Número de sílabas	1-dissílabo; 2-trissílabo
6	Posição na palavra	1-inicial; 2-medial

Assim, a etiquetagem da palavra “gata”, apresentada na Figura 19 (segunda camada) foi: gʌl11 e os códigos correspondentes foram:

- (1) – [g], plosiva velar vozeada;
- (2) – vogal antecedente [ʊ];
- (3) – vogal seguinte [a];
- (4) – 1, contextoônico;
- (5) – 1, dissílabo;
- (6) – 1, posição inicial na palavra.

### 3.4.2 Análise Acústica

As análises acústicas de todos os parâmetros abordados neste trabalho foram realizadas através do *software Praat*, versão 5.3.39, cuja distribuição é gratuita pelo *site* de seus desenvolvedores<sup>39</sup>.

As seguintes medições foram realizadas a partir dos dados etiquetados:

- (1) duração absoluta e relativa da consoante.
- (2) duração absoluta e relativa da oclusão da plosiva;
- (3) pico e momentos espectrais de *burst*;
- (4) duração absoluta e relativa do VOT em relação à palavra;
- (5) transição de F1 e F2
- (6) duração da vogal que precede e da que segue a plosiva;
- (7) taxa de elocução.

Cabe lembrar as medidas de duração relativa (da consoante, da oclusão, do VOT e das vogais) foram feitas com relação à palavra onde os segmentos em análise estavam inseridos.

Além das medições elencadas acima, realizamos uma análise qualitativa dos espectrogramas de cada um dos dados que apresentaram algum tipo de especificidade acústica, objetivando contabilizar e classificar as diferentes especificidades em nossos dados.

Abaixo, detalharemos os procedimentos adotados para a medição de cada um dos parâmetros aqui relacionados.

#### 3.4.2.1 Duração absoluta e relativa da consoante

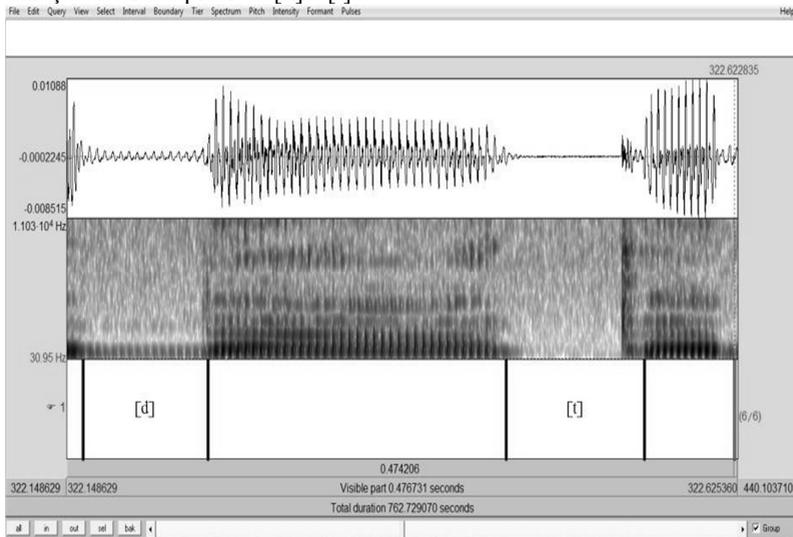
Para a coleta das medidas de duração absoluta e relativa, o cursor do ponto inicial foi posicionado no mesmo ponto inicial da duração da

---

<sup>39</sup> Boersma, P. e Weenik, D. *Praat- doing phonetics by computer*. Disponível em [www.praat.org](http://www.praat.org). Acesso em 27/05/2013.

oclusão e o cursor do ponto final foi posicionado no mesmo local do ponto final do VOT, obtendo-se, assim, a duração absoluta da consoante, conforme ilustra a Figura 20.

Figura 20 - Forma de onda e espectrograma da palavra "data". Em destaque: duração total das plosivas [d] e [t]



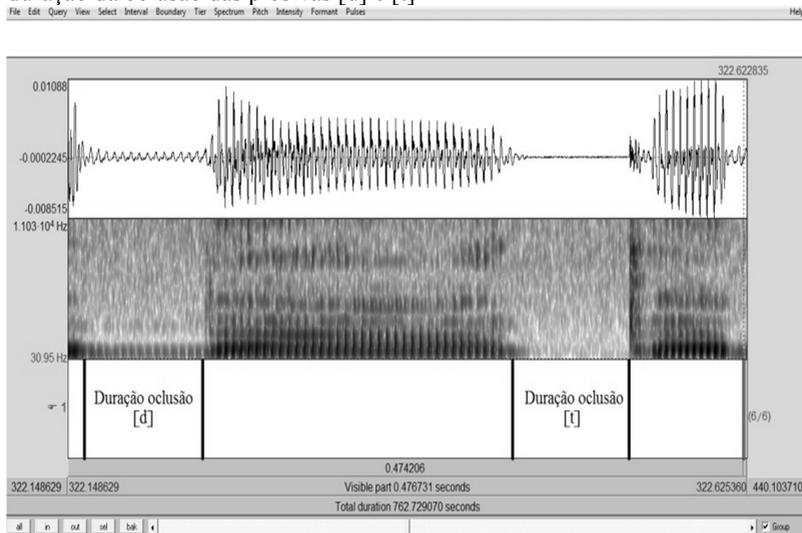
Fonte: Praat/Nossos dados

A duração relativa foi obtida pelo cálculo do percentual da duração absoluta do segmento plosivo tendo por referência a duração total da palavra ( $[\text{duração da consoante}/\text{duração da palavra}] \cdot 100$ ).

### 3.4.2.2 Duração da oclusão da plosiva

Para a medição da duração da oclusão, os cursores foram posicionados entre o último pulso visível da vogal que antecedia a plosiva até o início da liberação da oclusão, marcado antes da primeira explosão de *burst*. A duração da oclusão das plosivas vozeadas coincide com a medida de VOT. A Figura 21 traz exemplos da etiquetagem da duração da oclusão para plosivas vozeadas e não-vozeadas.

Figura 21 - Forma de onda e espectrograma da palavra "data". Em destaque: duração da oclusão das plosivas [d] e [t]



Fonte: Praat/Nossos dados

Para medirmos a oclusão de uma plosiva, devemos ser bastante criteriosos para que outras variações no espectro não interfiram nessa medição. Precisamos, assim, observar os *edge vibrations*, que como mencionado anteriormente, são vibrações que podem aparecer dentro da duração de uma plosiva não-vozeada, que são resquícios da vibração das pregas do seguimento anterior, representando um tipo do que chamamos de especificidade acústica.

Com base nisso, optamos por posicionar o cursor do ponto inicial no momento em que houve considerável diminuição na amplitude dos pulsos glotais da vogal anterior à plosiva, visível em oscilogramas, nos casos em que houve presença de *edge vibrations* no momento de oclusão da consoante não-vozeada. O mesmo procedimento foi utilizado na medição da duração total e relativa do segmento.

### 3.4.2.3 Especificidades acústicas

Uma inspeção visual de forma de onda e espectrograma foi realizada em todos os 2227 dados presentes em nosso banco de dados,

com o objetivo de avaliar quais dentre eles apresentavam uma das especificidades<sup>40</sup> relacionadas abaixo:

- (i) Plosiva não-vozeada totalmente vozeada
- (ii) Plosiva não-vozeada com *edge vibration*
- (iii) Plosiva não-vozeada com dupla explosão
- (iv) Plosiva não-vozeada com múltiplas explosões
- (v) Plosiva vozeada totalmente desvozeada
- (vi) Plosiva vozeada com interrupção de vozeamento na parte inicial da oclusão
- (vii) Plosiva vozeada com interrupção de vozeamento na parte medial da oclusão
- (viii) Plosiva vozeada com interrupção de vozeamento na parte final da oclusão
- (ix) Plosiva vozeada com dupla explosão
- (x) Plosiva vozeada com múltiplas explosões<sup>41</sup>

#### 3.4.2.4 Picos e momentos espectrais de *burst*

Foram utilizados os mesmos procedimentos de Forrest *et al.* (1988), Berti e Marino (2011) e Cristofolini (2013) para as medições do pico espectral e dos momentos espectrais de *burst*. São eles:

- pico espectral: maior pico de energia visível através da análise visual do espectro de frequência, valor fornecido pelo gráfico FFT<sup>42</sup>.
- centroide (*mean*): indica a média da concentração de energia do espectro. Corresponde ao primeiro momento espectral.
- variância (*variance*): aponta a variabilidade das frequências do espectro em relação à média. É o segundo momento espectral e corresponde ao quadrado do desvio padrão (HAYS, 1973).
- assimetria (*skewness*): medida de dispersão que indica o quão assimétrica é a distribuição em função da média. É o terceiro momento espectral. Em termos estatísticos, assimetria igual à

<sup>40</sup> Exemplos dessas especificidades podem ser observados no Capítulo 2, Seção 2.2.2.1.

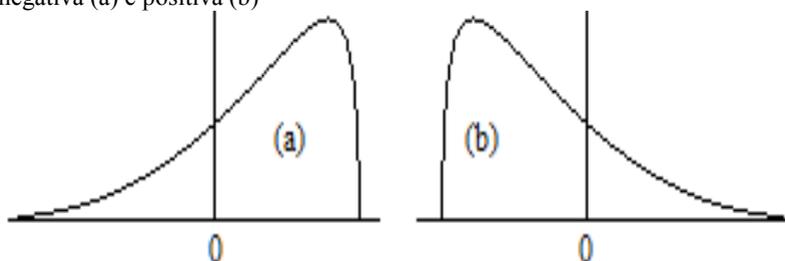
<sup>41</sup> A classificação das especificidades foi baseada na apresentada por Cristofolini (2013)

<sup>42</sup> FFT (*Fast Fourier Transform*): método utilizado para se calcular o espectro (a transformada de Fourier) de um sinal sonoro. É calculada a partir da decomposição da onda complexa em suas componentes senoidais, mostrada a partir de um gráfico de amplitude *versus* frequência (JOHNSON, 2003).

zero indica uma distribuição simétrica em torno da média. A assimetria negativa acontece quando a cauda do gráfico da distribuição se estende para a esquerda e a maior parte da distribuição se concentra a direita do gráfico, como na Figura 22a. A assimetria positiva apresenta maior extensão da cauda para a direita e maior parte da distribuição à esquerda do gráfico, conforme Figura 22b (HAYS, 1973; JONGMAN; WAYLAND; WONG, 2000; NEWELL; HANCOCK, 1984). De acordo com Jongman et al. (2000),

In phonetic terms, skewness refers to spectral tilt, the overall slant of the energy distribution. Positive skewness suggests a negative tilt with a concentration of energy in the lower frequencies. Negative skewness is associated with a positive tilt and a predominance of energy in the higher frequencies.<sup>43</sup>

Figura 22 – Gráfico de distribuição de frequências mostrando assimetria negativa (a) e positiva (b)



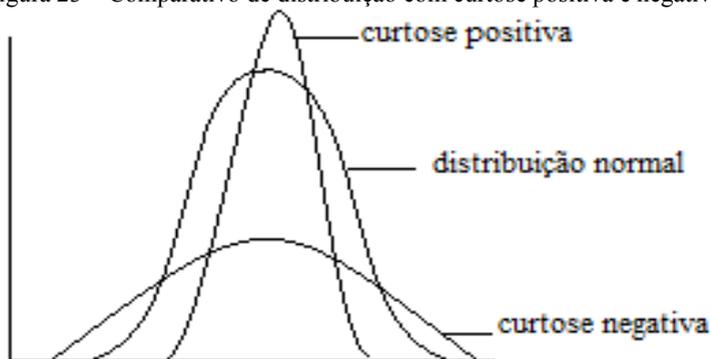
Fonte: Adaptado de Larson-Hall (2010), p.79.

- *curtose (kurtosis)*: medida que indica como se dá o pico de distribuição. A Figura 23 ilustra o gráfico comparativo de distribuição de frequências entre curtose positiva e negativa. Em termos de frequência de *burst*, curtose negativa indica um espectro achatado, com poucos picos ou sem picos definidos, enquanto uma curtose positiva indica um espectro bem definido, com picos altos ou ainda com múltiplos picos.

<sup>43</sup> “Em termos fonéticos, a assimetria se refere ao movimento no espectro, a inclinação geral da distribuição de energia. A assimetria positiva sugere uma movimentação negativa com a concentração de energia nas baixas frequências. A assimetria negativa está associada com uma movimentação positiva e a predominância de energia nas altas frequências.” (JONGMAN et al., 2000, p. 1253). Minha tradução.

Assim, um valor alto de curtose indica um espectro com mais picos e um valor baixo indica um espectro de característica mais achatada.

Figura 23 – Comparativo de distribuição com curtose positiva e negativa

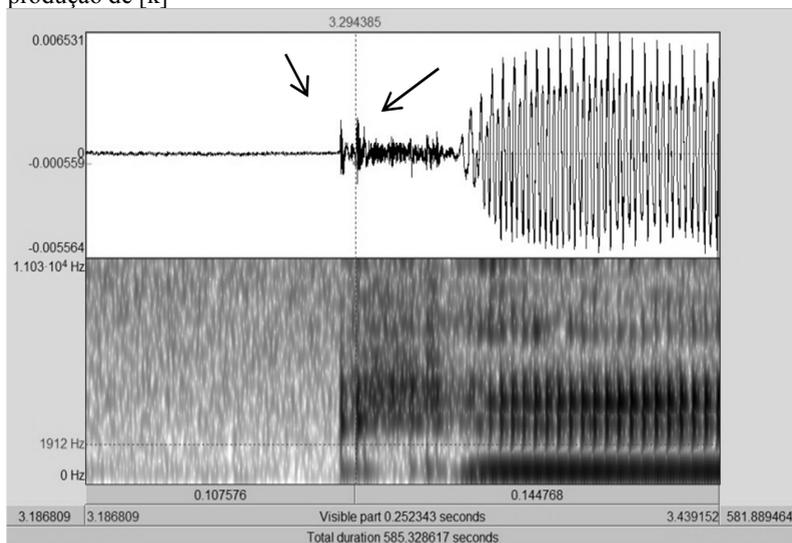


Fonte: Adaptado de Larson-Hall (2010), p.79.

As medições das características de *burst* foram obtidas manualmente através do *software* Praat e as configurações adotadas para o espectrograma seguiram o adotado por Cristofolini (2013) e Berti (2006), a saber: faixa de visualização (*view range*) de 0 a 11025 Hz; comprimento de janela (*window length*) de 0,02s; e faixa dinâmica (*dynamics range*) de 70 dB.

O ponto selecionado para a marcação do cursor para a extração das medidas foi obtido através de inspeção visual, posicionado na região em que foi detectado o maior pico de energia do *burst*, mesmo nos casos de múltiplos *bursts*, como nas plosivas velares (Figura 24).

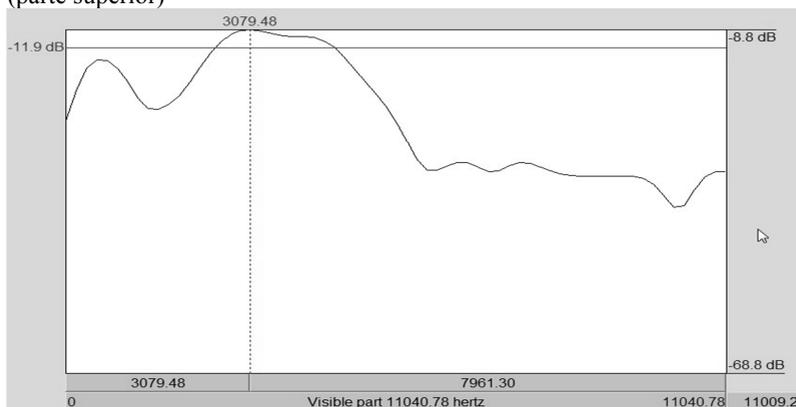
Figura 24 – Inspeção visual do pico de energia do *burst* para extração da FFT na produção de [k]



Fonte: Praat/Nossos dados

A partir dessa seleção, o espectro FFT que apresenta o espectro do ponto selecionado foi gerado através da opção *Spectrum* e *View Spectrum Slice* no Praat. Na janela “Praat Objects”, selecionamos o arquivo gerado e clicamos em *View & Edit*. A Figura 25 apresenta a FFT gerada para o segmento [k] mostrado na Figura 24.

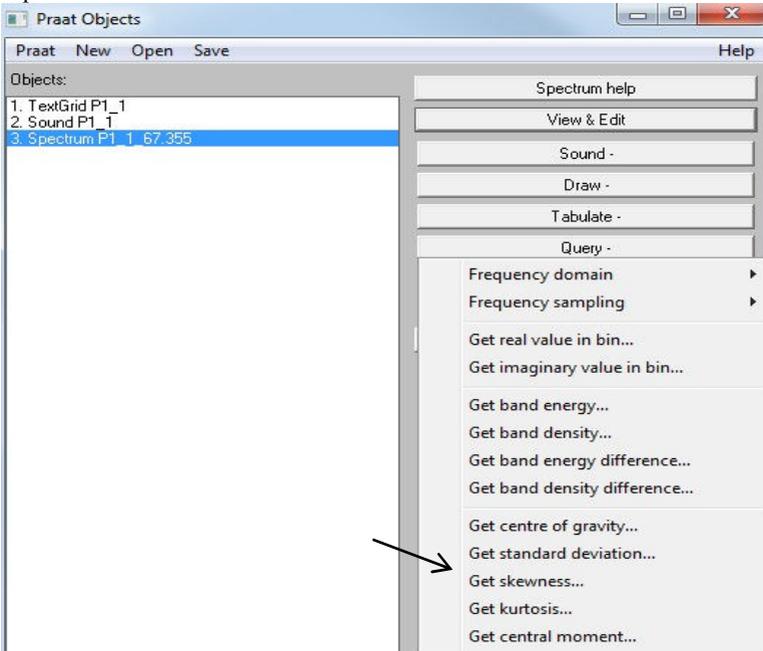
Figura 25 – FFT do segmento [k] com a indicação do pico espectral de 3079 Hz (parte superior)



Fonte: Praat/Nossos dados

Para a obtenção do centroide, voltamos à janela “Praat Objetes” e selecionamos o arquivo “Spectrumnomedoarquivo”. Na opção *Query*, obtêm-se todas as análises desejadas: *Get centre of gravity*, *Get standard deviation*, *Get skewness*, *Get kurtosis*, como demonstrado na Figura 26.

Figura 26 – Comandos no *software* Praat para obtenção dos momentos espectrais



Fonte: Praat/Nossos dados

Cabe lembrar que a variância corresponde ao quadrado do desvio padrão. Portanto, para obter-se o valor dessa medida, basta elevar ao quadrado o valor informado pelo *Praat* para *standard deviation*.

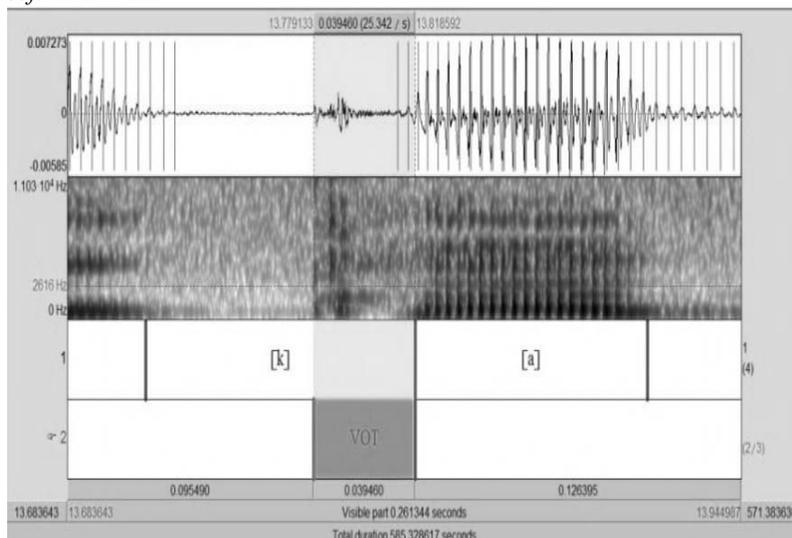
#### 3.4.2.5 Duração total e relativa de VOT

O VOT é definido como o momento entre a soltura da oclusão de uma plosiva e o início da vibração das pregas vocais do segmento adjacente.

Lisker e Abramson (1967) definem que o ponto de soltura, em espectrogramas, é claramente observado através de ruído de alta frequência.

Assim, optamos por selecionar o VOT da seguinte forma: no caso das plosivas não-vozeadas, posicionamos o cursor do ponto inicial na primeira explosão do *burst* e o cursor do ponto final na marca de 0 (*zero crossing*) do primeiro pulso regular da vogal seguinte (Figura 27).

Figura 27 – Exemplo de etiquetagem de VOT em plosiva não-vozeada [k] no software Praat

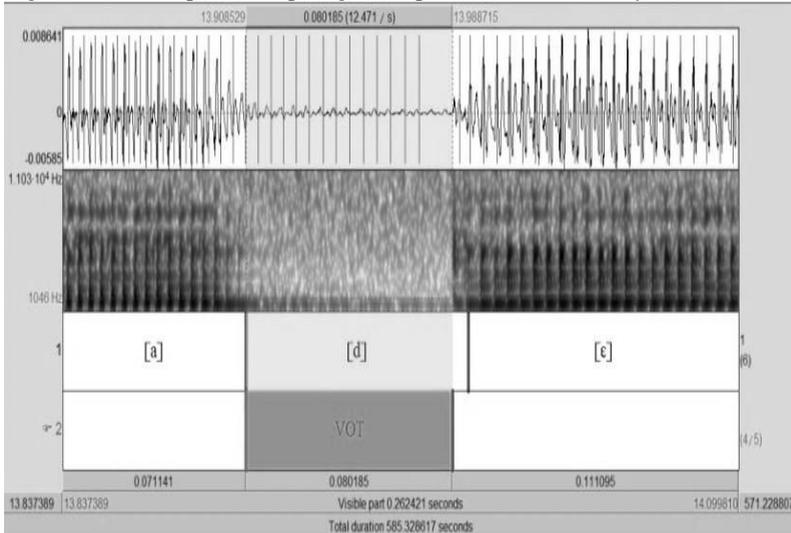


Fonte: Praat/Nossos dados

O cursor do ponto inicial das plosivas vozeadas foi posicionado quando detectamos visível atenuação nos pulsos periódicos da vogal anterior e, o cursor do ponto final, antes da primeira explosão do *burst*. A

Figura 28 ilustra a etiquetagem de uma plosiva vozeada. Nota-se, assim, que o VOT e a duração da oclusão da plosiva vozeada são coincidentes.

Figura 28 – Exemplo de etiquetagem de plosiva vozeada no *software Praat*



Fonte: Praat/Nossos dados

A duração relativa do VOT (em %) foi medida tendo como referência a duração total da palavra.

#### 3.4.2.6 Transição formântica

Para a averiguação da transição de F1 e F2 utilizamos um *script* (Apêndice C) capaz de selecionar, automaticamente, pontos dentro da transição CV. Por transição CV adotamos a definição de Kewley-Port (1982),

A formant transition was defined as the segment of a formant beginning at the onset of voicing and ending at the onset of the steady-state portion of the vowel (p. 380)<sup>44</sup>

Assim, selecionamos para cada plosiva o intervalo entre o final da plosiva e o início da vogal adjacente (primeiro terço da duração da vogal). Ficker (2003) coletou dados de transição em uma região similar, também seguindo o proposto por Kewley-Port (1982).

<sup>44</sup> “A transição formântica foi definida como o segmento de formantes iniciando no *onset* do vozeamento e terminando no *onset* da porção de região estável da vogal” (KEWLEY-PORT, 1982, p. 380). Minha tradução.

Utilizaremos os cinco pontos de frequência para mapear graficamente o desenho dos formantes na transição das plosivas para a vogal seguinte.

A análise de transição formântica será aplicada aos dados presentes no Quadro 2 e os contextos vocálicos da transição CV serão [ɐ] e [ʊ].

#### 3.4.2.7 Duração da vogal precedente/seguinte e qualidade vocálica

A medida da duração da vogal (tanto a que antecede quanto a que segue a plosiva) foi realizada através do cursor do ponto inicial posicionado no primeiro pulso regular, através da inspeção visual de oscilograma e do cursor do ponto final posicionado no último pulso regular da vogal.

Para a coleta da duração da vogal que antecede a plosiva, optamos por avaliar apenas a vogal [a]. A vogal [a] em posição tônica foi escolhida por ser a que mais facilmente nos possibilitou a formação desses pares.

A influência da vogal que segue a plosiva foi averiguada através de possíveis alterações na duração da plosiva e do seu VOT. Para tal fim, selecionamos palavras que apresentassem as plosivas, vozeadas ou não, seguidas das vogais átonas finais [ɐ] e [ʊ], através das palavras dispostas no Quadro 2.

#### 3.4.2.8 Taxa de elocução

O parâmetro taxa de elocução é medido em sílabas por segundo e foi calculado a partir da contagem do número de sílabas da palavra dividido pela duração absoluta da palavra. Essa medição foi feita nas duas repetições de cada informante.

### 3.5 Estatística utilizada na tese

Esta pesquisa utilizou-se, majoritariamente, de testes estatísticos de diferença, uma vez que as perguntas de pesquisa eram endereçadas à investigação de possíveis diferenças entre parâmetros acústicos e/ou contextos relacionados às plosivas vozeadas e não-vozeadas do PB.

O programa utilizado para a aplicação dos testes estatísticos foi o *SPSS Statistics*, versão 17. Primeiramente, testes de normalidade e homogeneidade foram aplicados, em conjunto com a observação de

curtose, assimetria e histograma, para se verificar se a distribuição dos dados era normal ou não, para a escolha entre testes paramétricos (distribuição normal) ou testes não paramétricos. Dependendo desses resultados, testes Anova unifatorial (teste paramétrico) ou Kruskal-Wallis (teste não-paramétrico) foram aplicados a avaliações entre três grupos. Quando houve diferença significativa no teste Anova Unifatorial, um teste *post hoc* foi aplicado (*Gabriel*)<sup>45</sup> para se observar entre quais grupos estavam as diferenças. Como o teste Kruskal-Wallis também não aponta diretamente entre quais membros encontram-se a diferença, testes subsequentes de Mann-Whitney com correção Bonferroni foram aplicados entre cada um dos três grupos.

Para avaliações entre dois grupos, Teste-t de amostras independentes (teste paramétrico) ou Mann-Whitney (teste não-paramétrico) foram utilizados. O nível de significância utilizado foi de  $p < 0,05$  e de  $p < 0,017$  no caso de testes com correção Bonferroni ( $0,05/3$ grupos,  $p = 0,017$ ). Contudo, para indicar o grau de significância nos testes, reportarmos os valores como apresentados na primeira coluna do Quadro 4.

Quadro 4 - Valores de  $p$  utilizados para se reportar os resultados

<b>Valor de <math>p</math></b>	<b>Resultado</b>
$p > 0,05$	não-significativo
$p < 0,05$	significativo
$p < 0,01$	significativo
$p < 0,001$	altamente significativo <sup>46</sup>
$p < 0,017$	significativo - correção Bonferroni para testes <i>post hoc</i> com três grupos

Ao final desta tese, no Apêndice D, trazemos um resumo detalhado de pressupostos teóricos relacionados à Estatística aplicada na pesquisa, bem como apresentamos as escolhas dos testes feitas para cada uma das perguntas de pesquisa, com os resultados obtidos. Esse resumo objetiva auxiliar aqueles leitores que, por ventura, tiverem interesse de observar o processo de raciocínio desenvolvido para a escolha dos testes aplicados a cada uma das perguntas de pesquisa. Objetiva também

<sup>45</sup> De acordo com Field (2009) e Martins (2011), o teste *post hoc* de Gabriel é um dos mais indicados para grupos que contenham diferente número de dados.

<sup>46</sup> Field (2009) relata que um valor de  $p$  menor do que 0,001 pode ser interpretado como altamente significativo (*highly significant*), pois é um valor de significância menor do que 0,05 (p.182)

auxiliar aqueles que possam ter pouca familiaridade com estatística e que queiram consultar conceitos básicos para o entendimento das escolhas dos testes aqui aplicados.

## 4 ANÁLISE DE DADOS E RESULTADOS

Este capítulo apresentará os resultados obtidos através da análise dos dados coletados para a presente pesquisa. A apresentação dos resultados será feita a partir de cada uma das perguntas de pesquisa desta tese, listadas no Capítulo 2.

### 4.1 TAXA DE ELOCUÇÃO

Conforme descrito anteriormente, é preciso atentar para a taxa de elocução apresentada pelos participantes, pois discrepâncias nessa medida podem ser responsáveis por diferenças encontradas em parâmetros de duração, como alongamento ou redução na duração dos segmentos. Como muitos dos parâmetros investigados nesta pesquisa são de ordem duracional, é preciso atentar possíveis diferenças na taxa de elocução das participantes. Assim, a Tabela 7 apresenta os dados relativos à taxa de elocução de cada uma das informantes que participaram desta pesquisa.

Tabela 7 – Valores médios, desvio padrão e faixa de distribuição da taxa de elocução por participante (sílabas/segundo)

<b>Participante</b>	$\bar{X}$	<b>Desvio Padrão</b>	<b>Faixa</b>
P1	5,2	0,9	3,4 – 7,6
P2	4,7	0,8	3,2 – 6,9
P3	4,7	0,7	3,3 – 6,8
P4	5,1	0,8	3,4 – 7,8
P5	4,5	0,7	3,0 – 7,4

De acordo com Klatt (1976), uma taxa de elocução considerada normal se estabelece entre 4 a 7 sílabas por segundo. Como podemos observar pela Tabela 7, os valores médios encontrados para as nossas participantes estão em conformidade com os valores apontados por Klatt (1976) como uma taxa de elocução considerada normal. Também verificamos que não houve diferença estatística dos valores da taxa de elocução entre as participantes, pois  $p > 0,05$ .

## 4.2 DURAÇÃO DA CONSOANTE

### 4.2.1 Duração da consoante e vozeamento

Conforme já relatado na Seção 2.2, a duração da consoante tende a ser mais longa para as plosivas não-vozeadas do que para as vozeadas, devido à dificuldade de se manter o vozeamento prolongado durante a produção de uma plosiva vozeada. Vimos também que, para o PB, essa tendência parece se manter na comparação entre plosivas homorgânicas (BARBOSA, 1996; FICKER, 2003, CRISTOFOLINI, 2013). Baseados nessas informações, a **Pergunta 1** questiona: Há diferença entre a duração das consoantes plosivas vozeadas e não-vozeadas?

Os resultados de duração absoluta e relativa da consoante, na comparação entre as categorias vozeada e não-vozeada, estão resumidos na Tabela 8.

Tabela 8 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com a comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas vozeadas e não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	Não-voz <sup>47</sup>	126(23)	58– 202	$U = 155442$	<b>&lt; 0,001</b>
	Voz	96 (22)	38 – 175		
<b>Duração Relativa</b>	Não-voz	25 (6)	9 – 59	$U = 219545$	<b>&lt; 0,001<sup>48</sup></b>
	Voz	19 (6)	6 – 39		

Os valores médios de duração da consoante encontrados para as plosivas não-vozeadas foram maiores do que para as vozeadas, de acordo com o indicado pela literatura (BARBOSA, 1996, FICKER, 2003, BARROCO *et al.*, 2007; CRISTOFOLINI, 2003). Testes Mann-Whitney revelaram que a duração da consoante não-vozeada é significativamente maior do que a da vozeada.

Na comparação entre os pares de plosivas homorgânicas, as plosivas não-vozeadas apresentaram, em todos os casos, valores médios de duração absoluta e relativa mais elevados do que as vozeadas,

<sup>47</sup> Os termos “não-vozeados” e “vozeados” foram abreviados para “não-voz” e “voz”, respectivamente, nas tabelas deste capítulo.

<sup>48</sup> Quando o valor de p for menor do que 0,05, indicando diferença significativa entre a comparação, destacamos o resultado em negrito nas tabelas apresentadas neste capítulo.

conforme a Tabela 9. Resultados semelhantes foram também encontrados por Ficker (2003).

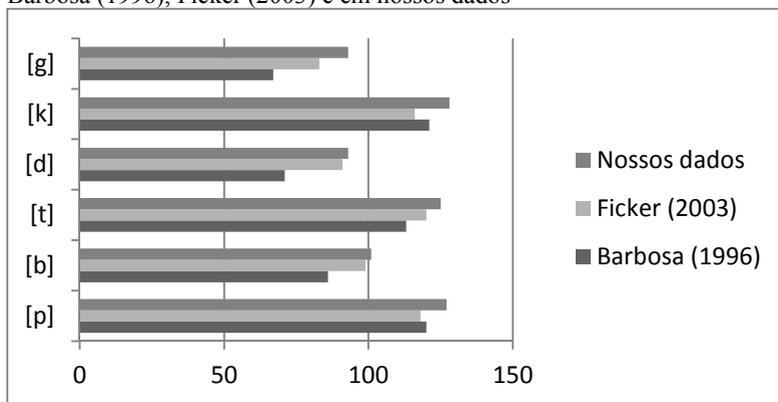
Tabela 9 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) entre os pares homorgânicos

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[p]	127 (23)	67 – 198	$U = 22978$	<b>&lt; 0,001</b>
	[b]	101 (21)	52 – 175		
	[t]	125 (23)	66 – 202	$t(648) = 18,34$	<b>&lt; 0,001</b>
	[d]	94 (22)	43 – 165		
	[k]	128 (24)	58 – 202	$U = 12993$	<b>&lt; 0,001</b>
[g]	93 (24)	38 – 171			
<b>Duração Relativa</b>	[p]	25 (6)	12 – 43	$U = 29259$	<b>&lt; 0,001</b>
	[b]	20 (5)	10 – 37		
	[t]	25 (6)	11 – 59	$U = 25891$	<b>&lt; 0,001</b>
	[d]	19 (6)	7 – 39		
	[k]	26 (6)	9 – 51	$U = 18421$	<b>&lt; 0,001</b>
[g]	19 (6)	6 – 39			

Os testes estatísticos de Mann-Whitney foram aplicados a cada uma das comparações entre os pares de plosivas homorgânicas. Todas as comparações se mostraram significativas, tanto com relação à duração absoluta da consoante, quanto à duração relativa (Tabela 9).

O Gráfico 1 mostra uma comparação entre os dados encontrados nas pesquisas de Barbosa (1996) e Ficker (2003) e os nossos.

Gráfico 1 - Resultados de média de duração absoluta da consoante (ms) em Barbosa (1996), Ficker (2003) e em nossos dados



Nota-se, no Gráfico 1, similaridade entre as médias de duração da consoante para as três pesquisas, especialmente entre as encontradas por Ficker (2003) e pelo nosso estudo. Assim, podemos afirmar que a duração da consoante se diferencia na produção das plosivas vozeadas e não-vozeadas do PB, podendo ser um dos parâmetros utilizados para a distinção desses segmentos.

Assim, considerando a **Hipótese 1** (de que haveria diferença entre as consoantes plosivas vozeadas e não-vozeadas, sendo que as últimas apresentariam valores de duração maiores do que as primeiras), concluímos que ela pode ser confirmada, uma vez que as plosivas não-vozeadas apresentaram maior duração absoluta e relativa do que as vozeadas e que as diferenças encontradas entre plosivas vozeadas e não-vozeadas e entre os pares homorgânicos são estatisticamente relevantes. Esse resultado, por corroborar outras pesquisas para o PB, pode ser um indicativo de que esse parâmetro possa ser usado por falantes do PB para diferenciar a produção de plosivas e que, provavelmente, também seja uma pista utilizada por ouvintes do PB para diferenciá-las.

#### 4.2.2 Duração da consoante e ponto de articulação

Outra questão a ser investigada com relação à duração da consoante é a possível influência do ponto de articulação. Assim, a **Pergunta 2.1** questiona: há influência do ponto de articulação na duração da consoante entre as plosivas não-vozeadas?

De acordo com a literatura, as plosivas não-vozeadas apresentam valores de duração de consoante bastante similares entre si, parecendo não haver consenso sobre qual dentre elas apresenta valores mais elevados. Barbosa (1996) e Cristofolini (2013)<sup>49</sup> encontraram valores de [t] < [p] < [k], enquanto Ficker (2003) encontrou valores de [k] < [p] < [t], como se pode observar através da Tabela 10. Além disso, testes estatísticos apontaram que não houve diferença significativa entre os valores absolutos de duração da consoante no estudo de Ficker (2003), na comparação entre plosivas não-vozeadas.

Tabela 10 – Valores médios de duração absoluta (ms) das plosivas do PB nos estudos de Barbosa (1996) e de Ficker (2003)

	[p]	[b]	[t]	[d]	[k]	[g]
<b>Barbosa (1996)</b>	120	86	113	71	121	67
<b>Ficker (2003)</b>	118	99	120	91	116	83

A Tabela 11 apresenta os resultados encontrados em nosso estudo. Pode-se observar que apesar de os valores médios de duração absoluta e relativa encontrados para a plosiva velar [k] serem mais elevados do que os demais, o teste Anova Unifatorial não apontou diferenças significativas entre a duração absoluta das consoantes plosivas não-vozeadas (Tabela 11). Esse resultado também foi observado por Ficker (2003). O teste Kruskal-Wallis também revelou que não há diferenças entre a duração relativa dessas mesmas consoantes (Tabela 11).

<sup>49</sup> Os resultados no estudos de Cristofolini (2013) são referentes à duração relativa da consoante em função da duração da palavra.

Tabela 11 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[p]	127 (23)	67 – 198	$F(2, 1012) = 1,03$	$> 0,05$
	[t]	126 (23)	66 – 202		
	[p]	127 (23)	67 – 198		
	[k]	128 (24)	58 – 202		
	[t]	126 (23)	66 – 202		
	[k]	128 (24)	58 – 202		
<b>Duração Relativa</b>	[p]	25 (6)	12 – 43	$H(2) = 3,23$	$> 0,05$
	[t]	25 (6)	11 – 59		
	[p]	25 (6)	12 – 43		
	[k]	26 (6)	9 – 51		
	[t]	25 (6)	11 – 59		
	[k]	26 (6)	9 – 51		

Já para as vozeadas, os resultados de estudos anteriores para o PB mostram que, quanto mais posterior o ponto de articulação, menor o valor médio de duração da consoante, conforme a Tabela 10. Esse mesmo resultado foi observado por Cristofolini (2013) para a medida de duração relativa das vozeadas. Ficker (2003) observou relevância estatística somente entre [b] e [g], dentre as vozeadas (o valor de [g] foi menor que o de [b], conforme a Tabela 10). A pesquisadora não encontrou diferenças significativas entre a duração das plosivas não-vozeadas. Assim, a Pergunta 2.2 questiona: há influência do ponto de articulação na duração da consoante entre as plosivas vozeadas?

A

Tabela 12 traz um resumo dos resultados da duração absoluta e relativa entre as plosivas vozeadas para nosso estudo. Os valores absolutos e relativos dessas consoantes seguiram a tendência encontrada por outras pesquisas (BARBOSA, 1996; FICKER, 2003; CRISTOFOLINI, 2013), ou seja, [b] > [d] > [g].

Tabela 12 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da consoante, juntamente com os resultados dos testes obtidos na comparação de duração absoluta (ms) e relativa (%) da consoante entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[b]	101 (21)	52 – 175	$U = 43568$	<b>&lt; 0,017</b>
	[d]	94 (22)	43 – 165		
	[b]	101 (21)	52 – 175	$U = 31912$	<b>&lt; 0,017</b>
	[g]	93 (24)	38 – 171		
	[d]	94 (22)	43 – 165	$U = 41179$	<b>&gt; 0,017</b>
	[g]	93 (24)	38 – 171		
<b>Duração Relativa</b>	[b]	20 (5)	10 – 37	$U = 47798$	<b>&lt; 0,017</b>
	[d]	19 (6)	7 – 39		
	[b]	20 (5)	10 – 37	$U = 34176$	<b>&lt; 0,017</b>
	[g]	19 (6)	6 – 39		
	[d]	19 (6)	7 – 39	$U = 40715$	<b>&gt; 0,017</b>
	[g]	19 (6)	6 – 39		

O teste Kruskal-Wallis mostrou que há diferença significativa na duração absoluta e relativa entre as três plosivas vozeadas ( $H(2) = 31,54$ ,  $p < 0,001$ ). No entanto, na comparação de pares de plosivas vozeadas, os testes Mann-Whitney com correção Bonferroni apontaram diferenças significativas na duração absoluta e relativa somente entre a consoante bilabial e a alveolar, e entre a bilabial e a velar (corroborando o resultado de FICKER, 2003), conforme podemos observar pela

Tabela 12. As plosivas alveolares e velares não mostraram diferenças significativas entre si.

Os resultados acima apontam para a confirmação das **Hipóteses 2.1 e 2.2** que consideravam que o ponto de articulação influenciaria somente a duração das consoantes plosivas vozeadas, tornando-se menor a medida que o ponto de articulação se posterioriza. Os testes estatísticos confirmaram que não há diferença entre os valores de

duração absoluta e relativa para as plosivas não-vozeadas (Hipótese 2.1) e que, para as vozeadas, [b] diferencia-se significativamente das demais plosivas vozeadas (Hipótese 2.2). A consoante bilabial vozeada apresentou valores maiores de duração absoluta e relativa do que as demais consoantes cujos pontos de articulação são mais posteriores. Esse resultado vai ao encontro de estudos anteriores do PB.

Com base nesses resultados, acreditamos que o parâmetro duração da consoante possa ser uma pista para a identificação das categorias de plosivas, isto é, vozeadas e não-vozeadas. Com relação ao ponto, a duração da consoante é distinta para as plosivas vozeadas, mas pode não ter influência tão expressiva na diferenciação das plosivas não-vozeadas entre si.

### 4.3 DURAÇÃO DA OCLUSÃO

#### 4.3.1 Duração da oclusão e vozeamento

Conforme apontado na literatura (BARROCO *et al.*, 2007; LADEFOGED; MADDIESON, 1996; LADEFOGED; JOHNSON, 2011, dentre outros), as plosivas não-vozeadas tendem a apresentar valores mais elevados de duração da oclusão. Baseando-se nessa informação, a **Pergunta 3** questiona: há diferenças entre a duração da oclusão de plosivas vozeadas e não-vozeadas?

Os resultados apresentados na Tabela 13 mostram que os valores médios encontrados para a duração absoluta e relativa da oclusão de plosivas não-vozeadas são maiores do que os valores médios obtidos para a duração absoluta e relativa da oclusão de plosivas vozeadas.

Tabela 13 – Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre plosivas vozeadas e não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	Não-voz	93 (26)	17 – 176	$U = 442402$	<b>&lt; 0,001</b>
	Voz	81 (22)	21 – 166		
<b>Duração Relativa</b>	Não-voz	19 (6)	3 – 41	$U = 480337$	<b>&lt; 0,001</b>
	Voz	16 (5)	4 – 36		

O teste Mann-Whitney revelou que a duração absoluta e relativa da oclusão das plosivas não-vozeadas é significativamente maior do que a das vozeadas (Tabela 13). Esse resultado corrobora estudos anteriores, tanto com relação a outras línguas, quanto para o PB (BARROCO et al., 2007; LADEFOGED; MADDIESON, 1996; MELO *et al.*, 2012, dentre outros).

Também encontramos valores médios de duração absoluta e relativa de oclusão maiores para as plosivas não-vozeadas, quando consideramos a comparação entre os pares de plosivas homorgânicas, cujos resultados estão dispostos na Tabela 14.

Tabela 14 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre os pares homorgânicos

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[p]	104 (23)	27 – 176	$U = 45708$	< <b>0,001</b>
	[b]	89 (21)	31 – 166		
	[t]	101 (23)	17 – 162	$U = 37812$	< <b>0,001</b>
	[d]	81 (21)	32 – 155		
	[k]	74 (21)	17 – 146	$U = 51001$	< <b>0,01</b>
	[g]	70 (18)	21 – 134		
<b>Duração Relativa</b>	[p]	21 (5)	6 – 40	$U = 48708$	< <b>0,001</b>
	[b]	18 (5)	7 – 34		
	[t]	20 (6)	4 – 41	$U = 46569$	< <b>0,001</b>
	[d]	17 (5)	5 – 36		
	[k]	15 (4)	3 – 31	$U = 18421$	< <b>0,01</b>
	[g]	14 (4)	4 – 31		

Os testes estatísticos revelaram que as diferenças na duração absoluta e relativa entre cada um dos pares homorgânicos foi significativamente maior para as plosivas não-vozeadas (Tabela 14). Esses resultados corroboram os achados de Melo et al. (2012) para o PB.

Desse modo, considerando-se a **Hipótese 3** de que a duração da oclusão em plosivas não-vozeadas tenderia a ser maior do que o mesmo momento em plosivas vozeadas, conclui-se que ela pode ser confirmada. As plosivas não-vozeadas apresentaram valores de duração de oclusão significativamente maiores do que as suas contrapartes vozeadas, tanto na comparação da duração da oclusão entre as duas categorias, isto é,

entre não-vozeadas e vozeadas, quanto na comparação do entre os pares de plosivas homorgânicas.

### 4.3.2 Duração da oclusão e ponto de articulação

Outro fator que pode influenciar a duração da oclusão das plosivas é o ponto de articulação. Assim, a **Pergunta 4.1** questiona: há influência do ponto de articulação na duração da oclusão entre as plosivas não-vozeadas?

Em relação às plosivas não-vozeadas, a Tabela 15 mostra que a plosiva velar foi aquela que apresentou menor duração absoluta e relativa da oclusão, corroborando os resultados de Seara (2000); Berti e Marino (2011) e Melo *et al.* (2012).

Tabela 15 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da oclusão entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[p]	104 (23)	27 – 176	$U = 67829$	$> 0,05$
	[t]	101 (23)	17 – 162		
	[p]	104 (23)	27 – 176	$U = 26357$	$< 0,017$
	[k]	74 (21)	17 – 146		
	[t]	101 (23)	17 – 162	$U = 28822$	$< 0,017$
	[k]	74 (21)	17 – 146		
<b>Duração Relativa</b>	[p]	21 (5)	6 – 40	$H(2) = 3,507$	$> 0,05$
	[t]	20 (6)	4 – 41		
	[p]	21 (5)	6 – 40		
	[k]	15 (5)	3 – 31		
	[t]	20 (6)	4 – 41		
	[k]	15 (5)	3 – 31		

O teste estatístico de Kruskal-Wallis apontou que há diferenças significativas entre a duração absoluta da oclusão ( $H(2) = 327,39$ ,  $p < 0,001$ ), mas não revelou diferença significativa na duração relativa da oclusão entre as plosivas não-vozeadas. Considerando os valores absolutos, os testes Mann-Whitney com correção Bonferroni mostraram

que há apenas diferença significativa na comparação entre a duração da oclusão da velar e as demais (Tabela 15).

Para as vozeadas, a **Pergunta 4.2** questiona: há influência do ponto de articulação na duração da oclusão entre as plosivas vozeadas?

Podemos notar, através da Tabela 16, que a bilabial foi a que apresentou maior duração absoluta e relativa, seguida por alveolar e velar, de acordo com os valores apresentados na Tabela 16. O teste Kruskal-Wallis revelou que há diferenças significativas na duração absoluta ( $H(2) = 145,29$ ,  $p < 0,001$ ) e relativa ( $H(2) = 102,49$ ,  $p < 0,001$ ) da oclusão entre as três plosivas vozeadas. Testes Mann-Whitney com correção Bonferroni foram aplicados a cada uma das comparações entre plosivas vozeadas e os resultados apresentados na Tabela 16 mostram que a diferença foi significativa entre todas as comparações.

Tabela 16 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para a duração da oclusão, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa da consoante entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração Absoluta</b>	[b]	89 (21)	31 – 166	$U = 55794$	<b>&lt; 0,017</b>
	[d]	81 (21)	43 – 165		
	[b]	89 (21)	31 – 166	$U = 25273$	<b>&lt; 0,017</b>
	[g]	70 (18)	21 – 134		
	[d]	81 (21)	43 – 165	$U = 38463$	<b>&lt; 0,017</b>
	[g]	70 (18)	21 – 134		
<b>Duração Relativa</b>	[b]	18 (5)	7 – 34	$U = 62558$	<b>&lt; 0,017</b>
	[d]	17 (5)	5 – 36		
	[b]	18 (5)	7 – 34	$U = 29835$	<b>&lt; 0,017</b>
	[g]	14 (4)	4 – 31		
	[d]	17 (5)	5 – 36	$U = 39587$	<b>&lt; 0,017</b>
	[g]	14 (4)	4 – 31		

As **Hipóteses 4.1** e **4.2** estabeleciam que, quanto mais posterior o ponto de articulação, menor tenderia a ser a duração da oclusão. De acordo com os resultados apresentados na Tabela 16, a Hipótese 4.1 se confirma parcialmente, pois não encontramos diferenças significativas entre os valores de duração da oclusão entre [p] e [t]. No entanto, podemos afirmar que a Hipótese 4.2 se confirmou, pois a duração da

oclusão das plosivas vozeadas diminui significativamente conforme o ponto de articulação vai se tornando mais posterior.

Em suma, a duração da oclusão parece se estabelecer como um parâmetro que se diferencia na produção conforme o vozeamento. Com relação à influência do ponto de articulação, vimos que quanto mais posterior é a região em que a plosiva é articulada, menor duração de oclusão ela apresenta (apesar de essa diferença ser pequena e não-significativa para [p] e [t]).

#### 4.4 ESPECIFICIDADES ACÚSTICAS

Como vimos anteriormente, nem todos os segmentos são produzidos de acordo com a descrição apresentada pela literatura. Dados que exibem, por exemplo, vozeamento dentro do momento de oclusão de uma plosiva não-vozeada (*edge vibrations*) ou ainda interrupção do vozeamento para uma plosiva vozeada apresentam as chamadas “especificidades acústicas”<sup>50</sup> (LSKER; ABRAMSON, 1964b; CRISTOFOLINI, 2008, 2013;). Esses dados não acontecem somente com a fala infantil, conhecida por ainda não apresentar todas as regularidades e refinamentos articulatórios, mas também na fala adulta, que já está estabilizada (CRISTOFOLINI, 2008; 2013). Assim, a **Pergunta 5** questiona: os dados de falantes adultos também apresentarão alguma especificidade acústica, como as apresentadas em dados de estudos com crianças falantes nativas do PB?

Dentre os 2227 dados válidos, 526 deles apresentaram algum tipo de especificidade, o que corresponde a 24% do total. A Tabela 17 traz os resultados de dados que apresentaram algum tipo de especificidade acústica, referentes a cada plosiva.

Tabela 17 - Total de dados e de especificidades por plosivas

	[p]	[t]	[k]	[b]	[d]	[g]
<b>Total de dados</b>	396	372	402	370	389	298
<b>Com especificidades</b>	65	78	233	15	39	96
<b>Com especificidades (%)</b>	16%	21%	58%	4%	10%	32%

<sup>50</sup> O termo “especificidades acústicas” foi cunhado por Cristofolini (2013).

Para as plosivas não-vozeadas (1170 no total), 32% dos dados apresentaram especificidades. Dentre esse percentual de dados em plosivas não-vozeadas, 17% ocorreram eram bilabiais, 21% eram alveolares e 62% eram velares.

Dentre os dados de plosivas vozeadas (1057 no total), 14% apresentaram algum tipo de especificidade, dos quais 10% eram plosivas bilabiais, enquanto 26% eram em plosivas alveolares e 64% velares.

Assim como Cristofolini (2013), grande parte dos dados de plosivas velares apresentaram especificidades (58% para [k] e 32% para [g]). **PelaErro! Fonte de referência não encontrada.** Tabela 18, é possível observar que, para a plosiva velar não-vozeada, o maior número de dados apresentou múltiplas explosões (57%), seguido por duplas explosões (37%), enquanto para a velar vozeada, a maior parte foi de dados com interrupção do vozeamento na parte final da oclusão, seguido por duplas explosões.

Tabela 18 – Total de dados com especificidade por plosiva, de acordo com o tipo apresentado<sup>51</sup>

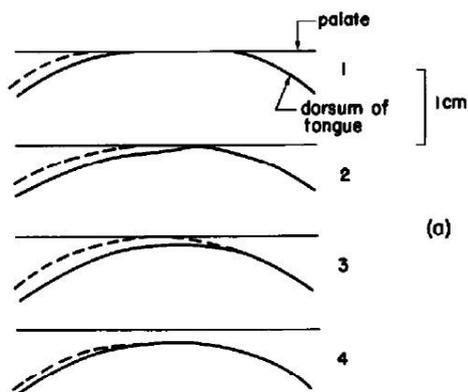
	[p]	[t]	[k]	[b]	[d]	[g]
<b>Edge vibration</b>	44 (68%)	47 (60%)	14 (6%)	x	x	x
<b>Dupla explosão</b>	3 (4%)	2 (3%)	86 (37%)	1 (7%)	6 (15%)	32 (33%)
<b>Múltiplas explosões</b>	18 (28%)	29 (37%)	133 (57%)	6 (40%)	5 (13%)	14 (15%)
<b>Int. voz. medial</b>	x	x	x	2 (13%)	1 (3%)	7 (7%)
<b>Int. voz. final</b>	x	x	x	6 (40%)	27 (69%)	43 (45%)
<b>Total de dados</b>	65 (100%)	78 (100%)	233 (100%)	15 (100%)	39 (100%)	96 (100%)

É possível que o ponto de contato entre os articuladores possa estar influenciando a produção de dados de plosivas velares com duplas ou múltiplas explosões. Segundo Stevens (2000), a área de contato entre o dorso da língua e o palato mole (articuladores ativo e passivo das velares, respectivamente) é relativamente maior do que para as demais plosivas. Assim, o tempo que eles levam para se distanciar totalmente,

<sup>51</sup> Trazemos exemplos de especificidades dos nossos dados no Capítulo 2.

liberando o ar retido atrás do bloqueio é maior do que para os articuladores envolvidos na produção dos outros tipos de plosivas. Isso faz com que o afastamento dos articuladores seja gradual, possibilitando que o ar saia gradativamente, à medida que o trato vai se abrindo, podendo gerar não uma, mas duas ou mais explosões, conforme exemplificado pela Figura 29.

Figura 29 - Abertura do trato oral em uma plosiva velar, focalizando o distanciamento entre os articuladores



Fonte: Stevens (2000, p. 329)

Nota-se, a partir da Figura 29, que o dorso da língua, representado pela linha cheia em curva, afasta-se gradativamente do palato (linha reta). Desse modo, quando os articuladores se distanciam pela primeira vez, há a primeira explosão. À medida que eles se afastam, podem ocorrer outras explosões, devido à forte pressão de ar retida atrás da articulação. Em seguida, quando os articuladores não mais se tocam, mas ainda estão bastante próximos, ocorre uma pequena fricção (já apontada por KLATT, 1975). Por fim, os articuladores se afastam e a língua assume a configuração necessária para a realização do próximo segmento.

No caso das plosivas vozeadas, houve um grande número de dados com interrupção de vozeamento na parte final da oclusão, tanto para as velares (45%), quanto para as demais (69% para [d] e 40% para [b]). Como vimos anteriormente, Clark e Yallop (1995) atribuem essa ocorrência a um “desequilíbrio aerodinâmico” entre as pressões supra e subglotais, que impede a continuidade do vozeamento por todo o segmento que deveria ser vozeado.

Os dados acima confirmam a **Hipótese 5** de que dados de falantes adultos apresentariam alguma especificidade, sendo o tipo mais comum a presença de *edge vibrations* nas plosivas não-vozeadas e interrupções de vozeamento nas vozeadas.

Cabe ressaltar que obtivemos um número considerável de especificidades para informantes adultos (24% do total), pois foi muito mais expressivo do que o encontrado por Cristofolini (2013) para os sujeitos adultos de sua pesquisa (1,3% do total de dados). Nossos resultados parecem bastante expressivos, uma vez que se acredita que a fala adulta seja estabilizada. Uma possível explicação para esse fato pode estar ligada à orientação de que não iríamos, de forma alguma, avaliar ou julgar a qualidade de dicção ou de qualquer outro aspecto que pudesse fazer as participantes acreditarem que suas gravações passariam por julgamento qualitativo. Esse é um ponto importante a ser esclarecido durante a gravação de dados de fala, pois acreditamos que isso possa diminuir a tensão da situação e também a preocupação do informante em produzir palavras bem articuladas, o que pode acabar conferindo ao dado de fala certa artificialidade.

Outro ponto a ser levantado é a possível influência de palavras menos recorrentes no *corpus* como “pícolo”, “pômulo” e “cítara” (apesar de serem em pequeno número), que podem ter causado certa hesitação na fala, acarretando em um aumento de dados com a presença de eventos acústicos distintos daqueles conhecidos para a fala “típica”. Sugerimos que estudos futuros que venham a lidar com as especificidades acústicas atentem para a escolha de palavras mais conhecidas, para se evitar possíveis efeitos de estranhamento por parte dos informantes. É possível que, com a investigação de dados de fala espontânea, esse número seja ainda maior, pois, nesse tipo de fala, há um menor cuidado por parte dos sujeitos avaliados durante a coleta de dados.

Por fim, é importante salientar que a presença de um número tão expressivo de dados, que apresentam características distintas daquelas apresentadas por livros ou manuais de fonética como dado “normal ou típico” (24% dos nossos dados), ressalta o caráter variável e gradiente da língua que está em constante mudança. A fala não é estática, pois se estabelece dentro de um *continuum* que vai caracterizar cada um dos segmentos, que se organizam de tal forma a permitir que os falantes de uma determinada língua sejam capazes de se comunicarem. Pesquisas mostram que alterações graduais em determinada característica ou parâmetro vai, aos poucos, modificando ou não o julgamento perceptual do ouvinte (SNOEREN, 2005; BENKÍ, 2001,

dentre outros). Portanto, não é a presença ou ausência de determinada característica que vai indicar a realização fonética deste ou daquele segmento de fala, mas sim um conjunto de parâmetros que são utilizados em menor ou maior grau a depender de como esse segmento foi articulado. Acreditamos ser, por essa razão, que encontramos dados que apresentam características que divergem do que é apresentado pela literatura como segmento “típico”, mas que ainda assim se constituem como segmentos de fala do PB, passíveis de identificação por qualquer outro falante dessa língua, independentemente de diferenças extralinguísticas como sexo, idade, dialeto, dentre outros.

Esses dados confirmam, em primeiro lugar, o caráter gradiente da língua. O desvozeamento de segmentos vozeados, por exemplo, não ocorre de maneira semelhante para todos os segmentos vozeados. Essa especificidade pode ocorrer no início, meio ou ainda no final do segmento, com diferentes graus de desvozeamento. Da mesma forma ocorrem os *edge vibrations*, que podem ocupar uma parte insignificante da duração da oclusão do segmento não-vozeado ou ainda ocupar quase a totalidade da duração da oclusão.

Em segundo lugar, esse resultado expressivo demonstra o caráter variável da língua. Apesar de coletarmos dados de fala controlada, obtivemos um quarto de dados divergentes daqueles apontados pela literatura como “típicos” ou “padrões” esperados para a produção de fala, atestando que a língua apresenta nuances que vão a modificando. De acordo com Labov (2008),

Por mais que esses estudos sejam perspicazes e produtivos, nem por isso eles nos permitem chegar mais perto dos dados fundamentais da língua em uso. Há muitas questões abertas que simplesmente não conseguimos responder. (...). Para lidar com a *língua*, temos de olhar para os dados da fala cotidiana o mais perto e diretamente possível, e caracterizar seu relacionamento com as teorias gramaticais de modo mais acurado que pudermos, corrigindo e adequando a teoria para que ela se ajuste ao objeto visado. (LABOV, 2008, p. 235-36).

É necessário que tais fenômenos não sejam encarados como “erros”, passíveis de serem descartados das análises linguísticas. Eles fazem parte do dia-a-dia da língua e demonstram a riqueza e a

diversidade que constitui esse organismo que está em constante modificação (LABOV, 2008).

#### 4.5 CARACTERÍSTICAS ESPECTRAIS DO *BURST*

Como visto anteriormente, o *burst* é o correlato acústico da soltura da oclusão. De acordo com Cooper *et al.* (1952) e Halle, Hughes e Radley (1957), as plosivas bilabiais apresentam *burst* característico nas baixas frequências, entre 500-1500 Hz, enquanto as plosivas alveolares nas regiões de altas frequências (acima de 4000 Hz) e velares nas regiões de médias frequências (entre 1500-4000 Hz). Por apresentarem dispersão de energia nas regiões de frequência que coincidem com a das vogais, as plosivas velares podem sofrer influência da vogal que a sucede, a depender do tipo de vogal. De acordo com o exposto acima, colocamos a **Pergunta 6**: as medidas de características de *burst* diferem entre diferentes pontos de articulação?

As seções abaixo apresentam os resultados referentes a cada uma das medidas de *burst* separadamente.

##### 4.5.1 Pico espectral

A Tabela 19 apresenta os valores médios para o pico espectral de plosivas não-vozeadas.

Tabela 19 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para o pico espectral, juntamente com os resultados dos testes obtidos na comparação de pico espectral entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Pico espectral</b>	[p]	2039 (764)	323 – 2975	$U = 328$	$> 0,017$
	[t]	2888 (1087)	1045 – 5075		
	[p]	2039 (764)	323 – 2975	$U = 308$	$< 0,017$
	[k]	1734 (919)	535 – 4438		
	[t]	2888 (1087)	1045 – 5075	$U = 279$	$< 0,017$
	[k]	1734 (919)	535 – 4438		

O teste Kruskal-Wallis apontou diferenças significativas entre o pico espectral das plosivas não-vozeadas ( $H(2) = 37,15$ ,  $p < 0,001$ ). Testes Mann-Whitney (com correção Bonferroni) mostraram que há diferenças significativas entre a velar e as demais (Tabela 19), que

apresentou o menor valor. De acordo com os valores de *burst* assinalados por Halle *et al.* (1957), as bilabiais deveriam ter apresentado o menor valor, o que não ocorreu nos nossos dados, pois foi a consoante velar que apresentou valores significativamente menores se comparados aos valores exibidos pelas demais consoantes. As bilabiais ultrapassaram a faixa estabelecida pelos autores que deveria ficar entre 500-1500 Hz. As alveolares, apesar de apresentarem valor médio maior do que as velares, se estabeleceram numa faixa de frequência bem abaixo dos 4000 Hz que era o valor proposto pelos pesquisadores. As velares foram as únicas que se encaixaram na faixa de 1500-4000 Hz, a estabelecida por Halle *et al.* (1957).

Para as plosivas vozeadas, os valores encontrados ratificam parcialmente a ordem proposta na literatura ( $[b] < [g] < [d]$ ), uma vez que a alveolar é a que apresenta maior duração. Assim podemos afirmar que  $[d] > [b]$  e  $[d] > [g]$ , conforme a Tabela 20. (HALLE *et al.*, 1957; COOPER *et al.*, 1952; LIBERMAN *et al.*, 1952; JOHNSON, 2003, RUSSO; BEHLAU, 1993; dentre outros). Contudo, não podemos afirmar que  $[g]$  apresenta pico espectral com valores maiores do que os apresentados por  $[b]$ , já que não houve diferença significativa entre  $[b]$  e  $[g]$  (Tabela 20).

Tabela 20 – Valores de média de pico espectral (Hz), de desvio padrão e de faixa de distribuição de pico espectral (Hz), juntamente com os resultados obtidos na comparação de pico espectral entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Pico espectral	[b]	1189 (732)	323 – 2296	$U = 328$	<b>&lt; 0,017</b>
	[d]	2324 (1230)	408 – 4481		
	[b]	1189 (732)	323 – 2296	$U = 308$	<b>&gt; 0,017</b>
	[g]	1745 (1115)	514 – 4629		
	[d]	2324 (1230)	408 – 4481	$U = 279$	<b>&lt; 0,017</b>
	[g]	1745 (1115)	514 – 4629		

O teste Kruskal-Wallis revelou diferença significativa entre os valores de pico espectral das vozeadas ( $H(2) = 16,06$ ,  $p < 0,001$ ). Testes Mann-Whitney com correção Bonferroni mostraram que os valores do pico espectral da alveolar foram os únicos que se diferenciaram significativamente dos demais, conforme os resultados apresentados na Tabela 20.

Ademais, a bilabial e a velar se enquadraram dentro das faixas de concentração de energia previstas por Halle *et al.* (1957), de 500-1500 Hz para bilabiais e 1500-4000 Hz para velares. A alveolar, apesar de apresentar o valor de pico espectral mais elevado, esse valor ficou abaixo da categoria de 4000 Hz, valor estabelecido por esses pesquisadores.

#### 4.5.2 Centroide

No que tange o centroide, Forrest *et al.* (1988), ao analisarem dados de língua inglesa, observaram que esse parâmetro se diferenciava entre bilabiais e alveolares, mas que era similar entre alveolares e velares.

A análise do centroide em nossos dados, para as plosivas não-vozeadas, revelou valores menores para as bilabiais e maiores para as alveolares (Tabela 21).

Tabela 21 – Valores de média do centroide (Hz), de desvio padrão e de faixa de distribuição do centroide (Hz), juntamente com os resultados dos testes obtidos na comparação do centroide entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Centroide	[p]	1172 (726)	240 – 2706	Gabriel*	< 0,001
	[t]	2190 (891)	577 – 3866		
	[p]	1172 (726)	240 – 2706	Gabriel*	< 0,05
	[k]	1671 (633)	550 – 3086		
	[t]	2190 (891)	577 – 3866	Gabriel*	< 0,05
	[k]	1671 (633)	550 – 3086		

\*Teste *post hoc*

O teste Anova Unifatorial apontou para diferenças significativas no centroide para plosivas não-vozeadas ( $F(2,111) = 14,12p < 0,001$ ). O teste *post hoc* de Gabriel revelou que a diferença entre cada uma das comparações do centroide das plosivas não-vozeadas foi significativa (Tabela 21).

A análise de centroide para [t] e [k] em Berti e Marino (2011) apresentou valores de 1888 Hz para [ta] e 1704 para [ka], mas sem encontrar diferenças significativas. Cristofolini (2013) encontrou valores bastante diferentes para [p] e [t] (344 Hz e 289 Hz, respectivamente), mas encontrou valores semelhantes aos de Berti e Marino (2011) para a

velar (1699 Hz). O valor encontrado para [k] em nosso estudo (1734 Hz) se aproxima dos valores encontrados por Berti e Marino (2011).

Para as plosivas vozeadas, a Tabela 22 mostra que o valor do centroide da bilabial também foi inferior ao da alveolar e o valor da alveolar foi o mais elevado dentre as três plosivas vozeadas.

Tabela 22 – Valores de média do centroide (Hz), de desvio padrão e de faixa de distribuição do centroide (Hz), juntamente com os resultados dos testes obtidos na comparação do centroide entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Centroide</b>	[b]	896 (563)	222 – 2040	$U = 281$	<b>&lt; 0,017</b>
	[d]	1749 (1076)	204 – 3900		
	[b]	896 (563)	222 – 2040	$U = 246$	> 0,017
	[g]	1095 (581)	330 – 2139		
	[d]	1749 (1076)	204 – 3900	$U = 515$	<b>&lt; 0,017</b>
	[g]	1095 (581)	330 – 2139		

Encontramos diferenças significativas no centroide para as plosivas vozeadas, conforme revelado pelo teste Kruskal-Wallis ( $H(2) = 10,66$ ,  $p < 0,01$ ). Testes Mann-Whitney com correção Bonferroni apontaram para diferenças significativas entre a alveolar e as demais (Tabela 22).

Como o centroide reflete a média de concentração de energia no espectro de frequências de *burst*, nossos resultados corroboram os achados de estudos anteriores, que relatam que a média das bilabiais se concentra em baixas frequências, das alveolares em altas frequências e das velares se concentra em valores entre bilabiais e alveolares (YAVAS, 2011; JOHNSON, 2003; COOPER *et al.*; 1952, dentre outros).

#### 4.5.3 Variância

A variância foi uma característica que não se mostrou relevante na diferenciação do ponto de articulação nos estudos de Forrest *et al.* (1988) e Berti e Marino (2011).

Para nossa pesquisa, o teste Kruskal-Wallis apontou para diferenças na variância entre as três plosivas não-vozeadas ( $H(2) = 22,48$ ,  $p < 0,001$ ) e entre as três vozeadas ( $H(2) = 10,66$ ,  $p < 0,01$ ). Conforme pode ser observado nas Tabela 23 e Tabela 24, os testes Mann-Whitney (com correção Bonferroni) revelaram que ela foi

significativa somente na comparação entre plosivas alveolares e velares, tanto entre as não-vozeadas ( $U = 322$ ,  $p < 0,017$ ), quanto entre as vozeadas ( $U = 396,5$ ,  $p < 0,017$ ).

Tabela 23 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para a variância, juntamente com os resultados dos testes obtidos na comparação da variância entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$	DP	Valor do teste	Valor de p
Variância	[p]	2023937	1721913	$U = 300$	$> 0,017$
	[t]	2390699	1005757		
	[p]	2023937	1721913	$U = 345$	$> 0,017$
	[k]	1226259	869058		
	[t]	2390699	1005757	$U = 322$	$< 0,017$
	[k]	1226259	869058		

Tabela 24 – Valores de média (Hz), de desvio padrão e de faixa de distribuição (Hz) para a variância, juntamente com os resultados os testes obtidos na comparação da variância entre as plosivas vozeadas

	Plosiva	$\bar{X}$	DP	Valor do teste	Valor de p
Variância	[b]	1394517	954478	$U = 292$	$> 0,017$
	[d]	2360246	1491311		
	[b]	1394517	954478	$U = 245$	$> 0,017$
	[g]	1222924	647465		
	[d]	2360246	1491311	$U = 396$	$< 0,017$
	[g]	1222924	647465		

Como outras pesquisas já apontaram e foi confirmado por nossos resultados, a variância parece não ser um parâmetro decisivo na diferenciação entre todos os pontos de articulação para o *burst*. Nossos resultados indicam que esse parâmetro pode talvez servir como uma pista a mais na identificação de plosivas alveolares.

#### 4.5.4 Assimetria

Conforme explicado na Seção 3.2.2, a assimetria negativa é aquela que apresenta concentração nas altas frequências. De acordo com os resultados mostrados nas Tabela 25 e Tabela 26, as alveolares [t] e [d] foram as que apresentaram a média mais baixa e, portanto, uma tendência a apresentar valores mais concentrados nas altas frequências,

apesar de não apresentarem valores médios negativos. As bilabiais foram aquelas que apresentaram a média mais alta, indicando concentração em frequências mais baixas.

Tabela 25 – Valores de média, de desvio padrão e de faixa de distribuição para a assimetria, juntamente com os resultados dos testes obtidos na comparação da assimetria entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Assimetria	[p]	5,9 (4)	1,4 – 16,5	$U = 116$	<b>&lt; 0,017</b>
	[t]	1,7 (1,4)	-1,1 – 5,1		
	[p]	5,9 (4)	1,4 – 16,5	$U = 368$	<b>&gt; 0,017</b>
	[k]	4,4 (2)	1,4 – 10,6		
	[t]	1,7 (1)	-1,1 – 5,1	$U = 232$	<b>&lt; 0,017</b>
	[k]	4,4 (2)	1,4 – 10,6		

O teste Kruskal-Wallis revelou diferença significativa na assimetria entre as plosivas não-vozeadas ( $H(2) = 37,54$ ,  $p < 0,001$ ) e vozeadas ( $H(2) = 28,18$ ,  $p < 0,001$ ). Conforme pode ser observado pelos resultados do teste Mann-Whitney (com correção Bonferroni) aplicado a cada uma das comparações entre as plosivas não-vozeadas (Tabela 25), a assimetria é significativamente diferente apenas entre a alveolar e as demais. O mesmo resultado foi revelado pelos testes estatísticos aplicados às plosivas vozeadas, conforme pode ser observado na Tabela 26.

Tabela 26 – Valores de média, de desvio padrão e de faixa de distribuição para a assimetria, juntamente com os resultados dos testes obtidos na comparação da assimetria entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Assimetria	[b]	6,0 (3)	1,8 – 12,7	$U = 292$	<b>&lt; 0,017</b>
	[d]	2,6 (2)	-0,6 – 7,0		
	[b]	6,0 (3)	1,8 – 12,7	$U = 245$	<b>&gt; 0,017</b>
	[g]	5,8 (4)	2,0 – 13,8		
	[d]	2,6 (2)	-0,6 – 7,0	$U = 396$	<b>&lt; 0,017</b>
	[g]	5,8 (4)	2,0 – 13,8		

Apesar de [t] apresentar valor médio de assimetria mais elevado do que para [k] no estudo de Berti e Marino (2011), a diferença foi significativa apenas quando eram seguidas pela vogal [u], mas não-significativa quando estavam diante da vogal [a]. Cristofolini (2013)

encontrou valores bastante distintos dos nossos e dos de Berti e Marino (2011), pois as alveolares foram as que apresentaram maior valor de assimetria.

Em vista desses resultados, acreditamos que a assimetria é ainda um ponto controverso para explicar a diferença entre os pontos de articulação das plosivas. No entanto, os resultados obtidos pela presente pesquisa confirmam que as alveolares tendem a apresentar concentração de energia nas altas frequências e as bilabiais nas baixas frequências (indo ao encontro dos resultados de pico espectral e centroide).

#### 4.5.5 Curtose

Por fim, as médias dos valores de curtose são apresentadas na Tabela 27 para as não-vozeadas e na Tabela 28 para as vozeadas.

Tabela 27 – Valores de média de curtose, de desvio padrão e de faixa de distribuição de curtose, juntamente com os resultados dos testes obtidos na comparação da curtose entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Curtose	[p]	59 (62)	5 – 232	$U = 115$	<b>&lt; 0,017</b>
	[t]	10 (9)	-1 – 37		
	[p]	59 (62)	5 – 232	$U = 354$	<b>&gt; 0,017</b>
	[k]	55 (55)	7 – 202		
	[t]	10 (9)	-1 – 37	$U = 136$	<b>&lt; 0,017</b>
	[k]	55 (55)	7 – 202		

Observamos diferença significativa na curtose entre as três plosivas não-vozeadas ( $H(2) = 42,36$ ,  $p < 0,001$ ) e entre as três vozeadas ( $H(2) = 40,22$ ,  $p < 0,001$ ). Através de testes Mann-Whitney (com correção Bonferroni), encontramos valores significativos somente entre as alveolares e as demais, tanto entre as não-vozeadas como entre as vozeadas. Esse resultado corrobora o achado de Berti e Marino (2011). Assim, as alveolares apresentaram valores de curtose significativamente mais baixos se comparados aos valores referentes às demais plosivas, de acordo com os resultados exibidos na Tabela 28. **Erro! Fonte de referência não encontrada.**

Tabela 28 – Valores de média de curtose, de desvio padrão e de faixa de distribuição de curtose, juntamente com os resultados dos testes obtidos na comparação da curtose entre as plosivas vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
Curtose	[b]	90 (95)	3 – 298	$U = 50$	< <b>0,017</b>
	[d]	11 (7)	-0,1 – 32		
	[b]	90 (95)	3 – 298	$U = 148$	> 0,017
	[g]	53 (46)	8 – 158		
	[d]	11 (7)	-0,1 – 32	$U = 122$	< <b>0,017</b>
	[g]	53 (46)	8 – 158		

A curtose é a medida que aponta se o espectro de frequência apresenta múltiplos picos (curtose positiva) ou se o espectro é mais achatado, com maior distribuição de valores (curtose negativa). Para nossos dados, a curtose foi sempre positiva, indicando que os espectros de todos os pontos de articulação apresentam mais pontos de pico de energia.

A partir dos resultados apresentados referentes às características espectrais de *burst*, podemos concluir que a **Hipótese 6**, que sugeria que as medidas do pico, do centroide, da variância, da assimetria e de curtose se diferenciariam com relação ao ponto de articulação, pode ser parcialmente aceita, pois o único parâmetro que diferenciou todas as plosivas entre si foi o centroide, para as plosivas não-vozeadas. Para as demais características de *burst* não houve conformidade para diferenciar todas as plosivas entre si, mas notamos uma tendência de diferenciarem as plosivas alveolares das demais.

Em suma, apesar de não encontrarmos resultados significativos para todas as características de *burst*, acreditamos que nossos resultados demonstram conformidade com aspectos já abordados por outros estudos, como a tendência das bilabiais de apresentarem concentração de energia em regiões mais baixas do que as demais consoantes, das velares em regiões de médias frequências e das alveolares em regiões um pouco mais elevadas do que as demais. Além disso, nossa pesquisa

investigou uma variedade dialetal do PB distinta daquelas apresentadas por outros estudos<sup>52</sup>.

#### 4.5.6 Relação entre *burst* e contexto vocálico adjacente

No estudo de Halle, Hughes e Radley (1957) com segmentos plosivos da língua inglesa, os pesquisadores observaram atenuação nos valores de frequência do pico espectral das plosivas velares quando elas eram seguidas por vogais posteriores. Enquanto o pico espectral das velares seguidas por vogais anteriores se estabeleceu entre 2000-4000 Hz (vogais anteriores do inglês apresentam F2 entre 1200-1500Hz), o das velares seguidas por vogais posteriores ocorreu em regiões mais baixas, pois o F2 dessas vogais fica abaixo de 1200Hz. Em vista disso, a **Pergunta 7** levanta a questão da possível influência da vogal seguinte nas características espectrais das plosivas: as medidas de *burst* podem sofrer alteração de acordo com o tipo de vogal que segue a plosiva?

Nosso *corpus* foi composto por plosivas seguidas de [ɐ] e [ʊ]. A Tabela 29 apresenta os valores médios, o desvio padrão e os resultados obtidos com testes Mann-Whitney para cada característica espectral, de acordo com a vogal que segue cada tipo de plosiva.

---

<sup>52</sup> Ficker (2003) e Barbosa (1996) observaram o dialeto paulistano; Klein (1999) e Cristofolini (2008;2013) coletaram dados com informantes na região de Florianópolis. Todas as pesquisas utilizadas como fonte de referência avaliaram dialetos distintos do avaliado por nós. Portanto, é possível que as diferenças possam ou não se aplicar ao PB ou ainda possam estar relacionadas com questões dialetais, as quais não observamos neste estudo.

Tabela 29 – Valores médios e de desvio padrão para as características espectrais de *burst*, de acordo com a vogal seguinte a cada tipo de plosiva, juntamente com os resultados dos testes obtidos na comparação entre o contexto vocálico, para cada tipo de plosiva

	Plosiva	Vogal	$\bar{X}$	DP	Valor do teste	Valor de p
<b>Pico</b>	Bil.	[e]	1402	916	$U = 68$	$> 0,05$
		[o]	1676	870		
	Alv.	[e]	2193	245	$t(63) = -4,56$	$< 0,001$
		[o]	3306	1225		
	Velar	[e]	1839	138	$U = 210$	$< 0,01$
		[o]	1744	1624		
<b>Centroide</b>	Bil.	[e]	719	341	$U = 22$	$< 0,001$
		[o]	1337	408		
	Alv.	[e]	1837	623	$t(63) = -3,1$	$< 0,01$
		[o]	2465	935		
	Velar	[e]	1684	510	$U = 104$	$< 0,001$
		[o]	924	355		
<b>Variância</b>	Bil.	[e]	1333754	927002	$U = 26$	$< 0,01$
		[o]	3192199	1565774		
	Alv.	[e]	2906755	1057445	$t(63) = 1,5$	$> 0,05$
		[o]	2514267	978887		
	Velar	[e]	1286984	710234	$U = 308$	$> 0,05$
		[o]	1177317	1000007		
<b>Assimetria</b>	Bil.	[e]	6,3	2,9	$U = 43$	$> 0,05$
		[o]	3,8	0,5		
	Alv.	[e]	2,4	0,8	$t(63) = 6,4$	$< 0,001$
		[o]	1,0	0,9		
	Velar	[e]	3,8	1,8	$U = 182$	$< 0,001$
		[o]	6,4	2,8		
<b>Curtose</b>	Bil.	[e]	86	71	$U = 49$	$> 0,05$

	[ʊ]	28	8		
Alv.	[e]	13	6	$t(63) =$ 3,8	< 0,001
	[u]	7	5		
Velar	[e]	40	33	$U = 250$	< 0,05
	[u]	92	70		

Observando a Tabela 29, sob o ponto de vista do pico espectral, os testes estatísticos mostraram que há atenuação significativa da frequência nas alveolares seguidas pela vogal [e]. Esse resultado corrobora os achados de Berti e Marino (2011) para o PB. Já para as velares, houve atenuação significativa do pico espectral quando a vogal posterior [u] seguia a plosiva. A bilabial não apresentou atenuação significativa para nenhum dos contextos vocálicos analisados.

De forma geral, a análise estatística apontou para diferenças significativas na maioria das comparações entre o contexto vocálico para cada um dos momentos espectrais das plosivas, em especial para as alveolares e velares. Para as bilabiais, os testes só apontaram diferenças significativas entre os contextos vocálicos somente para a variância e centroide. As diferenças significativas apontadas pelos testes estatísticos, no entanto, ora apresentavam valores mais atenuados quando a plosiva era seguida pela vogal anterior, ora pela vogal posterior. Assim, acreditamos que a influência contexto vocálico no *burst* das plosivas do PB pode ser um fator relevante, apesar de não sistemático.

Cabe ainda lembrar que, em um estudo sobre os padrões espectrais de plosivas diante de vogais do inglês, Blumstein e Stevens (1979) reportaram que, apesar de o contexto vocálico adjacente exercer certa influência sobre os padrões espectrais das plosivas, ele opera de maneira limitada. Em face dos nossos resultados, é possível que o mesmo ocorra para as plosivas do PB. Para fazer inferências mais contundentes, seria preciso conduzir uma pesquisa com um volume maior de dados e de participantes, com as demais vogais do PB, o que não foi possível neste estudo.

Desse modo, a **Hipótese 7** que colocava que existiriam diferenças entre as características do *burst* das plosivas seguidas por diferentes vogais para o PB e apontava que plosivas seguidas por vogais posteriores poderiam apresentar menores valores de frequência (valores mais atenuados) do que aquelas seguidas por vogais anteriores, pode ser parcialmente aceita, uma vez que a vogal posterior atenuou

significativamente os valores de frequência de pico espectral e de centroide no caso das plosivas velares.

## 4.6 VOT E ASPIRAÇÃO

O VOT é, sem dúvida, o parâmetro mais utilizado pelos pesquisadores quando se quer observar o comportamento das plosivas nas mais diversas línguas.

De forma geral, estudos têm mostrado que existe considerável diferença entre o VOT de plosivas homorgânicas, bem como entre as plosivas de uma mesma categoria (ALVES, 2011; ALVES et al., 2008; CHO; LADEFOGED, 1999; CLARK; YALLOP, 1995; FICKER, 2003; KLEIN, 1999; LADEFOGED; JOHNSON, 2011; LISKER; ABRAMSON, 1964b). Algumas dessas diferenças entre as plosivas de uma mesma língua ou de línguas distintas são atribuídas, como vimos anteriormente, à presença ou não do parâmetro aspiração.

### 4.6.1 VOT e vozeamento

Tendo em vista as informações acima descritas, a **Pergunta 8** questiona: há diferenças entre o VOT de plosivas vozeadas e não-vozeadas?

Os resultados dispostos na Tabela 30 mostram que o valor médio de VOT das vozeadas foi (-)82 ms<sup>53</sup>, o que o enquadra na categoria de pré-sonorização; enquanto para as não-vozeadas, o valor médio foi 31 ms, o que as enquadra na categoria de retardo curto.

Tabela 30 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas vozeadas e não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
VOT	Não-voz	31 (19)	6 – 103	$U = 44590$	< <b>0,001</b>
	Voz	(-)82 (23)	(-)166 – (-) 29		

<sup>53</sup> Para a análise estatística, o sinal negativo do VOT das plosivas vozeadas, que denota que o vozeamento é anterior à soltura da oclusão. Levamos em consideração apenas o seu valor em módulo. Ex.: Valor = -75,56 ms, valor em módulo = |75,56 |

<b>VOT</b>	Não-voz	6 (4)	1 – 30	$U = 55300$	<b>&lt; 0,001</b>
<b>relativo</b>	Voz	17 (5)	5 – 36		

Os testes Mann-Whitney mostraram que a diferença entre esses valores é significativa, indicando que, como exposto nas pesquisas acima citadas, o VOT pode ser um parâmetro que distingue esses dois tipos de plosivas.

Realizamos também a análise do comportamento do VOT entre as plosivas homorgânicas. A Tabela 31 mostra os resultados obtidos por essa investigação.

Tabela 31 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa de VOT entre os pares homorgânicos

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>VOT</b>	[p]	20 (8)	6 – 54	$U = 31$	<b>&lt; 0,001</b>
	[b]	(-)92 (21)	(-)166 – (-41)		
	[t]	21 (7)	8 – 47	$U = 93$	<b>&lt; 0,001</b>
	[d]	(-)83 (22)	(-)155 – (-32)		
	[k]	52 (16)	21 – 103	$U = 19329$	<b>&lt; 0,001</b>
	[g]	(-)70 (19)	(-)134 – (-29)		
<b>VOT relativo</b>	[p]	4 (1)	1 – 13	$U = 96$	<b>&lt; 0,001</b>
	[b]	18 (5)	8 – 34		
	[t]	4 (2)	1 – 10	$U = 303$	<b>&lt; 0,001</b>
	[d]	17 (6)	5 – 36		
	[k]	10 (4)	4 – 24	$U = 21863$	<b>&lt; 0,001</b>
	[g]	14 (5)	5 – 31		

Em todas as comparações entre plosivas homorgânicas, o valor modular das vozeadas foi maior do que o das não-vozeadas. Testes Mann-Whitney revelaram que a diferença entre cada um dos pares é significativa (Tabela 31), confirmando o resultado da análise geral entre plosivas não-vozeadas e vozeadas. Encontramos resultados que inserem todas as não-vozeadas na categoria de retardo curto (LISKER; ABRAMSON, 1964a). As plosivas [p] e [t], por apresentarem valores médios de VOT menores do que 35 ms, podem ser classificadas como

não-aspiradas e o VOT de [k], por estar entre 35 e 55 ms, insere essa plosiva na categoria de levemente aspirada, de acordo com a classificação de Cho e Ladefoged (1999) apresentada, anteriormente, na Tabela 4. Às plosivas vozeadas foram atribuídos valores negativos, pois o vozeamento foi anterior à soltura da oclusão, sendo assim inseridas na categoria de pré-sonorização (LISKER; ABRAMSON, 1964).

Esses resultados confirmam a **Hipótese 8**, que previa que plosivas não-vozeadas seriam classificadas como retardo curto, sem aspiração ou com leve aspiração, e que as vozeadas seriam classificadas como pré-sonorização, como previsto pela literatura para o PB.

#### 4.6.2 VOT e ponto de articulação

No que tange a comparação entre os diferentes pontos de articulação, observemos, em primeiro lugar, as plosivas não-vozeadas, conforme questiona a **Pergunta 9.1**: há diferenças entre os valores de VOT entre as plosivas não-vozeadas?

De acordo com os resultados mostrados na Tabela 32, os valores médios de VOT das plosivas não-vozeadas bilabiais e alveolares foram bastante semelhantes e a média da velar se distanciou de forma acentuada das demais. O resultado do teste Kruskal-Wallis apontou para diferenças significativas na duração absoluta,  $H(2) = 614,60$ ,  $p < 0,001$  e relativa do VOT,  $H(2) = 595,04$ ,  $p < 0,001$ , entre as plosivas não-vozeadas. Testes Mann-Whitney (com correção Bonferroni) foram aplicados a cada uma das comparações e os resultados são mostrados pela Tabela 32 **Erro! Fonte de referência não encontrada.**

Tabela 32 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas não-vozeadas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
VOT	[p]	20 (8)	6 – 54	$U = 50653$	$> 0,017$
	[t]	21 (7)	8 – 47		
	[p]	20 (8)	6 – 54	$U = 2820$	$< 0,017$
	[k]	52 (16)	21 – 103		
	[t]	21 (7)	8 – 47	$U = 2064$	$< 0,017$
	[k]	52 (16)	21 – 103		
VOT relativo	[p]	4 (1)	1 – 13	$U = 50795$	$> 0,17$
	[t]	4 (2)	1 – 10		

[p]	4 (1)	1 – 13	$U = 3976$	$< 0,017$
[k]	10 (4)	4 – 24		
[t]	4 (2)	1 – 10	$U = 3441$	$< 0,017$
[k]	10 (4)	4 – 24		

A comparação feita através de testes estatísticos mostrou que não há diferença significativa entre [p] e [t], tanto para duração absoluta quanto para relativa de VOT. Com relação à comparação da velar com as demais, testes estatísticos apontaram para diferenças significativas entre os valores de VOT dessa plosiva e das demais não-vozeadas.

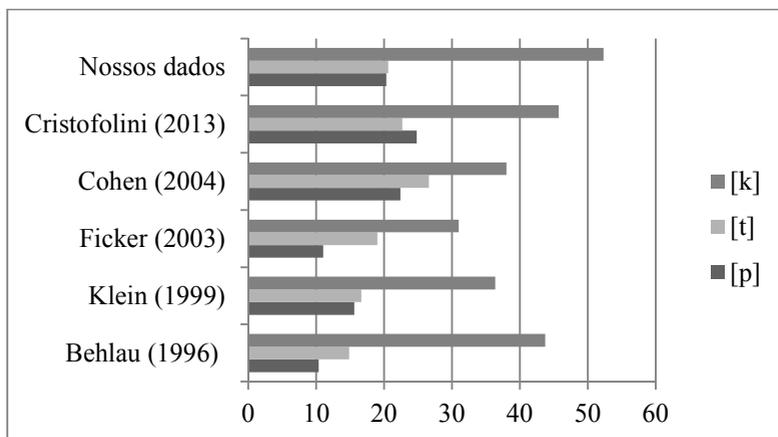
É possível observar também, a partir dos valores médios de VOT para as não-vozeadas (Tabela 32), que esse parâmetro segue parcialmente a ordem [p] < [t] < [k], como encontrada por outros pesquisadores para o PB (ALVES, 2011; ALVES et al., 2008; BEHLAU, 1986; COHEN, 2004; FICKER, 2003), já que nossos dados mostraram que ([p] e [t]) < [k], pois se não há diferenças significativas entre bilabiais e alveolares, não podemos considerar que os valores de VOT de [p] sejam menores do que os valores referentes à consoante [t].

As plosivas [p] e [t] enquadram-se na categoria não-aspirada, proposta por Cho e Ladefoged (1999)<sup>54</sup>, uma vez que nenhum dos valores ultrapassou o limite para essa categoria, que se estabelece em 35 ms. A plosiva velar, por ultrapassar esse valor, adentra a categoria de levemente aspirada, com valores entre 35-55 ms. Esse resultado corrobora os achados de Klein (1999), Ficker (2003); Alves et al. (2008) e Alves (2011), dentre outros.

O Gráfico 2 apresenta uma comparação entre os nossos resultados e outros encontrados em pesquisas anteriores para o PB.

Gráfico 2 - Resultados de média de VOT (em ms) em Cristofolini (2013), Cohen (2006), Ficker (2003), Klein (1999), Behlau (1996) e em nossos dados para plosivas não-vozeadas

<sup>54</sup> Ver classificação na Tabela 4, no Capítulo 2, Seção 2.2.2.1.



A partir do Gráfico 2, podemos notar que os valores encontrados pelo nosso estudo para plosivas não-vozeadas são semelhantes aos encontrados por outros pesquisadores, com relação às plosivas bilabiais e alveolares, e se mostrou ligeiramente mais elevado com relação às velares. Nota-se, também, que os valores médios para as plosivas velares, na maioria dos estudos, ultrapassou o limite de 35ms e podem, então, ser classificadas como levemente aspiradas. Os valores médios de [p] e [t], para a maior parte das pesquisas, seguiram o padrão  $[p] < [t]$  – podemos notar inversão nos valores somente no estudo de Cristofolini (2013) – e também são bastante similares entre si. Assim, podemos afirmar que nossos resultados corroboram os estudos acima no que diz respeito às plosivas não-vozeadas do PB.

Esses resultados indicam que a **Hipótese 9.1** pode ser aceita, pois previa que as alveolares e bilabiais seriam não-aspiradas e que as velares apresentariam VOT mais longo do que as demais, podendo ser inserida na categoria de leve aspiração, assim como em Klein (1999), Alves *et al.* (2008), Alves (2011), Cristofolini (2013), dentre outros .

A respeito das plosivas vozeadas, a **Pergunta 9.2** questiona: há diferenças entre os valores de VOT entre as plosivas vozeadas?

A Tabela 33 apresenta os valores médios de duração absoluta e relativa do VOT, juntamente com os valores dos testes estatísticos e de significância, observados na comparação de duração absoluta e relativa do VOT entre as plosivas vozeadas.

Tabela 33 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas vozeadas

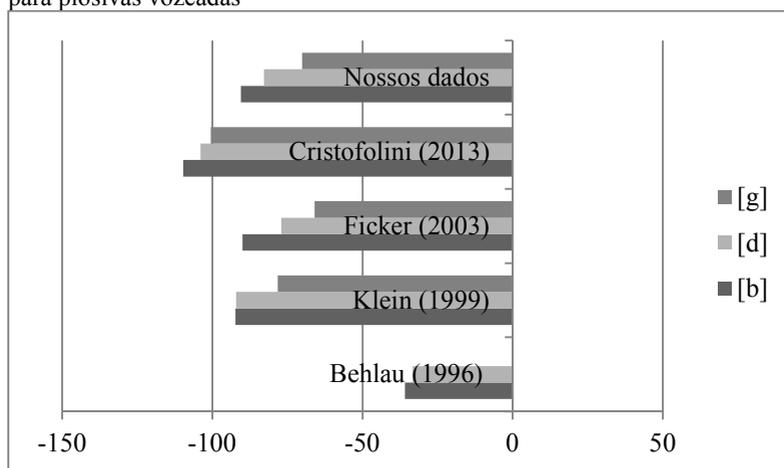
	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
VOT	[b]	(-)92 (20)	(-)166 – (-)41	$U = 42335$	< <b>0,017</b>
	[d]	(-)83 (22)	(-)155 – (-)32		
	[b]	(-)92 (20)	(-)166 – (-)41	$U = 18184$	< <b>0,017</b>
	[g]	(-)70 (19)	(-)134 – (-)29		
	[d]	(-)83 (22)	(-)155 – (-)32	$U = 27590$	< <b>0,017</b>
	[g]	(-)70 (19)	(-)134 – (-)29		
VOT Relativo	[b]	18 (5)	8 – 34	$U = 45360$	< <b>0,017</b>
	[d]	17 (6)	5 – 36		
	[b]	18 (5)	8 – 34	$U = 21682$	< <b>0,017</b>
	[g]	14 (5)	5 – 31		
	[d]	17 (6)	5 – 36	$U = 29635$	< <b>0,017</b>
	[g]	14 (5)	5 – 31		

Todos os valores mostrados na Tabela 33 inserem as plosivas vozeadas do nosso estudo na categoria de pré-sonorização. O teste Kruskal-Wallis apontou para diferenças significativas entre as plosivas vozeadas nos valores de duração do VOT,  $H(2) = 125,83$ ,  $p < 0,001$ , e de duração relativa do VOT em função da duração da palavra,  $H(2) = 87,06$ ,  $p < 0,001$ . Em sequência, testes Mann-Whitney com correção Bonferroni foram aplicados a cada uma das comparações entre as plosivas vozeadas. Os resultados apresentados na Tabela 33 mostram

que os valores médios de VOT entre as plosivas vozeadas diferem-se significativamente e, portanto, corroboram resultados encontrados por estudos anteriores (BEHLAU, 1996 *apud* FICKER, 2003; CRISTOFOLINI, 2013; FICKER, 2003; KLEIN, 1999). Pode-se observar, através da Tabela 33, que as bilabiais foram aquelas que apresentaram o maior valor médio de VOT e que as velares apresentaram o menor, seguindo, assim, a ordem [b] > [d] > [g]. Desta forma, a **Hipótese 9.2** que previa que os valores de VOT para as plosivas vozeadas se dariam na ordem [b] > [d] > [g] e que eles teriam valores negativos por iniciarem o vozeamento antes da soltura da oclusão, pode ser confirmada, visto que todas essas plosivas foram consideradas pré-sonorização e que [b] mostrou o maior valor de VOT (tanto médio como relativo) e que [g] apresentou o menor valor de VOT.

O Gráfico 3 apresenta uma comparação entre os valores médios encontrados em nosso estudo e em outras pesquisas para as plosivas vozeadas do PB.

Gráfico 3 – Resultados de média de VOT (em ms) em Cristofolini (2013), Cohen (2006), Ficker (2003), Klein (1999), Behlau (1996) e em nossos dados, para plosivas vozeadas



Nota-se, pelo Gráfico 3, que nossos resultados são bastante similares aos encontrados por outros pesquisadores para o PB. Todos os valores dos estudos apresentados no Gráfico 3 demonstram que as

plosivas vozeadas dessa língua estão inseridas na categoria de pré-sonorização.

### 4.6.3 Aspiração

Outra característica importante que pode estar presente na produção de uma plosiva é a aspiração. Estudos anteriores apontam que as plosivas velares do PB podem apresentar um leve grau de aspiração que as diferenciam das demais plosivas do PB (KLEIN, 1999; COHEN, 2004; ALVES *et al*, 2008; ALVES, 2011; CRISTOFOLINI, 2013, dentre outros). Assim, a **Pergunta 10** questiona: há diferenças entre o VOT das plosivas que apresentaram aspiração e o VOT das que não apresentaram?

Conforme visto acima, a aspiração ocorre, eventualmente, para as plosivas não-vozeadas do PB. Dessa forma, analisaremos a seguir as diferenças entre as plosivas não-aspiradas (com menos de 35ms de duração de VOT) e com aspiração (acima de 35 ms). Nosso objetivo é observar, especialmente, se as categorias aspiradas e não-aspiradas são bem definidas, isto é, se elas se afastam da fronteira definida por Cho e Ladefoged (1999) de 35ms ou se elas se aproximam dessa fronteira, mostrando que esse limite seria um valor limite, mas que mostra que há uma região *fuzzy* (ALBANO, 2001) entre elas e indicando que as fronteiras são o que Albano (2001) chama de fronteiras vagas.

A Tabela 34 apresenta os diferentes percentuais de ocorrências entre essas plosivas nos nossos dados.

Tabela 34 – Percentual de dados com aspiração dentre as plosivas não-vozeadas

	Total	Não-aspiradas	Com aspiração
[p]	396 (100%)	376 (95%)	20 (5%)
[t]	372 (100%)	348 (94%)	24 (6%)
[k]	402 (100%)	53 (13%)	349 (87%)
<b>Total</b>	<b>1170 (100%)</b>	<b>777 (66%)</b>	<b>393 (34%)</b>

De forma semelhante a Alves (2011) que encontrou 82% de dados de plosivas velares inseridos na categoria de leve aspiração, encontramos para a mesma plosiva, 87% de casos em que houve aspiração, dos quais 67% foram classificados como levemente aspirados (valores entre 35-55 ms) e 33% com aspiração (valores entre 55-95 ms).

Os valores apresentados na Tabela 35 mostram a diferença entre as médias encontradas para cada plosiva, separando-as em

aspiradas ou não-aspiradas e os respectivos resultados dos testes estatísticos.

Tabela 35 – Valores de média de duração absoluta (ms) e relativa (%), de desvio padrão e de faixa de distribuição de duração absoluta (ms) e relativa (%) para o VOT, de acordo com a presença ou não de aspiração, juntamente com os resultados dos testes obtidos na comparação de duração absoluta e relativa do VOT entre as plosivas sem e com aspiração

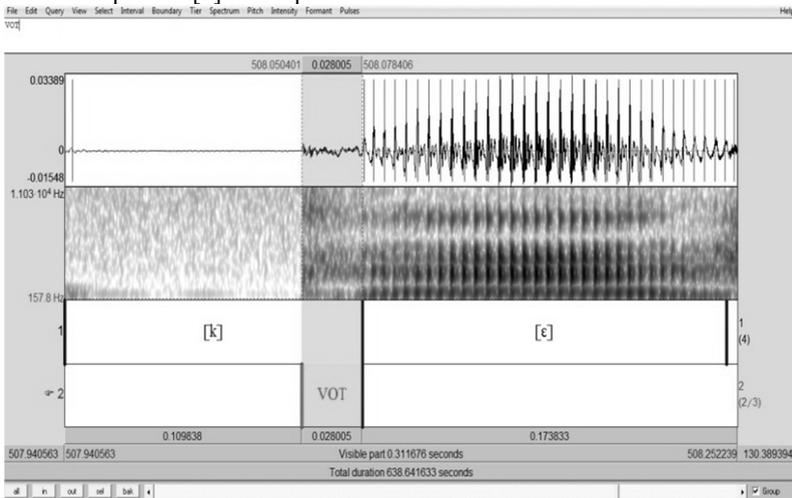
	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
VOT	[p]	19 (7)	6 – 35	$U = 22$	< 0,001
	[p <sup>h</sup> ]	41 (6)	35 – 54		
	[t]	20 (6)	8 – 35	$U = 42$	< 0,001
	[t <sup>h</sup> ]	41 (3)	35 – 47		
	[k]	30 (4)	21 – 35	$U = 38$	< 0,001
	[k <sup>h</sup> ]	55 (15)	35 – 103		
VOT relativo	[p]	4 (1)	1 – 9	$U = 270$	< 0,001
	[p <sup>h</sup> ]	8 (2)	5 – 13		
	[t]	4 (1)	1 – 7	$U = 162$	< 0,001
	[t <sup>h</sup> ]	7 (1)	6 – 8		
	[k]	6 (1)	4 – 10	$U = 812$	< 0,001
	[k <sup>h</sup> ]	11 (4)	5 – 24		

Esses resultados demonstram que a diferença entre os valores das plosivas com e sem aspiração são significativos, podendo assim confirmar a **Hipótese 10**, que colocava que a diferença entre os valores de plosivas com e sem aspiração seria significativo. No entanto, isso não significa que essa diferença seja distintiva (em termos fonológicos), uma vez que, para o PB, não há distinção entre plosivas com ou sem aspiração. Cabe ainda ressaltar que a média encontrada para cada uma das plosivas não-vozeadas com aspiração encontra-se dentro da

categoria de levemente aspirada ou ainda na categoria aspirada (somente plosivas velares).

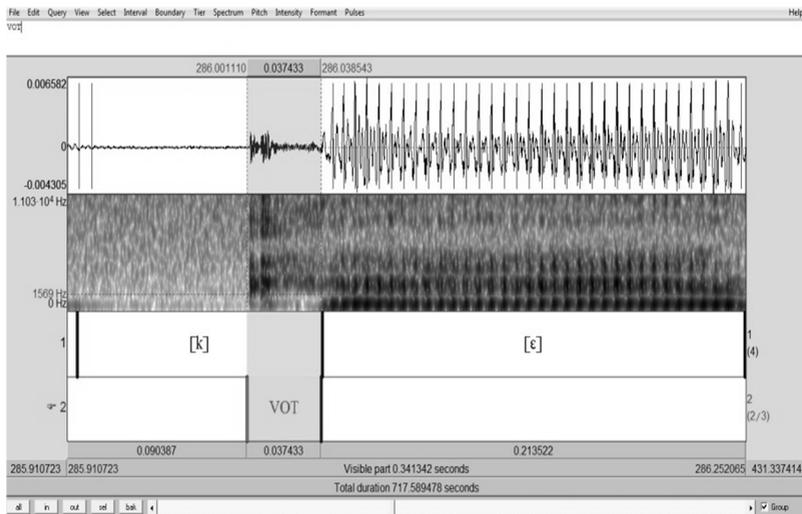
Pode-se observar a diferença do VOT de uma plosiva com ou sem aspiração através da forma de onda, na qual é possível visualizar um certo ruído para aquelas plosivas com a aspiração, ou ainda a partir do espectrograma, que apresenta uma região mais acinzentada, indicativa de ruído, para a plosiva com aspiração, conforme podemos visualizar através das Figura 30 e Figura 31. A Figura 30 mostra, em destaque, o VOT de [k] com 28 ms, uma plosiva não-aspirada, enquanto a Figura 31 mostra, em destaque, o VOT da plosiva [k] levemente aspirada, com 37 ms.

Figura 30 - Forma de onda e espectrograma da palavra "quéchua". Em destaque: o VOT da plosiva [k] não-aspirada



Fonte: Nossos dados

Figura 31 - Forma de onda e espectrograma da palavra "quero". Em destaque: o VOT da plosiva [k] levemente aspirada



Fonte: Nossos dados

Como já apontado por estudos anteriores (ALVES, 2011; ALVES et al., 2008; KLEIN, 1999), a presença da aspiração no PB não depende do tipo de vogal que a segue. Encontramos plosivas velares levemente aspiradas ou aspiradas diante dos mais variados contextos vocálicos.

Por fim, queremos ressaltar que os valores médios aqui encontrados parecem apontar para categorias distintas que podem definir que as plosivas apresentam leve aspiração ou que não apresentam aspiração. No entanto, encontramos dados que apresentaram valores de VOT de 34,99ms ou de 35,01ms, por exemplo. Qual a diferença entre essas plosivas em termos de aspiração? Como um ouvinte as identifica? Essas questões nos levam à reflexão sobre a definição de fronteiras ou limites que na verdade não são definitivos. Essas regiões conhecidas como “fuzzy” ou “nebulosas” precisam ser observadas e descritas, uma vez que demonstram a gradiência da língua (ALBANO, 2001). Apesar de a aspiração não ser distintiva no PB, é preciso analisar como tais nuances são produzidas e se o ouvinte é capaz de distingui-las, para que possamos reavaliar como podem participar efetivamente da caracterização dos segmentos plosivos.

Essa reflexão se mostra importante também em estudos com percepção de fala, quando se busca investigar os limiares acústicos entre um ou outro segmento. De acordo com Albano (2001),

(...) o vago pode ser conhecido, mapeado, modelado com extremo rigor. O primeiro passo para tanto é reconhecer limites, fronteiras objetivamente indistintas, vagas (p. 11).

Acretimos, assim, que é possível se identificar características acústicas que podem definir este ou aquele segmento, mas sem esquecer que essas características operam dentro de um *continuum* de fala (que, por sua vez, demonstram a gradiência nos gestos articulatórios envolvidos na produção dos segmentos) e que atuam em conjunto durante a produção dos segmentos de fala.

## 4.7 TRANSIÇÃO FORMÂNTICA

A região em que há a transição entre a consoante plosiva e a vogal, como descrito na seção 2.2.4, pode ser uma pista importante na distinção entre plosivas vozeadas e não-vozeadas, através da transição de F1, e entre diferentes pontos de articulação, com a transição de F2. De acordo com a literatura, a região de *onset* de F1 e F2 pode conter informações que são pertinentes para diferenças no vozeamento ou ponto de articulação, respectivamente (COOPER *et al.*, 1952; LIBERMAN *et al.*, 1958; KEWLEY-POR, 1982, dentre outros). É nessa região, por exemplo, que se encontra o *locus* consonantal, isto é, a região de concentração de energia na transição de F2, que pode caracterizar uma plosiva de acordo com o seu ponto de articulação. Além do ponto de início da transição (*onset*), observaremos também a curva de transição de F1 e de F2 formada pelos cinco pontos de coleta, dentre o início (momento 1 coletado ao final da produção da consoante plosiva) e o final (momento 5 coletado na região estável da vogal adjacente). De acordo com a literatura, a direção da curva de transição (se ascendente ou descendente), também é diferente de acordo com o vozeamento ou ponto de articulação (COOPER *et al.*, 1952; DELATTRE; LIBERMAN; COOPER, 1955; HAGGARD; AMBLER; CALLOW, 1970; LIBERMAN; DELATTRE; COOPER, 1958; STEVENS; KLATT, 1974, dentre outros).

### 4.7.1 Transição de F1

De acordo com a literatura, a transição de F1 pode auxiliar na distinção entre o vozeamento de uma consoante plosiva (COOPER *et al.*, 1952). Conforme vimos anteriormente, as transições de F1

consideradas ascendentes, com valores de F1 mais baixos, são representativas de plosivas vozeadas e transições mais planas, com valores de F1 mais altos apontam para plosivas não-vozeadas (COOPER *et al.*, 1952; DELATTRE; LIBERMAN; COOPER, 1955; LIBERMAN; DELATTRE; COOPER, 1958; BENKI, 2001; FICKER, 2003).

Assim, a **Pergunta 11** questiona: uma vez que F1 está relacionado ao tipo de vozeamento, há diferenças, em termos de frequência, entre [p] x [b]; entre [t] x [d]; e entre [k] x [g]?

Conforme a Tabela 36, os valores de F1 no *onset* ou momento inicial de coleta (F1\_1) de plosivas não-vozeadas, quando comparadas às suas contrapartes homorgânicas, com base nos testes Mann-Whitney, não apontaram diferenças significativas entre os pares. Assim, não podemos afirmar que a diferença entre os valores de *onset* de F1 encontradas para o nosso estudo foi capaz de diferenciá-las acusticamente.

Tabela 36 – Resultados para a transição de F1 em Hz no primeiro e último momento de coleta, juntamente com o resultado dos testes obtidos na comparação de F1\_1 entre os pares de plosivas homorgânicas<sup>55</sup>

		Vozeadas F1_1	Não-vozeadas F1_1	Valor do teste <sup>56</sup>	Valor de p
F1	Bil.	450	463	$t(60) = -0,3$	$> 0,05$
	Alv.	433	421	$t(72) = 0,8$	$> 0,05$
	Velar	434	456	$t(62) = -2,0$	$> 0,05$

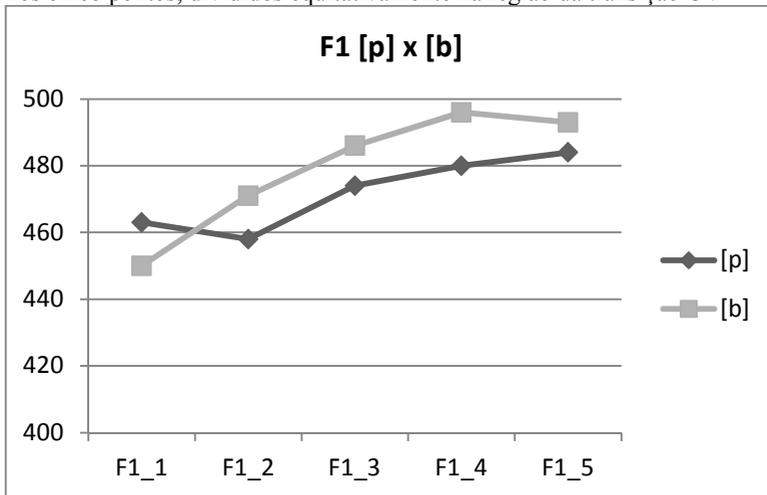
Em um estudo para o PB, Ficker (2003) avaliou valores de *onset* de F1 e observou diferenças significativas entre os pares homorgânicos. Ao analisar a diferença entre o momento de *onset* de F1 e o momento final da transição, a pesquisadora observou transições mais acentuadas para as plosivas vozeadas, justamente por encontrar valores de *onset* de F1 mais baixos para esses segmentos.

Podemos observar a curva que representa a transição de F1 das plosivas bilabiais vozeadas e não-vozeadas para nosso estudo através do Gráfico 4. Cada ponto na curva se refere a cada um dos cinco momentos de coleta.

<sup>55</sup> O F1\_1 refere-se ao primeiro momento de coleta, posicionado na transição entre consoante e vogal e o F1\_5 refere-se ao último momento, posicionado próximo à região estável da vogal seguinte.

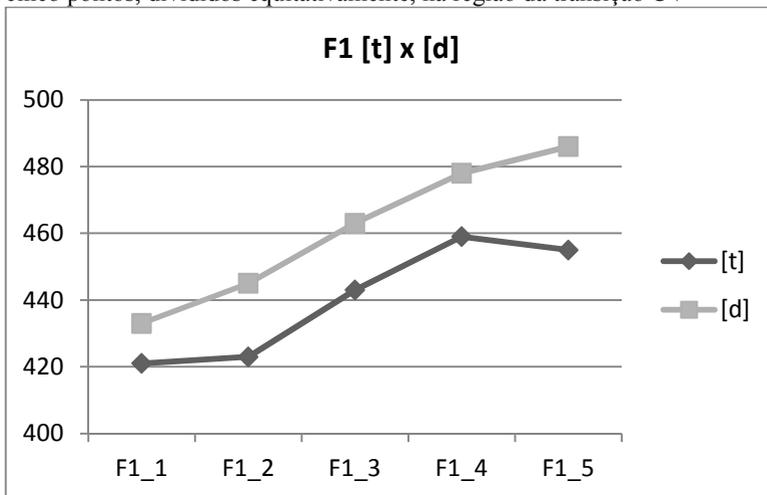
<sup>56</sup> Testes aplicados aos valores de início de F1 (F1\_1) entre os pares de plosivas homorgânicas.

Gráfico 4 – Plotagem dos valores médios de F1 (Hz) para [p] x [b], coletados nos cinco pontos, divididos equitativamente na região da transição CV



Ambas as transições de [b] e de [p], apresentam-se ascendentes, sendo mais acentuada para a plosiva vozeada, corroborando os resultados de Ficker (2003). Podemos observar a mesma tendência de ascensão para a comparação entre as plosivas alveolares, no Gráfico 5.

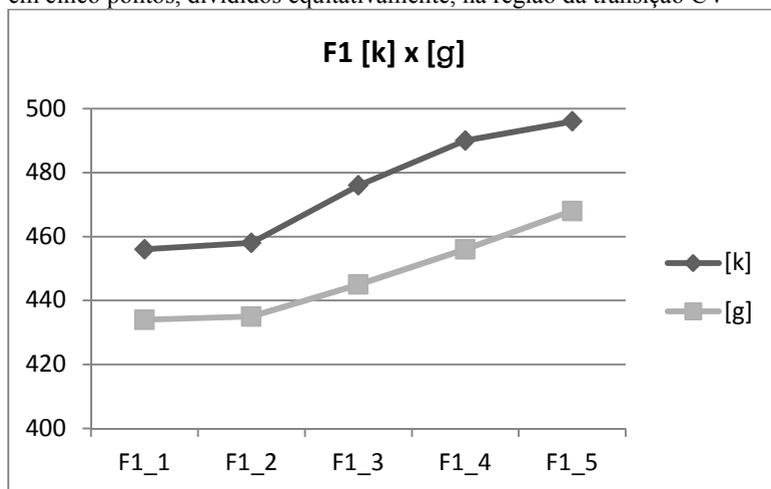
Gráfico 5 – Plotagem dos valores médios de F1 (Hz) para [t] e [d], coletados em cinco pontos, divididos equitativamente, na região da transição CV



Os dois gráficos da transição de F1 das alveolares também mostram tendências ascendentes, conforme encontrado por estudos anteriores (FICKER, 2003, BENKI, 2001).

A única transição que não se comportou como o esperado foi a das plosivas velares que, como se pode observar no Gráfico 6, é mais acentuada para plosiva não-vozeada.

Gráfico 6 – Plotagem dos valores médios de F1 (Hz) para [k] e [g], coletados em cinco pontos, divididos equitativamente, na região da transição CV



A partir dos resultados obtidos, concluímos que a **Hipótese 11**, que previa que as plosivas vozeadas iriam apresentar valores de início de coleta (*onset*) mais baixos do que as não-vozeadas foi parcialmente confirmada. Apesar dos testes estatísticos não apontarem para diferenças significativas entre esses valores, os gráficos representando as curvas de transição de F1 para bilabiais e alveolares foram mais acentuados para

as plosivas vozeadas, conforme já encontrado pelo estudo de Ficker (2003) para o PB. Acreditamos que nossos resultados, apesar de reiterarem tendências de ascensão e valores de F1 já apontados pela literatura, ainda não são conclusivos para se afirmar que a transição de F1 é capaz de diferenciar acusticamente as plosivas com relação ao vozeamento no PB.

#### 4.7.2 Transição de F2

Como relatado anteriormente, a transição de F2 é um dos parâmetros que pode apontar para a diferença entre pontos de articulação. A transição de F2 das bilabiais tende a se concentrar em baixas frequências; a das alveolares em médias (em torno de 1800 Hz); e a das velares é dependente do tipo da vogal que segue essa consoante, podendo ser maior do que as demais plosivas quando seguida por vogais anteriores e podendo sofrer atenuação quando seguida por vogal posterior (CLARK; YALLOP, 1995; COOPER et al., 1952; DELATTRE; LIBERMAN; COOPER, 1955; LIBERMAN; DELATTRE; COOPER, 1952; FANT, 1973, dentre outros). Para o estudo desse parâmetro, lançamos a **Pergunta 12.1**: Há diferenças em F2 entre as plosivas não-vozeadas?

No que concerne a transição de F2, vamos investigar a relação entre as plosivas, de acordo com o contexto vocálico seguinte, a saber, as vogais finais [ɐ] e [u]. Em termos acústicos, essas duas vogais diferem entre si pela frequência de seus formantes: [ɐ] tem F2 em torno de 1300 Hz e a vogal [u] tem F2 em torno de 820 Hz, de acordo com Seara (2000). Esses valores serão importantes, especialmente na comparação entre a ascendência e descendência da curva de F2, uma vez que a transição irá apontar para diferentes locais a depender do F2 da vogal adjacente.

Começaremos a análise pela comparação entre as três plosivas não-vozeadas. A Tabela 37 apresenta os resultados encontrados para cada uma delas.

Tabela 37 – Resultados para a transição de F2 (Hz) no primeiro e último momento de coleta, juntamente com os resultados dos testes obtidos na comparação de F2\_1 (*onset*) entre os pares de plosivas não-vozeadas<sup>57</sup>

---

<sup>57</sup> O F2\_1 refere-se ao primeiro momento de coleta, posicionado na transição entre consoante e vogal e o F2\_5 refere-se ao último momento, posicionado próximo a região estável da vogal seguinte.

	CV	F2_1	Valor do teste	Valor de p
F2	[pɐ]	1396	$U = 109^{58}$	< 0,017
	[tɐ]	1761		
	[pɐ]	1396	$U = 68$	< 0,017
	[kɐ]	1746		
	[tɐ]	1761	$U = 134$	> 0,05
	[kɐ]	1746		
	[pʊ]	1399	$H(2) = 3,25$ p > 0,05	
	[tʊ]	1567		
	[pʊ]	1399		
	[kʊ]	1227		
	[tʊ]	1567		
	[kʊ]	1227		

Para o contexto de vogal seguinte [ɐ], o teste Kruskal-Wallis apontou para diferenças significativas nos valores de F2 no momento inicial de coleta (F2\_1) entre as plosivas não-vozeadas ( $H(2) = 12,53$ ,  $p < 0,01$ ). Os resultados dos testes Mann-Whitney com a bilabial apresentando valores significativamente mais baixos do que as demais (alveolares e velares).

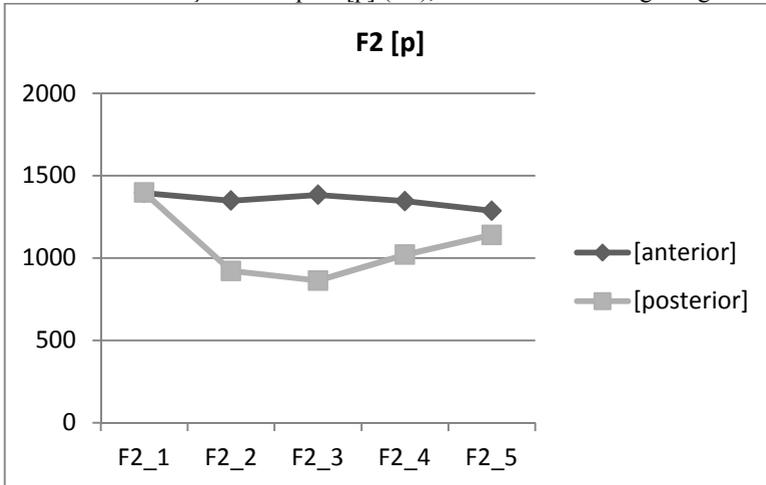
Com relação às análises no contexto consoante+[ʊ], não encontramos diferenças significativas entre os valores de F2 para as plosivas não-vozeadas, conforme o teste Kruskal-Wallis ( $H(2) = 3,25$ ,  $p > 0,05$ ).

O que podemos avaliar a partir da Tabela 37 é uma atenuação nos valores de frequência da plosiva velar quando seguida pela vogal posterior [ʊ], conforme apontado nos estudos de Delattre *et al.* (1955), Fant (1973) e Berti e Marino (2011), para o PB.

Além das informações concernentes ao *onset* (F2\_1), a curva de transição também pode mostrar certas particularidades referentes ao ponto de articulação. Encontramos valores de F2 relativamente altos para [p], o que não era esperado, pois a literatura afirma que, de forma geral, a transição de F2 para as bilabiais deve ser ascendente (DELATTRE; LIBERMAN; COOPER, 1955) e não descendente, conforme mostrado pelo Gráfico 7.

<sup>58</sup> Teste realizado somente entre os valores de F2\_1, momento que marca o início da transição de F2, ainda na zona de transição. F2\_5 mostra os valores de final de transição, já dentro da zona estável da vogal seguinte.

Gráfico 7 – Transição de F2 para [p] (Hz), de acordo com a vogal seguinte



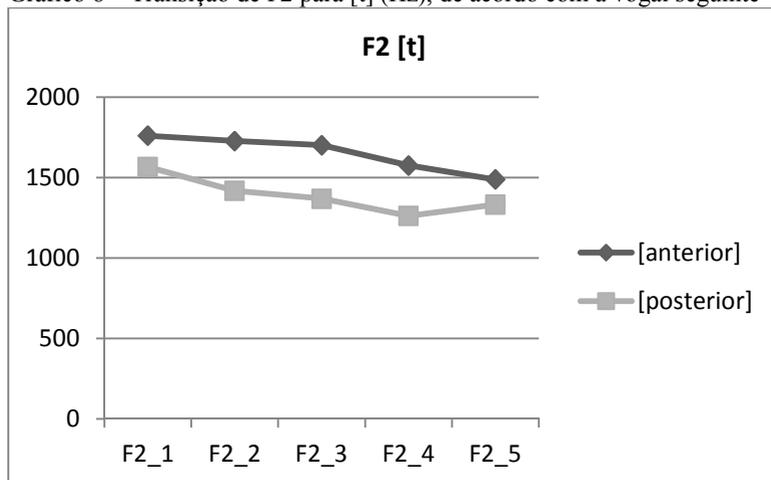
Com relação à comparação da transição de F2\_1 entre os diferentes contextos vocálicos adjacentes, o Gráfico 7 mostra que, para a bilabial, os valores estão em torno de 1400 Hz, tanto para [pɐ] quanto para [pʊ]. Aplicando o teste Mann-Whitney ao F2 entre [pɐ] x [pʊ], não encontramos diferenças significativas. Como Ficker (2003) já havia encontrado valores em torno de 1300 Hz para o PB para [pa], esse resultado pode estar indicando um possível *locus* consonantal para as plosivas bilabiais no PB.

Com relação à plosiva [t], notamos diferença entre os valores de F2\_1 dependendo do contexto vocálico. A transição [ta] apresentou média de 1761 Hz, valor bastante aproximado ao que a literatura aponta como seu *locus* consonantal, 1800 Hz (BLUMSTEIN; STEVENS, 1979) e a transição [tu] apresentou média de 1567 Hz. A estatística apontou para diferenças significativas entre a transição [tɐ] e [tu] ( $t(26) = 3,43, p < 0,01$ ). Essa atenuação de valores de transição de F2 quando a plosiva encontra-se diante de vogal posterior também foi observada por Lehiste e Peterson (1961), que encontraram valores mais elevados para a vogal anterior em [ti] = 1825 Hz e mais baixo para a vogal posterior em [tɔ] = 1600 Hz.

Não obstante, a transição de F2 da consoante alveolar foi sempre descendente, corroborando os estudos de Delattre *et al.* (1955) para o

inglês e Ficker (2003), para o PB. O Gráfico 8 mostra como se deu a transição de F2 para as alveolares.

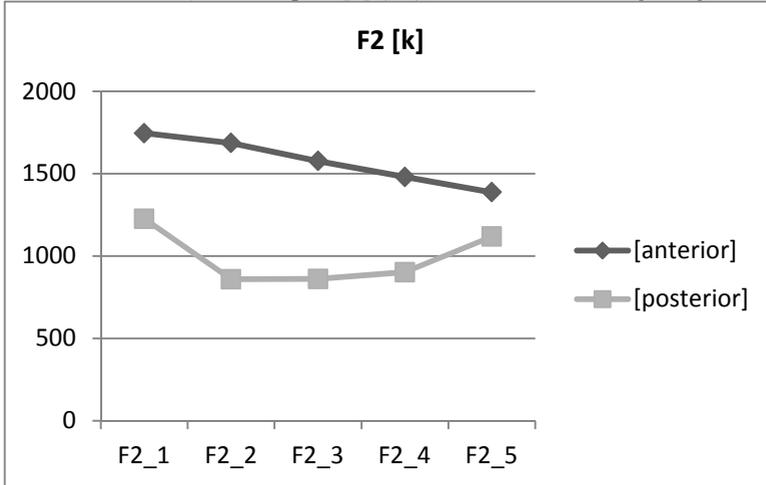
Gráfico 8 – Transição de F2 para [t] (Hz), de acordo com a vogal seguinte



Por fim, a transição de F2 de [k] também seguiu o encontrado pelos autores supracitados, com curva descendente, como pode ser visualizado no Gráfico 9. É possível observar similaridade entre os valores de transição de F2 para a vogal anterior<sup>59</sup> entre o nosso estudo (média de 1746 Hz para [kɐ]), com os encontrados por Ficker (2003), com média de 1725 Hz para [ka].

<sup>59</sup> Cabe salientar que Ficker (2003) não avaliou a transição de F2 no contexto de vogal posterior.

Gráfico 9 - Transição de F2 para [k] (Hz), de acordo com a vogal seguinte



Testes estatísticos mostraram que há diferenças significativas entre os valores iniciais de transição de F2 na comparação entre o contexto de vogal anterior e posterior ( $U = 107$ ,  $p < 0,01$ ). O valor médio de transição de F2 de [k $\upsilon$ ], 1228 Hz, se aproxima do valor encontrado por Lehiste e Peterson (1961) para a transição de [ku] no inglês, de 1200 Hz.

Com relação às plosivas vozeadas, a **Pergunta 12.2** questiona: há diferenças em F2 entre as plosivas vozeadas?

A Tabela 38 mostra que a plosiva [b] foi a que apresentou valores mais baixos de transição de F2 no contexto de vogal seguinte [e] e que a velar foi a que apresentou o menor valor de F2 no contexto de vogal [u], repetindo o resultado encontrado para as plosivas não-vozeadas.

Tabela 38 – Resultados para a transição formântica de F2 (Hz) no primeiro e último momento de coleta, juntamente com os resultados dos testes obtidos na comparação de F2\_1 entre os pares de plosivas vozeadas<sup>60</sup>

	CV	F2_1	F2_5	Valor do teste	Valor de p
F2	[b $\upsilon$ ]	1440	1343	$F(2,42) = 21,91$	< 0,001
	[d $\upsilon$ ]	1748	1612		

<sup>60</sup> O F2\_1 refere-se ao primeiro momento de coleta, posicionado na transição entre consoante e vogal e o F2\_5 refere-se ao último momento, posicionado próximo a região estável da vogal seguinte.

[bɐ]	1440	1343		
[gɐ]	1890	1466		
[dɐ]	1748	1612		
[gɐ]	1890	1466		
[bʊ]	1377	1148	$U = 91^{61}$	$> 0,017$
[dʊ]	1718	1325		
[bʊ]	1377	1111	$U = 45$	$> 0,017$
[gʊ]	1064	1035		
[dʊ]	1718	1325	$U = 10$	$< 0,017$
[gʊ]	1064	1035		

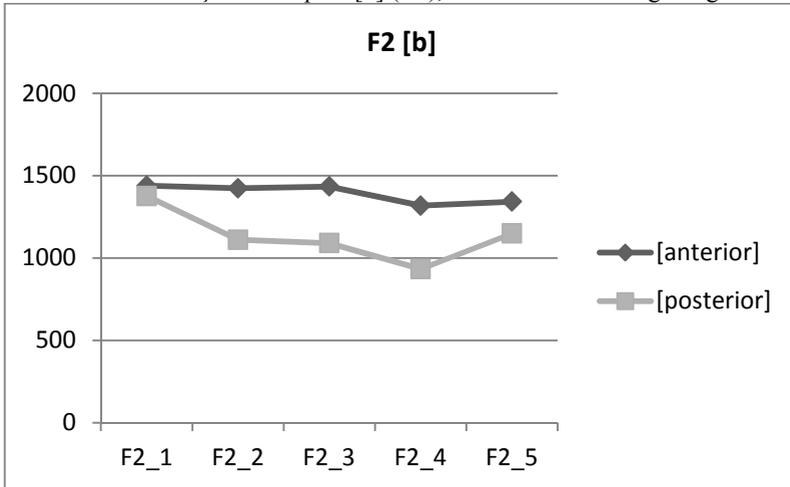
O teste Anova de amostras independentes revelou diferenças significativas entre o F2\_1 das plosivas vozeadas seguidas por [ɐ] ( $F(2,42) = 21,91$ ,  $p < 0,001$ ). O Teste *Post Hoc* de Gabriel mostrou diferenças significativas entre a bilabial e as demais, resultado semelhante ao das não-vozeadas. A análise de F2\_1 entre as vozeadas seguidas da vogal [ʊ] revelou que há diferença significativa ( $H(2) = 19,10$ ,  $p < 0,001$ ). Testes Mann-Whitney com correção Bonferroni aplicados nas comparações entre elas mostrou diferenças significativas apenas entre a alveolar e as demais (Tabela 38).

No que tange à comparação na transição de F2 das plosivas entre os diferentes contextos vocálicos adjacentes, os valores encontrados para F2\_1 entre [bɐ] e [bʊ] são bastante próximos entre si, indicando que essa consoante não sofre alteração no valor inicial de transição com a alteração da vogal seguinte ( $U = 89$ ,  $p > 0,05$ ). Da mesma forma que a sua contraparte não-vozeada, as transições de F2 de [b] mostram valores na região de 1400 Hz, valor semelhante ao encontrado por Ficker (2003) para [ba] – 1370 Hz. Essa informação confirma que as plosivas bilabiais do PB podem apresentar concentração de energia de F2, ou seja, o seu *locus* consonantal em 1400 Hz.

Conforme se observa no Gráfico 10, a curva de transição é praticamente plana para [bɐ] e descendente para [bʊ], justamente apontando para a região estável de cada uma das vogais que seguem as plosivas.

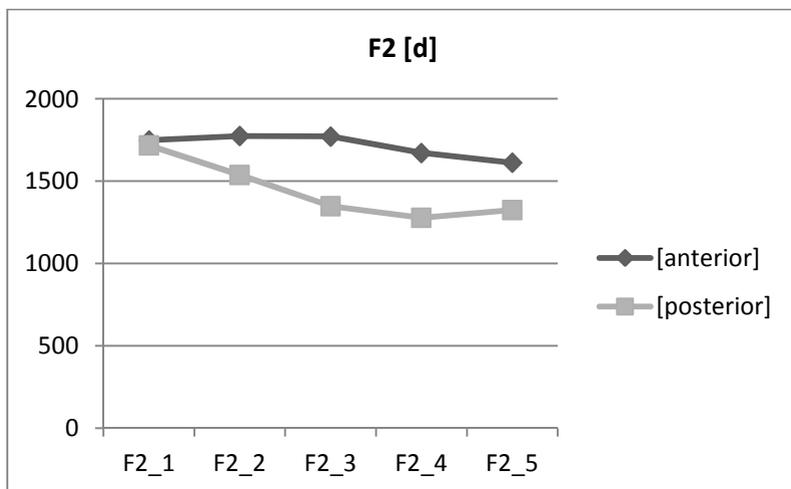
<sup>61</sup> Teste realizado somente entre os valores de F2\_1, momento que marca o início da transição de F2, ainda na zona de transição. F2\_5 mostra os valores de final de transição, já dentro da zona estável da vogal seguinte.

Gráfico 10 - Transição de F2 para [b] (Hz), de acordo com a vogal seguinte



A plosiva alveolar [d] apresentou valores de F2\_1 em torno de 1800 Hz, tanto para [tɐ] quanto para [tu], reiterando o fato de que este pode ser o *locus* consonantal para esse tipo de plosiva. O Gráfico 11 apresenta a curva de transição de F2 para a plosiva alveolar vozeada.

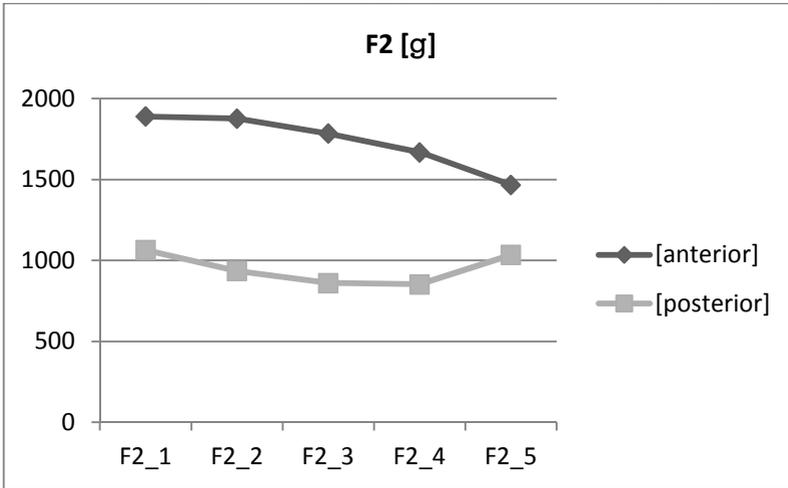
Gráfico 11 – Transição de F2 para [d] (Hz), de acordo com a vogal seguinte



Um teste Mann-Whitney foi aplicado às alveolares, de acordo com o contexto vocálico seguinte, e mostrou não haver diferenças significativas entre os valores de F2 inicial para [dɐ] e [dʊ]. Nota-se, também, que ambas as transições analisadas foram descendentes, corroborando os resultados de outras pesquisas (COOPER *et al.*, 1952; LIBERMAN *et al.*, 1952; FANT, 1973; FICKER, 2003; dentre outras).

Com relação à [g], os resultados da Tabela 38 mostraram que a vogal posterior à plosiva influencia drasticamente no valor referente ao momento inicial e na direção da transição de F2, da mesma forma que para a sua contraparte não-vozeada. No caso de [g], observa-se que, quando seguido pela vogal anterior, o seu valor de transição de F2 inicial mantém-se nas regiões de alta frequência, em 1890 Hz, mas quando diante da vogal posterior, esse valor sofre uma atenuação, caindo para 1064 Hz. O teste Mann-Whitney mostrou que essa diferença é significativa ( $t(17) = p < 0,001$ ). Esse resultado reitera a constatação de que a transição de plosivas velares é diferenciada dependendo da qualidade da vogal que as seguem desde o momento inicial de transição, indo ao encontro dos achados de Cooper *et al.* (1952), Liberman *et al.* (1952), Delattre *et al.* (1955), Fant (1973), dentre outros. Essa diferença pode ser visualizada através da comparação entre as curvas de F2 para [gɐ] e [gʊ] no Gráfico 12.

Gráfico 12 – Transição de F2 para [g] (Hz), de acordo com a vogal seguinte



As **Hipóteses 12.1** e **12.2** previam diferenças significativas entre os pontos de articulação de plosivas vozeadas e não-vozeadas. Como vimos, encontramos diferenças significativas entre [pɐ] e as demais não-vozeadas seguidas por [ɐ]; entre [bɐ] e as demais vozeadas seguidas por [ɐ] e entre [dʊ] e as demais vozeadas seguidas por [ʊ], o que confirma parcialmente essas Hipóteses. Elas também previam que haveria influência do contexto vocálico adjacente, especialmente no caso de das velares seguidas por vogal posterior, fato que foi confirmado tanto para a [k] quanto para [g].

Apesar de não encontramos resultados significativos para todas as análises de F2, observamos certas tendências em nossos dados que foram apontados por outras pesquisas como, por exemplo, *locus* consonantal para plosivas alveolares em torno de 1800 Hz, ordem de valores de frequência na transição de F2 da plosiva para a vogal anterior de: [bilabial] < [alveolar] < [velar] e de: [velar] < [bilabial] < [alveolar] para plosiva seguida de vogal posterior. Além disso, observamos um possível *locus* consonantal para as bilabiais no PB em torno de 1400 Hz.

No entanto, não podemos afirmar com exatidão a extensão da influência da transição de F2 na determinação do ponto de articulação das plosivas do PB. Conforme Kewley-Port (1982), pesquisas como as de Cooper et al. (1952), Liberman et al. (1952) e Delattre et al. (1955), com dados de fala sintética, e como as de Öhman (1966), Lehiste e

Peterson (1961) e Fant (1973), com fala natural, não conseguiram comprovar que a transição é um fator preponderante na distinção do ponto de articulação. Em seu estudo, Kewley-Port (1982) analisa as plosivas vozeadas do inglês seguidas pelas vogais [i, ɪ, e, ε, æ, a, o, u], a partir da gravação de um sujeito adulto. Segundo a autora,

“For the onset frequencies of the F1, F2, and F3 transitions, /b/ was statistically grouped separately from /d/ and /g/ averaged over all vowel contexts. This tendency, however, was by no means a reliable acoustic correlate of place of articulation (...). We conclude from the results of this investigation that invariant correlates of place of articulation are unlikely to be observed in the formant transition of initial stop consonants”. (KEWLEY-PORT, 1982, p386-7)<sup>62</sup>

Acreditamos que o parâmetro transição formântica é ainda muito controverso dentro da análise de consoantes plosivas. No nosso entendimento, esse é um ponto que merece um olhar mais aprofundado para se determinar a sua influência na distinção entre os diferentes pontos de articulação, tanto em termos de produção do segmento quanto de percepção. É provável que características frequenciais da transição de F2 como ascendência e descendência, *locus* consonantal, e outros aspectos possam ser pistas acústicas que, correlacionadas a outros parâmetros, possam ajudar no julgamento perceptual de um ou outro segmento, mas que sozinhos não sejam capazes de distinguir uma plosiva de outra.

#### 4.8 CONTEXTO VOCÁLICO ADJACENTE

Sabe-se que o contexto adjacente é um fator que pode influenciar ou alterar um segmento. Um exemplo disso é o resultado encontrado para a transição de F2 de plosivas velares, relatado na seção anterior. Essas consoantes apresentaram atenuação de transição de F2 quando

---

<sup>62</sup> “Para as frequências de início de transição de F1, F2 e F3, /b/ foi separado estatisticamente de /d/ e /g/, avaliados em todos os contextos vocálicos. Essa tendência, no entanto, não se mostrou um correlato acústico confiável de ponto de articulação (...). Nós concluímos com os resultados dessa investigação que correlatos invariantes de ponto de articulação são improváveis de serem observados na transição formântica de plosivas iniciais.” (KEWLEY-PORT, 1982, p386-7). Minha tradução.

seguidas pela vogal posterior se comparada à mesma plosiva seguida por uma vogal anterior.

Contudo, não sabemos se essa influência ocorre em apenas um sentido, ou seja, se a vogal influencia a plosiva ou se a vogal adjacente também pode sofrer interferência do tipo de plosiva que a segue ou antecede.

#### 4.8.1 Duração da vogal no contexto VC

Estudos apontam que a duração da vogal pode ser influenciada pelo tipo de vozeamento da consoante que a segue ou precede (KLATT, 1973; LADEFOGED; JOHNSON, 2011; LADEFOGED; MADDIESON, 1996; LISKER, 1957; STEVENS; KLATT, 1974; LADEFOGED, 2001; FICKER, 2003).

Desse modo, a **Pergunta 13** questiona: há diferença na duração das vogais que antecedem as plosivas não-vozeadas e plosivas vozeadas?

Para a vogal que antecede a plosiva, os resultados de média de duração encontrados foram de 162 ms quando seguida por plosivas não-vozeadas e 172 ms para a mesma vogal seguida por vozeadas, como pode ser observado pela Tabela 39. Cabe lembrar que analisamos a vogal tônica [a], inserida na primeira sílaba de palavra, como em “cAPa”.

Tabela 39 - Valores médios de duração absoluta (ms) e relativa da vogal (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração da vogal [a] no contexto VC, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa da vogal [a] quando seguida por plosivas vozeadas ou não-vozeadas

	Contexto	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração absoluta [a]</b>	Não-voz	162 (28)	100 – 232	$t(284) = -3,3$	<b>&lt; 0,01</b>
	Voz	172 (25)	92 – 230		
<b>Duração relativa [a]<sup>63</sup></b>	Não-voz	37 (6)	24 – 52	$t(284) = -6,0$	<b>&lt; 0,001</b>
	Voz	41 (5)	29 – 52		

O Teste-t de amostras independentes apontou que a diferença entre a duração absoluta e relativa da vogal [a] que antecede vozeadas e não-vozeadas é significativo (Tabela 39), indicando que a plosiva pode

<sup>63</sup> Duração relativa da vogal anterior em função da duração da palavra

influenciar a produção da vogal que a antecede, uma vez que quando seguida por consoante vozeada, a duração da vogal tende a ser maior do que quando ela está diante de plosiva não-vozeada.

Avaliamos também a duração absoluta e relativa da vogal diante de cada uma das plosivas e comparamos essa duração entre os pares de plosivas homorgânicas. Os resultados dessa análise estão dispostos na Tabela 40.

Tabela 40 – Valores médios de duração absoluta (ms) e relativa da vogal (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) da vogal [a] no contexto VC, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa da vogal entre os pares de plosivas homorgânicas

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração absoluta [a]</b>	[p]	144 (21)	100 – 205	$t(85) = -5,1$	<b>&lt; 0,001</b>
	[b]	168 (22)	112 – 207		
	[t]	178 (27)	115 – 232	$t(105) = 1,1$	> 0,05
	[d]	172 (28)	92 – 226		
	[k]	162 (24)	109 – 220	$t(90) = -2,9$	<b>&lt; 0,01</b>
	[g]	177 (22)	123 – 230		
<b>Duração relativa [a]</b>	[p]	33 (4)	24 – 42	$t(85) = -6,9$	<b>&lt; 0,001</b>
	[b]	39 (4)	30 – 47		
	[t]	41 (5)	28 – 52	$t(105) = 0,1$	> 0,05
	[d]	41 (5)	31 – 51		
	[k]	36 (4)	25 – 48	$t(90) = -5,0$	<b>&lt; 0,001</b>
	[g]	41 (5)	29 – 52		

Observando a Tabela 40, verificamos que a duração da vogal que precede as plosivas não apresentou diferenças significativas diante de plosivas alveolares. Nos demais pares homorgânicos bilabial e velar, a diferença existente na duração da vogal foi significativa e maior diante da plosiva vozeada, conforme esperado. Uma possível explicação para esse fato pode estar relacionado a não-concordância de vozeamento entre o segmento precedente e que sucede a tônica para as plosivas alveolares nos pares análogos dispostos no Quadro 2. Para os demais pares, há concordância no vozeamento, pois as consoantes que são adjacentes à tônica são sempre não-vozeadas, como por exemplo em **capa** e **faca**, mas não para **data**, que apresenta uma vozeada antecedendo à tônica. O mesmo ocorre para as vozeadas em **baba** e **baga**, mas não em **fada**. Como tanto a consoante que antecede quanto a que sucede uma

vogal pode exercer influência sobre a vogal, acreditamos que o fato da não-conformidade de vozeamento entre os pares análogos analisados podem explicar o fato de que as alveolares foram as únicas que não apresentaram diferenças significativas entre a duração da vogal tônica.

Esse resultado confirma parcialmente a **Hipótese 13**, que previa que a duração da vogal que precede a plosiva tende a ser maior quando ela se encontra diante de plosivas vozeadas. Somente diante de plosivas alveolares não houve diferenças na duração da vogal. O resultado referente às plosivas bilabiais e velares corrobora estudos anteriores para outras línguas e para o PB (BARROCO *et al.*; 2007; KLATT, 1973; LADEFOGED; JOHNSON, 2011; LADEFOGED; MADDIESON, 1996; LISKER, 1957; STEVENS; LOUSADA; MARTINS; JESUS, 2006; KLATT, 1974; BRITO, 2000 *apud* FICKER, 2003; BRITTO, 2010; FICKER, 2003).

#### 4.8.2 Duração da vogal no contexto CV

Com relação à duração da vogal seguinte à plosiva, comparamos a duração das vogais átonas finais [ɐ] e [ʊ], inseridas após contexto vozeado e não-vozeado, para avaliarmos se há influência do vozeamento na duração dessas vogais, conforme questiona a **Pergunta 14**: há diferença na duração das vogais que seguem as plosivas não-vozeadas e plosivas vozeadas?

Os resultados apresentados na Tabela 42 demonstram que, para ambas as vogais analisadas, as suas médias de duração foram significativamente maiores quando se encontravam seguindo a plosiva vozeada. Os Testes-t de amostras independentes confirmaram que a duração absoluta e relativa das vogais que seguem as plosivas [ɐ] e [ʊ] é significativamente mais longa quando essas vogais estão diante de plosivas vozeadas.

Tabela 41 - Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração das vogais [e] e [u] no contexto CV, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa dessas vogais quando seguidas por plosivas vozeadas ou não-vozeadas

	Contexto	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração absoluta [e]</b>	Não-voz	54 (13)	27 – 95	$t(162) = -5,7$	<b>&lt; 0,01</b>
	Voz	67 (15)	33 – 98		
<b>Duração relativa [e]<sup>64</sup></b>	Não-voz	12 (3)	7 – 24	$t(162) = -7,0$	<b>&lt; 0,001</b>
	Voz	16 (3)	8 – 24		
<b>Duração absoluta [u]</b>	Não-voz	40 (11)	17 – 64	$t(120) = -4,6$	<b>&lt; 0,001</b>
	Voz	50 (13)	24 – 83		
<b>Duração absoluta [u]</b>	Não-voz	9 (2)	4 – 14	$t(120) = -6,0$	<b>&lt; 0,001</b>
	Voz	12 (3)	7 – 18		

Analizamos também a duração absoluta e relativa das vogais [e] e [u] dentre os pares de plosivas homorgânicas que as antecediam. Os resultados são apresentados na Tabela 42.

<sup>64</sup> Duração relativa da vogal anterior em função da duração da palavra

Tabela 42 - Valores médios de duração absoluta (ms) e relativa (%), de desvio padrão, de faixa de distribuição da duração absoluta (ms) e relativa (%) para a duração das vogais [ɐ] e [ʊ] no contexto CV, juntamente com os resultados dos testes obtidos na comparação da duração absoluta e relativa das vogais [ɐ] e [ʊ] entre os pares homorgânicos

	Plosiva	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
<b>Duração absoluta</b> [ɐ]	[p]	63 (9)	41 – 81	$t(42) = -2,3$	<b>&lt; 0,05</b>
	[b]	68 (7)	54 – 77		
	[t]	51 (10)	27 – 70	$t(54) = -1,9$	<b>&lt; 0,05</b>
	[d]	58 (16)	33 – 86		
	[k]	50 (13)	28 – 82	$t(58) = -6,8$	<b>&lt; 0,001</b>
	[g]	73 (13)	56 – 95		
<b>Duração relativa</b> [ɐ]	[p]	15 (2)	12 – 19	$t(42) = -3,3$	<b>&lt; 0,01</b>
	[b]	17 (2)	14 – 21		
	[t]	12 (2)	8 – 16	$t(54) = -2,8$	<b>&lt; 0,01</b>
	[d]	14 (3)	8 – 19		
	[k]	11 (3)	7 – 20	$t(58) = -7,3$	<b>&lt; 0,001</b>
	[g]	17 (3)	12 – 23		
<b>Duração absoluta</b> [ʊ]	[p]	46 (9)	29 – 64	$t(38) = -2,6$	<b>&lt; 0,01</b>
	[b]	56 (14)	32 – 78		
	[t]	37 (10)	20 – 55	$t(48) = -2,3$	<b>&lt; 0,01</b>
	[d]	43 (8)	24 – 64		
	[k]	35 (12)	17 – 61	$t(29) = -4,2$	<b>&lt; 0,001</b>
	[g]	54 (12)	31 – 73		
<b>Duração relativa</b> [ʊ]	[p]	10 (2)	6 – 14	$t(38) = -2,8$	<b>&lt; 0,01</b>
	[b]	12 (3)	8 – 18		
	[t]	9 (2)	6 – 12	$t(48) = -3,3$	<b>&lt; 0,01</b>
	[d]	11 (2)	7 – 14		
	[k]	8 (3)	4 – 13	$t(29) = -5,3$	<b>&lt; 0,001</b>
	[g]	13 (3)	8 – 17		

Para todas as comparações da duração absoluta e relativa das vogais [ɐ] e [ʊ], de acordo com a plosiva que a antecedia (comparação entre os pares homorgânicos), os valores foram maiores quando essas vogais eram precedidas por consoantes vozeadas. Os Testes-t de amostras independentes revelaram que a duração absoluta e relativa das vogais [ɐ] e [ʊ] é significativamente maior quando elas seguem as plosivas vozeadas (Tabela 42).

Os resultados acima relatados confirmam a **Hipótese 14**, que previa que a duração das vogais fosse maior quando elas fossem antecedidas por consoantes vozeadas. Esse resultado vai ao encontro do postulado pela literatura, tanto para o PB, quanto para outras línguas (BRITTO, 2010; FICKER, 2003; KLATT, 1973, 1976; LADEFOGED; JOHNSON, 2011; LADEFOGED; MADDIESON, 1996; LISKER, 1957; LOUSADA; MARTINS; JESUS, 2006; STEVENS; KLATT, 1974).

#### 4.8.3 Influência da vogal na duração da consoante e do VOT

Outro ponto interessante sobre o contexto vocálico se refere à possível influência que a vogal que segue a plosiva pode exercer sobre a produção da consoante, a depender da altura dessa vogal. Esse aspecto é abordado pela **Pergunta 15** questiona: existe diferença na duração da consoante e de seu VOT dependendo da altura da vogal que a segue?

No caso das não-vozeadas, Klein (1999) e Alves (2011) encontraram valores maiores de VOT quando essas plosivas se encontravam diante de vogais altas, apesar de que a última pesquisadora não encontrou diferenças estatísticas entre as comparações.

Tabela 43 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração absoluta da consoante e do VOT (ms), juntamente com os resultados do testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com o contexto vocálico seguinte

		Vogal	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
[p]	Duração consoante	[e]	103 (12)	79 – 127	$t(45) = -3,3$	< 0,01
		[u]	117 (16)	89 – 147		
	VOT	[e]	16 (5)	5 – 29	$U = 113$	< 0,001
		[u]	26 (12)	10 – 57		
[t]	Duração consoante	[e]	105 (8)	89 – 124	$t(45) = -0,4$	> 0,05
		[u]	105 (14)	77 – 132		
	VOT	[e]	20 (6)	8 – 32	$t(45) = -1,6$	> 0,05
		[u]	23 (6)	16 – 38		
[k]	Duração consoante	[e]	105(12)	81 – 124	$t(47) = -2,3$	< 0,05
		[u]	113 (14)	82 – 137		
	VOT	[e]	39 (8)	21 – 53	$t(47) = -3,9$	< 0,001
		[u]	50 (10)	36 – 65		

Testes estatísticos revelaram que há diferenças significativas para a comparação entre [ɐ] e [ʊ] para [p] e [k], mas não para [t], conforme pode ser visto na Tabela 43. Nossos resultados reiteram os achados das pesquisas anteriores, para as plosivas bilabiais e velares não-vozeadas, pois os valores médios de duração da consoante e do VOT foram maiores quando essas plosivas eram seguidas pela vogal alta [ʊ]. Esse comportamento também foi encontrado por Port e Rotunno (1979), Klein (1999) e Alves (2011), sendo os dois últimos estudos relacionados ao PB.

Analisando a duração da consoante e VOT das plosivas vozeadas, os resultados apresentados na Tabela 44 também confirmam estudos anteriores (PORT; ROTUNNO, 1979; KLEIN, 1999; ALVES, 2011). A duração absoluta das consoantes alveolares e velares e do VOT dessas plosivas foi maior quando essas consoantes se encontravam diante da vogal alta.

Tabela 44 - Valores médios de duração da consoante e VOT (ms), de desvio padrão, de faixa de distribuição da duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e VOT para cada plosiva vozeada, de acordo com o contexto vocálico seguinte

	Vogal	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p
[b]	Duração consoante	[ɐ]	38 – 105	$t(37) = -1,9$	$> 0,05$
		[ʊ]	66 – 119		
	VOT	[ɐ]	(-)98 – (-)31	$t(37) = -1,4$	$> 0,05$
		[ʊ]	(-)107 – (-)60		
[d]	Duração consoante	[ɐ]	60 – 80	$t(45) = -5,9$	$< 0,001$
		[ʊ]	70 – 95		
	VOT	[ɐ]	(-)75 – (-)54	$t(45) = -6,5$	$< 0,001$
		[ʊ]	(-)87 – (-)65		
[g]	Duração consoante	[ɐ]	52 – 83	$t(33) = -3,7$	$< 0,001$
		[ʊ]	69 – 87		
	VOT	[ɐ]	(-)74 – (-)39	$t(47) = -2,7$	$< 0,05$
		[ʊ]	(-)66(8) – (-)51		

Testes estatísticos mostraram que a bilabial vozeada foi a única consoante que não apresentou diferença significativa entre os seus valores médios de duração da consoante e de VOT, de acordo com o

contexto vocálico seguinte. Para a alveolar e velar, os valores de duração da consoante e VOT foram significativamente maiores quando eram seguidas pela vogal alta [u]. Os resultados acima confirmam parcialmente a **Hipótese 15**, que previa que a duração da consoante e do VOT seriam maiores quando as plosivas se encontrassem diante das vogais altas.

Como Port e Rotunno (1979) para a língua inglesa, Klein (1999) encontrou diferenças significativas entre as comparações de média de duração de VOT com diferentes contextos vocálicos ([a] x [o], [a] x [e]), enquanto Alves (2011) não encontrou sistematicidade entre a comparação estatística do mesmo parâmetro em relação a vogais altas, médias e baixas, apesar de encontrar valores maiores em contexto de vogal alta. Acreditamos que nossos dados, apesar de não apresentarem resultados significativos entre todas as comparações, corroboram esses achados e mostram que pode haver uma tendência para a influência da qualidade vocálica na duração da consoante e VOT das plosivas do PB, cujos valores médios se mostram maiores quando a plosiva é seguida pela vogal alta.

#### 4.9 POSIÇÃO NA PALAVRA

O contexto posição na palavra foi apontado como um fator importante na produção de plosivas, especialmente no caso da língua inglesa, uma vez que as plosivas não-vozeadas e vozeadas, em posição inicial, são diferenciadas pela presença de aspiração nas primeiras (retardo longo) e ausência de aspiração nas últimas (retardo curto, sem batimento de pregas).

No entanto, para o PB, esse contexto parece não influenciar de forma sistemática o VOT dos segmentos plosivos (KLEIN, 1999; ALVES, 2011). Alves e Dias (2010) afirmam, contudo, que há considerável apagamento das vogais finais, levando ao alongamento no VOT das plosivas não-vozeadas que ocupam essa posição.

Com o intuito de observar se esse parâmetro pode realmente afetar ou não a produção das plosivas do PB, lançamos a **Pergunta 16.1**: o contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas não-vozeadas do PB?

A Tabela 45 apresenta as diferenças estatísticas encontradas no valor médio de duração e de VOT entre uma mesma plosiva inserida em contexto inicial e medial.

Tabela 45 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com a posição na palavra

	Posição	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p	
[p]	<b>Duração consoante</b>	Início	123 (20)	86 – 176	$t(321) = -3,2$	<b>&lt; 0,01</b>
		Meio	130 (21)	81 – 176		
	<b>VOT</b>	Início	21(8)	8 – 41	$U = 10143$	<b>&lt; 0,001</b>
		Meio	18 (6)	6 – 42		
[t]	<b>Duração consoante</b>	Início	126 (21)	80 – 176	$t(290) = -0,6$	<b>&gt; 0,05</b>
		Meio	125 (21)	77 – 172		
	<b>VOT</b>	Início	20 (6)	9 – 47	$U = 9886$	<b>&gt; 0,05</b>
		Meio	21 (8)	8 – 43		
[k]	<b>Duração consoante</b>	Início	137(21)	75 – 184	$t(341) = 7,6$	<b>&lt; 0,001</b>
		Meio	120 (22)	66 – 184		
	<b>VOT</b>	Início	55 (19)	21 – 112	$U = 12001$	<b>&lt; 0,01</b>
		Meio	50 (15)	23 – 102		

No caso da duração da consoante, cada uma das plosivas obteve um resultado distinto: para a bilabial, a duração foi significativamente maior em contexto medial. Para a velar, foi observada duração significativamente maior quando a plosiva se encontrava em contexto inicial. Para a alveolar, não houve diferença significativa na duração da plosiva considerando a posição na palavra.

No que diz respeito ao VOT, observamos o mesmo comportamento, já apontado em Alves (2011). Nossos resultados mostraram que a bilabial e a velar apresentaram VOT significativamente maior em contexto inicial. Para a alveolar não foram observadas diferenças significativas.

Em relação às vozeadas, a **Pergunta 16.2** questiona: o contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas vozeadas do PB?

Os resultados para as vozeadas se mostraram um pouco diferentes do que a categoria anterior. Para essas plosivas, a duração absoluta da consoante foi sempre significativamente maior quando inseridas em posição inicial, conforme os resultados apresentados na Tabela 46. Esse resultado é semelhante aos apresentados em Klatt (1976) para a língua inglesa. Esse autor observou que as plosivas vozeadas dessa língua são de 10 a 30 ms mais longas em posição inicial. Para nossos dados, as consoantes vozeadas em posição inicial foram, em média, 7,3 ms mais longas do que as em posição medial.

Tabela 46 – Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva vozeada, de acordo com a posição na palavra

	Posição	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p	
[b]	<b>Duração consoante</b>	Início	103 (21)	55 – 175	$t(321) = 2,2$	< 0,05
		Meio	98 (20)	52 – 153		
	<b>VOT</b>	Início	(-)93(21)	(-)166 – (-)48	$U = 11334$	> 0,05
		Meio	(-)89 (20)	(-)137 – (-)41		
[d]	<b>Duração consoante</b>	Início	97 (24)	46 – 165	$U = 10660$	< 0,05
		Meio	89 (18)	43 – 133		
	<b>VOT</b>	Início	(-)87 (25)	(-)155 – (-)38	$t(135) = 3,2$	< 0,01
		Meio	(-)80 (18)	(-)123 – (-)34		
[g]	<b>Duração consoante</b>	Início	95(23)	45 – 171	$t(254) = 3,2$	< 0,01
		Meio	86 (21)	38 – 138		
	<b>VOT</b>	Início	(-)72 (20)	(-)134 – (-)33	$t(246) = 2,1$	< 0,05
		Meio	(-)67 (17)	(-)119 – (-)21		

Da mesma forma que a duração da consoante, o VOT das plosivas vozeadas apresentou duração significativamente maior em contexto inicial, à exceção da bilabial (Tabela 46).

Baseado nos resultados apresentados, podemos concluir que a **Hipótese 16.1** (que previa que a posição na palavra exerceria influência sobre a duração da consoante e do VOT das plosivas não-vozeadas) foi parcialmente confirmada, pois não encontramos sistematicidade na influência da posição da palavra sobre a duração da consoante e VOT das plosivas não-vozeadas e nem sempre os valores foram significativamente diferentes. No entanto, a **Hipótese 16.2** pode ser confirmada, pois os resultados para as plosivas vozeadas se mostraram sistemáticos para as avaliações de duração de consoante e VOT, cujos valores se mostraram sempre maiores em início de palavra. Apenas para os resultados concernentes ao VOT da consoante bilabial não foram observadas diferenças significativas.

Cabe ressaltar que, tanto em posição inicial quanto em medial, as plosivas não-vozeadas do PB apresentaram o seu VOT dentro da categoria de retardo curto (Tabela 45), sem presença de vozeamento; e as vozeadas dentro da categoria de pré-sonorização (Tabela 46), pois o vozeamento se inicia antes da explosão. Com relação à duração da consoante, as plosivas vozeadas apresentaram valores significativamente menores do que as suas contrapartes não-vozeadas nos dois contextos avaliados.

Assim, a posição na palavra não pode ser apontada como um parâmetro que diferencia as duas categorias de plosivas no PB, diferentemente da língua inglesa que, em posição inicial, apresenta falta de vozeamento para plosivas não-vozeadas, que se diferenciam apenas pela duração do VOT (com aspiração para as primeiras e sem aspiração para as segundas). Acreditamos, então, que a posição na palavra pode ser um parâmetro que se diferencia apenas na produção das plosivas vozeadas do PB e que, talvez, possa auxiliar na distinção entre os pontos de articulação dessa categoria.

#### 4.10 TONICIDADE

O contexto tonicidade influencia as plosivas de maneira diferente, a depender do tipo de consoante que estamos investigando.

No que diz respeito às consoantes não-vozeadas, pesquisas feitas para a língua inglesa indicam que o contexto pré-tônico privilegia o aumento da duração da plosiva (KLATT, 1976) e que o VOT é aumentado quando a plosiva se encontra em contexto tônico (GIMSON, 1980; LISKER; ABRAMSON, 1964c, 1967).

Estudos no PB indicam que os maiores valores de VOT são encontrados quando as plosivas se encontram em contexto pós-tônico (e, portanto, átono), especialmente para as plosivas bilabiais e alveolares, enquanto o tônico propicia um aumento no VOT das velares. (KLEIN, 1999; ALVES et al., 2008; ALVES, 2011).

Assim, a **Pergunta 17.1** questiona: o contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas não-vozeadas do PB?

De acordo com os resultados dispostos na Tabela 47, no que concerne a duração da consoante para a plosiva não-vozeada, todos os valores indicam que elas apresentam maior duração quando inseridas em contexto tônico. Os testes Mann-Whitney realizados entre os contextos de tonicidade para cada plosiva confirmam que a diferença entre a duração da consoante é significativamente maior quando a plosiva está inserida no contexto tônico. Esse resultado diverge daquele encontrado por Klatt (1976) para a língua inglesa. Não encontramos estudos para o PB, até o presente momento, que avalie a influência da tonicidade na duração total da consoante.

Tabela 47 - Valores médios (ms), de desvio padrão, de faixa de distribuição para a duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva não-vozeada, de acordo com a tonicidade da sílaba

	Sílaba	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p	
[p]	<b>Duração consoante</b>	Tônica	137 (17)	97 – 176	$U = 1974$	<b>&lt; 0,001</b>
		Átona	108 (12)	81 – 135		
	<b>VOT</b>	Tônica	19 (7)	6 – 38	$U = 10828$	<b>&gt; 0,05</b>
		Átona	21 (7)	8 – 41		
[t]	<b>Duração consoante</b>	Tônica	136 (17)	97 – 176	$U = 2187$	<b>&lt; 0,001</b>
		Átona	108 (13)	77 – 139		
	<b>VOT</b>	Tônica	19 (6)	8 – 41	$U = 6402$	<b>&lt; 0,001</b>
		Átona	23 (8)	9 – 47		
[k]	<b>Duração consoante</b>	Tônica	136 (23)	85 – 184	$U = 7257$	<b>&lt; 0,001</b>
		Átona	117 (19)	66 – 167		
	<b>VOT</b>	Tônica	53 (17)	23 – 104	$U = 12896$	<b>&gt; 0,05</b>
		Átona	52 (18)	21 – 112		

A duração do VOT das não-vozeadas só mostrou diferenças significativas entre contexto tônico e átono para a consoante alveolar. Nesse caso, os resultados seguem parcialmente a tendência encontrada por outras pesquisas para o PB, isto é, plosivas alveolares apresentaram maior VOT quando inseridas em contexto átono. Para as bilabiais e velares, o conjunto de dados analisados no presente estudo não mostrou diferenças significativas entre contexto átono e tônico Assim, embora os resultados de Klein (1999), Alves *et al.* (2008) e Alves (2011) apontassem para diferenças entre velares, sendo o VOT maior quando em contexto tônico, e para diferenças entre alveolares, sendo nesse caso, maior em contexto átono, os testes Mann-Whitney aplicados às plosivas do presente estudo revelaram que a diferença nos valores de VOT de plosivas em contexto átono ou tônico foi significativa somente para a [t].

Alguns autores mencionam que a quantidade de aspiração<sup>65</sup> também pode ser influenciada pela tonicidade da sílaba na qual a plosiva está inserida, no sentido de que a aspiração é mais acentuada em contexto tônico (GIMSON, 1980; LISKER; ABRAMSON, 1964c). Para averiguar se isso ocorre também em nossos dados, realizamos uma análise com a plosiva [k] que, como visto anteriormente, apresentou um

<sup>65</sup> Para haver aspiração, o VOT da plosiva deve apresentar valores maiores que 35 ms. Quanto maior o valor do VOT, mais aspiração a plosiva apresenta.

elevado número de plosivas com aspiração (87% do total). Considerando as médias de VOT encontradas para velares inseridas em contexto tônico (55,82 ms) e em contexto átono (54,90ms), foi rodado o teste Mann-Whitney que mostrou que não existem diferenças significativas entre a duração de VOT nos dois contextos de tonicidade ( $U = 5264$ ,  $p > 0,05$ ). Assim, concluímos que, para os nossos dados, parece não haver relação entre a quantidade de aspiração e a tonicidade.

Já a **Pergunta 17.2** questiona: o contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas vozeadas do PB?

Com relação a essas plosivas, os testes Mann-Whitney mostraram que, para cada uma das plosivas vozeadas, a duração da consoante é significativamente maior quando elas estão inseridas em contexto tônico, conforme os resultados apresentados na Tabela 48.

Ao estudar a relação entre o VOT e a tonicidade, Klein (1999) encontrou que o valor médio de VOT é maior quando as plosivas vozeadas estão inseridas em contexto tônico. Como se pode observar através dos resultados apresentados na Tabela 48, encontramos a mesma tendência de Klein (1999), pois os testes Mann-Whitney mostraram que há diferenças significativas entre a duração do VOT de plosivas em contexto átono e tônico.

Tabela 48 - Valores médios de duração da consoante e VOT (ms), de desvio padrão, de faixa de distribuição da duração da consoante e VOT (ms), juntamente com os resultados dos testes obtidos na comparação de duração de consoante e do VOT para cada plosiva vozeada, de acordo com a tonicidade da sílaba

	Sílaba	$\bar{X}$ (DP)	Faixa	Valor do teste	Valor de p	
[b]	<b>Duração cons.</b>	Tônica	111 (18)	57 – 175	$U = 4374$	< 0,001
		Átona	88 (16)	52 – 131		
	<b>VOT</b>	Tônica	(-101 (18)	(-166 – (-)52	$U = 4068$	< 0,001
		Átona	(-78 (15)	(-121 – (-)41		
[d]	<b>Duração cons.</b>	Tônica	104 (19)	51 – 165	$U = 2467$	< 0,001
		Átona	76 (13)	43 – 109		
	<b>VOT</b>	Tônica	(-93 (19)	(-155 – (-)41	$U = 2383$	< 0,001
		Átona	(-66 (13)	(-101 – (-)34		
[g]	<b>Duração cons.</b>	Tônica	99 (21)	55 – 171	$U = 3652$	< 0,001
		Átona	80 (20)	38 – 136		
	<b>VOT</b>	Tônica	(-76 (20)	(-150 – (-)20	$U = 3584$	< 0,001
		Átona	(-58 (17)	(-119 – (-)14		

Os resultados anteriores confirmam parcialmente a **Hipótese 17.1**, que previa que a duração da consoante seria maior em contexto átono e que o VOT de [p] e [t] seria maior em contexto átono e de [k] em contexto tônico. A **Hipótese 17.2** se confirma parcialmente em relação à duração da consoante. Diferentemente de estudos anteriores para outras línguas (KLATT, 1976), a duração da consoante foi significativamente maior para consoantes vozeadas em contexto tônico. Com relação ao VOT, encontramos resultados que confirmam essa hipótese, uma vez que os valores de VOT das vozeadas foram sempre maiores em contexto tônico.

Apesar de avaliar a questão da tonicidade, avaliando possíveis diferenças entre plosivas inseridas em sílabas tônicas e átonas, não investigando questões correlacionando tonicidade e posição na palavra, como, por exemplo, através da investigação de possíveis diferenças entre plosivas em sílaba pré e pós-tônica. Assim, sugerimos que pesquisas futuras observem possíveis diferenças em plosivas inseridas nesses contextos e ainda uma possível correlação entre a variável posição na palavra e tonicidade.

#### 4.11 RESUMO DOS RESULTADOS

Observamos nas análises apresentadas anteriormente que muitos dos parâmetros acústicos e contextos linguísticos envolvidos na produção de plosivas no PB são utilizados por falantes nativos dessa língua na caracterização de cada uma das três plosivas. Esses parâmetros podem não somente ser utilizados na caracterização da produção desses segmentos, mas, possivelmente, podem ser utilizados como pistas pelos ouvintes para identificar cada um desses segmentos. A Tabela 49 traz um resumo dos principais resultados encontrados através da análise dos dados coletados pela presente pesquisa para cada um dos parâmetros e contexto avaliados. Cabe ressaltar que os valores dispostos nessa tabela se referem às medias encontradas para cada uma das plosivas avaliadas a depender do parâmetro em análise. Por questões de espaço, não foram incluídos os valores de duração relativa, que podem ser encontrados nas subseções do Capítulo 4.

Tabela 49 - Resumo dos resultados para cada plosiva por parâmetro avaliado

	[p]	[t]	[k]	[b]	[d]	[g]
<b>Dur. cons.<sup>66</sup> (ms)</b>	127	125	128	101	94	93
<b>Dur.Oclusão (ms)</b>	104	101	74	89	81	70
<b>Pico (Hz)</b>	2039	2888	1734	1189	2324	1745
<b>Centroide (Hz)</b>	1172	2190	1671	896	1749	1095
<b>Variância<sup>67</sup> (Hz)</b>	*	*	*	*	*	*
<b>Assimetria</b>	5,9	1,7	4,4	6,0	2,6	5,8
<b>Curtose</b>	59	10	55	90	11	53
<b>VOT (ms)</b>	20	21	52	(-)92	(-)83	(-)70
<b>Aspiração (%)</b>	5	6	87	-	-	-
<b>F1 (Hz)</b>	463	421	456	450	433	434
<b>F2_1 (Hz): [e]</b>	1396	1761	1746	1440	1748	1890
<b>F2_1 (Hz): [u]</b>	1399	1567	1227	1377	1718	1064
<b>Dur. vogal [a]: VC<sup>68</sup></b>	144	178	162	168	172	177
<b>Dur. vogal [e]: CV</b>	63	51	50	68	58	73
<b>Dur. vogal [u]: CV</b>	49	37	35	56	43	54
<b>Dur. cons.: C[e]<sup>69</sup></b>	103	105	105	77	72	68
<b>Dur. cons.: C[u]</b>	117	105	113	88	83	78

<sup>66</sup> Dur. Cons. = duração total da consoante

<sup>67</sup> Obtivemos valores extremamente altos de variância, que não conseguimos apresentar na tabela acima. Para acessá-los, ver Tabelas 23 e 24.

<sup>68</sup> Aqui são reportados os valores médios (ms) da duração da vogal [a] diante de cada uma das consoantes plosivas mostradas na tabela.

<sup>69</sup> Aqui são reportados os valores médios da duração da consoante de cada uma das plosivas mostradas na tabela, de acordo com o contexto vocálico seguinte, a saber, [e] ou [u].

<b>VOT: C[ɐ]</b> <sup>70</sup>	16	20	39	(-)70	(-)64	(-)58
<b>VOT: C[u]</b>	26	23	50	(-)79	(-)75	(-)66
<b>Posição inicial: dur.cons. (ms)</b>	123	126	137	103	97	95
<b>Posição medial: dur.cons. (ms)</b>	130	125	120	98	89	86
<b>Posição inicial: VOT. (ms)</b>	21	20	55	(-)93	(-)87	(-)72
<b>Posição inicial: VOT. (ms)</b>	18	21	50	(-)89	(-)80	(-)67
<b>Sílaba tônica: dur.cons.(ms)</b>	137	136	136	111	104	99
<b>Sílaba átona: dur.cons.(ms)</b>	108	108	117	88	76	80
<b>Sílaba tônica: VOT(ms)</b>	19	19	53	(-)101	(-)93	(-)76
<b>Sílaba átona: VOT(ms)</b>	21	23	52	(-)78	(-)66	(-)58

Esses resultados mostram as características e peculiaridades encontradas em um falar do PB, apontando para aquilo que pode ser utilizado por falantes nativos dessa língua ao produzirem os segmentos plosivos. Vimos que os parâmetros de duração apresentaram características similares àqueles encontrados para outras pesquisas do PB, com falantes oriundos de outras regiões, mostrando que esses parâmetros podem ser específicos do PB.

No entanto, encontramos também resultados divergentes daqueles encontrados por outras pesquisas do PB para alguns parâmetros avaliados (por exemplo, transição de F1), o que pode sinalizar que eles

<sup>70</sup> Aqui são reportados os valores médios de VOT de cada uma das plosivas mostradas na tabela, de acordo com o contexto vocálico seguinte, a saber, [ɐ] ou [u].

são utilizados distintamente pelos diferentes falares do PB ou que ainda necessitam de uma investigação mais aprofundada para que conclusões mais concretas possam ser feitas. A Tabela 50 apresenta os principais resultados relacionados às especificidades acústicas encontradas dentre os nossos dados.

Tabela 50 - Resumo dos resultados para cada plosiva por ocorrência de especificidades acústicas

	[p]	[t]	[k]	[b]	[d]	[g]	Total
<b>Total de dados</b>	396	372	402	370	389	298	2227
<b>Com especificidades</b>	65	78	233	15	39	96	526
<b>Com especificidades (%)</b>	16%	21%	58%	4%	10%	32%	24%

Por fim, observamos que a fala é gradiente, no sentido de que ela pode apresentar especificidades que não são apontadas por livros e manuais. Observamos essa gradiência, por exemplo, nos resultados de VOT encontrados, que são “escolhidos” pelos falantes dentro de uma faixa de valores que caracteriza plosivas bilabiais e alveolares não-vozeadas como não-aspiradas e velares como levemente aspiradas, fato que comprova que a fala se dá dentro de um *continuum*. A fala é gradiente também, por mostrar que ela não é dicotômica, ou seja, por apresentar ou não certa característica, mas por apresentar singularidades ou especificidades distintas daquelas esperadas como padrão pela literatura. A língua está em constante modificação, fato mostrado pela quantidade considerável de dados que apresentaram eventos acústicos (as chamadas “especificidades) distintos do considerado como “típico” ou “padrão”.



## 5 CONCLUSÃO

Sem sombra de dúvida, a fonética acústica é uma área que, apesar de estar sendo estudada desde meados do século passado, é ainda um vasto campo da ciência a ser explorado. No PB, a demanda por pesquisas nessa área é ainda maior, tanto no campo segmental como no suprasegmental. Foi a partir da década de 80 que um maior número de estudos na área começou a ser desenvolvido (BEHLAU, 1986; KLEIN, 1999, dentre outros). Com o passar dos anos, mais e mais pesquisadores começaram a se interessar pela área e já podemos ver hoje o surgimento ou consolidação de periódicos relacionados a essa área de concentração, que impulsionam o desenvolvimento científico e acadêmico em Fonética e Fonologia no Brasil, através da interdisciplinaridade entre pesquisadores das áreas de Linguística, Fonoaudiologia, Ciência da Computação, Engenharia, dentre outras.

Apesar de a investigação fonética e acústica ter respaldo em diversos campos de pesquisa, tais como Linguística Forense, Linguística aplicada à LM ou à LE, Fonoaudiologia, Engenharia de Telecomunicações, dentre outros, é ainda reduzido o número de pesquisas e pesquisadores (especialmente, linguistas) que se aventuram por esse caminho para o PB, se comparado à grande quantidade de pesquisas para outras línguas como, por exemplo, o inglês e o francês.

Assim, pretendemos, com esta tese, dar continuidade a trabalhos realizados anteriormente na área de Fonética Acústica, almejando ampliar os conhecimentos sobre a produção dos segmentos plosivos no PB. Nosso objetivo primordial foi desenvolver uma investigação dos principais parâmetros acústicos utilizados na produção desses segmentos, bem como observar a possível influência de contextos linguísticos sobre eles, tendo como foco a fala de informantes monolíngues adultos, coletada com mulheres, de uma mesma região do Brasil e dentro de uma mesma faixa etária. Acreditamos que esses fatores conferem a esta tese um caráter de ineditismo, visto que grande parte dos estudos relacionados ao PB que coletou fala adulta, utilizou-a como comparativo (grupo controle) com a fala infantil, fala atípica ou ainda com outras línguas (em especial na comparação entre LM e LE) e não como foco principal de investigação.

## 5.1 IMPLICAÇÕES PARA A PRODUÇÃO DE PLOSIVAS

Nossa investigação observou parâmetros relacionados a aspectos de duração, dentre eles a duração da consoante, da oclusão, do VOT, do segmento e das vogais adjacentes; parâmetros frequenciais, como características espectrais de *burst* e a transição formântica (transição de F1 e F2); análise qualitativa da forma de onda e espectrograma das plosivas com o intuito de se observar possíveis especificidades acústicas e, por fim, possível influência de contextos linguísticos tais como posição na palavra e tonicidade.

A duração da consoante seguiu o apontado pela literatura, pois observamos diferenças estatísticas entre a duração absoluta e relativa entre plosivas vozeadas e não-vozeadas, sendo que as primeiras mostraram valores menores do que as últimas. Esse resultado se aplica a comparação entre as duas categorias de forma geral e entre cada um dos pares homorgânicos. A análise entre os diferentes pontos de articulação de uma mesma categoria na literatura mostra que, no caso das não-vozeadas, a ordem na duração da consoante é  $[t] < [p] < [k]$ , enquanto, para as vozeadas, é  $[g] < [d] < [b]$ . No entanto, os testes estatísticos mostraram resultados não significativos entre as não-vozeadas e somente entre  $[b]$  e as demais para as vozeadas. Dessa forma não ratificamos a ordem na duração entre consoantes não-vozeadas com diferentes pontos de articulação. Podemos confirmar apenas a ordem  $[b] > [t]$  e  $[g]$  para as plosivas vozeadas. Concluimos assim que a duração da consoante é um fator importante na distinção entre o vozeamento, mas não parece ser tão relevante para o ponto de articulação.

Com relação à duração da oclusão, testes estatísticos confirmaram que a diferença entre a duração absoluta e relativa da oclusão de plosivas vozeadas e não-vozeadas é significativa, sendo maior para as não-vozeadas. Esse resultado se estende à comparação entre os pares de plosivas homorgânicas. Com relação ao ponto de articulação, dentre as não-vozeadas, a velar foi a que apresentou menor valor médio, diferenciando-se estatisticamente das demais. Não houve diferença significativa entre a bilabial e a alveolar. Para as vozeadas, a bilabial foi a que apresentou maior valor médio de duração da oclusão, seguida por alveolar e velar. Todas as comparações apresentaram resultados significativos, indicando que, para essas consoantes, quanto mais posterior é o ponto de articulação, menor é o valor da duração da

oclusão e que este pode ser um fator importante na distinção entre esse grupo de plosivas.

Encontramos dentre os nossos dados as chamadas especificidades acústicas, conforme apontadas por Cristofolini (2008; 2013). Elas ocorreram em 24% dos dados. As especificidades com maior ocorrência dentre as não-vozeadas foram *edge vibration* (28% dos casos) e múltiplas explosões (48%), especialmente dentre as velares. No caso das vozeadas, as especificidades mais frequentes foram dupla explosão (26%) e interrupção do vozeamento na porção final do segmento (51%). Como avaliamos dados de fala controlada, acreditamos que esses percentuais possam ser ainda maiores em fala espontânea. Visto que quase um quarto dos nossos dados apresentaram características diferentes daquelas descritas pela literatura como típicas, podemos afirmar que esse fato corrobora pesquisas anteriores como Cristofolini (2008; 2013), Meneses (2012), McMurray *et. al* (2002), dentre outras, que mostram que esses fenômenos reafirmam o caráter gradiente e variável da fala. Assim, é preciso que se atente para tais fenômenos também na investigação de questões relacionadas a aspectos perceptuais da fala como, por exemplo, até que ponto definições de limiares entre categorias podem ser estabelecidos (a partir do julgamento perceptual de ouvintes) e como lidar com regiões *fuzzy* – regiões de limite de fronteiras – conforme apontado por Albano (2001), questão que está diretamente ligada ao que a autora aponta como “o problema da discretização dos contínuos físicos envolvidos na produção e percepção de fala” (ALBANO, 2001, p. 28).

Com relação às características espectrais de *burst*, avaliamos o pico e os momentos espectrais (centroide ou média, variância, assimetria e curtose). Segundo a literatura, o *burst* é distinto com relação ao ponto de articulação. Desse modo, comparamos, primeiramente, o pico espectral entre os diferentes pontos de articulação dentre as categorias vozeadas e não vozeadas. A estatística revelou que os valores de pico são significativamente distintos entre [k] e as demais não-vozeadas e entre [d] e as demais vozeadas. Plosivas bilabiais não-vozeadas apresentaram valores mais elevados do que a faixa esperada (500-1500 Hz), enquanto a bilabial vozeada se enquadrou dentro dessa faixa. As plosivas alveolares apresentaram valores de pico mais baixos do que o esperado (menores do que 4000 Hz). As duas plosivas velares apresentaram valores dentro da faixa esperada (1500-4000Hz).

Com relação às demais características espectrais, a centroide mostrou que a média de concentração de energia é, de fato, menor para as bilabiais e maior para as alveolares e foi significativa para todas as

comparações entre as plosivas não-vozeadas e entre as vozeadas. A variância não se mostrou um fator de influência nas plosivas, pois encontramos diferenças significativas somente entre alveolares e as demais. A assimetria e a curtose também foram distintas somente entre as alveolares. Também não observamos sistematicidade quanto à influência do contexto vocálico sobre as características espectrais de *burst*, pois a vogal [ɐ] atenuou os valores de pico espectral para as alveolares e a vogal [ʊ] para as velares.

O VOT, o parâmetro mais observado pelos pesquisadores que investigam as plosivas, seguiu as tendências apontadas por outras pesquisas. As plosivas não-vozeadas em nosso estudo foram incluídas na categoria de retardo curto (0-60 ms). Com relação à aspiração, [p] e [t] apresentaram, substancialmente, valores na faixa de 0-35 ms, o que as enquadra como não-aspiradas. A velar acabou extrapolando essa faixa, com 87% dos valores acima de 35 ms, dos quais 67% apresentavam VOT na faixa de 35-55 ms, sendo consideradas como levemente aspiradas, e 33% na faixa de 55-95 ms, sendo consideradas aspiradas. Testes estatísticos apontaram que o VOT de [p] e [t] não são distintos entre si, corroborando estudos anteriores para o PB, enquanto o de [k] difere significativamente dos demais. As vozeadas foram classificadas dentro da categoria de pré-sonorização, por apresentarem vibração de pregas antes do momento de soltura da oclusão, com valores negativos. Encontramos diferenças significativas entre o VOT de todas as consoantes vozeadas, indicando a seguinte ordem para o VOT dessas consoantes: [b] > [d] > [g].

É conhecido que a transição de F1 diferencia vozeamento, que plosivas vozeadas apresentam valores de F1 mais baixos do que as não-vozeadas e que, portanto, exibem transições mais acentuadas. Nossos resultados mostraram que todas as transições foram ascendentes, como esperado, indicando que não há alteração no formato da curva de transição quando se modifica a qualidade vocálica. Apesar disso, testes estatísticos não apontaram para diferenças significativas quanto à transição de F1 para as plosivas em análise. De acordo com esses resultados, acreditamos que a influência da transição de F1 na identificação de plosivas do PB é limitada para os nossos dados.

A transição de F2 é distinta conforme o ponto de articulação e, por esse motivo, avaliamos primeiramente a relação entre plosivas não-vozeadas e posteriormente entre as vozeadas.

Dentre as não-vozeadas seguidas por vogal [ɐ], observamos que [p] apresentou valores de frequência de *onset* de F2 significativamente

menores do que [t] e [k]. Com relação às mesmas plosivas seguidas pela vogal [u], não encontramos diferenças significativas entre [pʊ], [tʊ] e [kʊ]. A avaliação de cada uma das plosivas não-vozeadas de acordo com a vogal seguinte mostrou que: [pɐ] x [pʊ] não apresentaram diferença significativa nas suas frequências de onset de F2, com valores em torno de 1400 Hz, corroborando os resultados de Ficker (2003) para o PB. Os gráficos de transição se mostraram descendentes, diferentemente do encontrado na literatura. Quando a consoante alveolar é seguida pela vogal [ɐ], apresenta atenuação significativa, corroborando os resultados encontrados por Lehiste e Peterson (1961). Quando seguida por [ʊ], essa plosiva apresenta valores próximos ao que a literatura aponta como *locus* consonantal (em torno de 1800 Hz). O gráfico de curva de F2 para [tɐ] e [tʊ] mostraram-se descendentes, conforme descrito pela literatura. Por fim, a análise das frequências do *onset* de F2 das plosivas velares, mostrou diferenças significativas entre [kɐ] e [kʊ], com atenuação de frequências para a plosiva diante de vogal [ʊ], resultado que corrobora estudos anteriores.

A transição de F2 para as plosivas vozeadas, na análise entre os pontos para cada contexto vocálico, mostrou que, para as consoantes seguidas por [ɐ], a bilabial apresentou valor significativamente menor do que as demais e que, para as consoantes seguidas por [ʊ], a velar apresentou valores significativamente mais baixos do que as demais. Com relação à comparação de cada uma das plosivas entre os contextos vocálicos [ɐ] e [ʊ], observamos que: para [b] e [p], não houve diferença significativa entre os valores de *onset* de F2, com valores em torno de 1400 Hz, corroborando Ficker (2003) e mostrando que pode haver um *locus* consonantal de F2 para as bilabiais no PB. As curvas para a bilabial foram descendentes, diferentemente do esperado. Também não encontramos diferenças significativas na comparação entre [tɐ] e [tʊ], com os valores de frequência de *onset* de F2 se concentrando na região descrita como seu *locus* consonantal (1800 Hz), indo ao encontro dos achados de Ficker (2003) e apontando para a manutenção do *locus* do PB na mesma região que é apresentada para demais línguas. Por fim, a comparação entre [gɐ] e [gʊ] mostrou valores de *onset* de F2 significativamente menores para a plosiva seguida pela vogal posterior, indicando que a vogal posterior influencia na transição de F2 para essa consoante. As curvas de transição de F2 foram descendentes para a velar seguida pelas duas vogais em análise.

Observa-se assim que, no que tange os dois parâmetros frequenciais (transição de F1 e de F2), nossos resultados foram ao encontro de alguns aspectos relacionados pela literatura. Os casos em que não encontramos diferenças estatísticas entre as comparações, divergindo de alguns estudos, não desmerecem o teor dos resultados, mas mostram que esses casos são, possivelmente, parâmetros que talvez precisem de mais aprofundamento para o PB em trabalhos futuros (análises com um maior volume de dados e para outros dialetos), visto que são poucos aqueles que investigaram a transição de F1 e F2 do PB.

Avaliamos também a relação entre os segmentos plosivos do PB e as vogais adjacentes. A duração da vogal [a] que precede a plosiva (contexto VC) foi sempre maior quando ela era seguida por uma consoante vozeada, mostrando diferenças estatisticamente significativas entre a comparação com a mesma duração dessa vogal diante de uma plosiva não-vozeada. Na comparação entre os pares homorgânicos, somente as alveolares não apresentaram diferenças significativas entre si.

A duração da vogal que segue a plosiva (contexto CV) foi significativamente maior quando a vogal era precedida por consoante vozeada, tanto na comparação geral, quanto entre os pares homorgânicos.

No que tange à influência da vogal que sucede as plosivas, avaliamos os contextos [ \_ɐ ] e [ \_ʊ ]. Nossos resultados indicam que a duração da consoante e do VOT são maiores quando a plosiva é sucedida pela vogal alta (resultados significativos), com exceção de [t] e [b].

Outros contextos avaliados na produção das plosivas em nossas análises foram a posição que a plosiva ocupa na palavra e a tonicidade. Observamos que a posição que a plosiva ocupa na palavra foi importante para as plosivas vozeadas, mas não para as não-vozeadas. As primeiras apresentam maior duração da consoante e VOT quando estão inseridas em início de palavra, enquanto as últimas não apresentaram relação consistente entre maior duração de consoante e VOT conforme a posição que ocupam na palavra.

Por fim, observamos que as plosivas não-vozeadas apresentam maior duração quando inserida em sílaba tônica (em detrimento do contexto átono, sendo a diferença significativa para todas as plosivas dessa categoria). No entanto, a alveolar [t] foi a única que apresentou maior duração quando inserida em contexto átono. Para as vozeadas, tanto a duração da consoante quanto do VOT foram sempre

significativamente maiores quando essas consoantes estavam inseridas em sílabas tônicas. O Quadro 5 apresenta um resumo dos resultados encontrados para cada uma das hipóteses lançadas pela presente pesquisa.

Quadro 5 - Resultados para cada uma das hipóteses levantadas pela presente pesquisa

<b>Hipótese<sup>71</sup></b>	<b>Resultado</b>
<b>H1:</b> Haverá diferença entre as consoantes plosivas vozeadas e não-vozeadas, sendo que as últimas apresentarão valores de duração de consoante maiores do que as primeiras. Essa tendência se manterá na comparação entre os pares de plosivas homorgânicas.	confirmada
<b>H2.1:</b> Não haverá diferença entre os três pontos de articulação com relação à duração das plosivas não-vozeadas.	confirmada
<b>H2.2:</b> O ponto de articulação influenciará a duração das consoantes vozeadas, pois, na medida em que o ponto vai se tornando mais posterior, diminui-se a duração da consoante.	confirmada
<b>H3:</b> A duração da oclusão em plosivas não-vozeadas será maior do que o mesmo momento em plosivas vozeadas, mesmo na comparação entre plosivas homorgânicas.	confirmada
<b>H4.1:</b> Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.	parcialmente confirmada
<b>H4.2:</b> Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.	confirmada
<b>H5:</b> Os dados de falantes adultos apresentarão especificidades, sendo o tipo mais comum a presença de edge vibrations nas plosivas não-vozeadas e interrupções de vozeamento nas vozeadas.	confirmada

<sup>71</sup> Um resumo com as perguntas de pesquisa e suas respectivas hipóteses está inserido ao final do Capítulo 2.

<p><b>H6:</b> Os valores de características espectrais de burst se diferenciarão com relação ao ponto de articulação.</p>	parcialmente confirmada
<p><b>H7:</b> Existirá diferenças no burst das plosivas seguidas por diferentes vogais para o PB, sendo que plosivas seguidas por vogais posteriores podem apresentar maiores valores de frequência do que aquelas seguidas por vogais anteriores.</p>	parcialmente confirmada
<p><b>H8:</b> O VOT das plosivas não-vozeadas apresentará valores positivos e as plosivas não-vozeadas poderão ser incluídas na categoria de não-aspiradas ou levemente aspiradas, além da categoria de retardo curto. O VOT de plosivas vozeadas do PB apresentará valores negativos, indicando pré-sonorização.</p>	confirmada
<p><b>H9.1:</b> O valor da plosiva velar não-vozeada se diferenciará das demais plosivas, pois apresentará maiores valores de VOT, enquanto os valores de VOT poderão ser similares para as plosivas bilabiais e alveolares. As plosivas velares serão aquelas com mais chance de apresentar valores de VOT dentro da categoria de leve aspiração.</p>	confirmada
<p><b>H9.2:</b> Os valores de VOT para plosivas vozeadas seguirão a sequência [b] &gt; [d] &gt; [g], apresentarão valores negativos (por possuírem pré-sonorização) e diferentes entre si.</p>	confirmada
<p><b>H10:</b> Existirá diferença significativa entre o VOT das plosivas aspiradas e não-aspiradas.</p>	confirmada
<p><b>H11:</b> A região de F1 para plosivas vozeadas apresentará valores mais baixos do que para as não-vozeadas. Espera-se encontrar transições mais acentuadas para as plosivas vozeadas.</p>	parcialmente confirmada
<p><b>H12.1:</b> Se encontrará diferenças</p>	parcialmente confirmada

<p>significativas entre os valores de transição de F2 para os diferentes pontos de articulação. Ademais, haverá influência da vogal que segue a plosiva na transição de F2, sobretudo no caso das plosivas velares.</p>	
<p><b>H12.2:</b> Se encontrará diferenças significativas entre os valores de transição de F2 para os diferentes pontos de articulação. Ademais, haverá influência da vogal que segue a plosiva na transição de F2, sobretudo no caso das plosivas velares.</p>	parcialmente confirmada
<p><b>H13:</b> A duração das vogais que antecedem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.</p>	parcialmente confirmada
<p><b>H14:</b> A duração das vogais que seguem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.</p>	confirmada
<p><b>H15:</b> A duração da consoante e de seu VOT serão maiores quando a vogal adjacente é uma vogal alta (em detrimento da vogal baixa).</p>	parcialmente confirmada
<p><b>H16.1:</b> A posição na palavra poderá influenciar a duração total da consoante e VOT de plosivas não-vozeadas.</p>	parcialmente confirmada
<p><b>H16.2:</b> A posição na palavra poderá influenciar a duração total da consoante e VOT de plosivas vozeadas.</p>	confirmada
<p><b>H17.1:</b> Se o PB seguir a mesma tendência para outras línguas (KLATT, 1976), as consoantes em posição átona terão maior duração do que em posição tônica. No que tange o VOT, o contexto átono propiciará maiores valores de VOT no caso de bilabiais e alveolares e oônico propiciará a maior duração do VOT de consoantes velares.</p>	parcialmente confirmada

<p><b>H17.2:</b> Se o PB seguir a mesma tendência para outras línguas (KLATT, 1976), as consoantes em posição átona terão maior duração do que em posição tônica. Além disso, as plosivas vozeadas inseridas em sílabas tônicas apresentarão valores maiores de VOT.</p>	<p>parcialmente confirmada</p>
--	--------------------------------

Assim, podemos observar que muitos dos parâmetros apontados pela literatura como importantes na distinção da produção dos segmentos plosivos se confirmaram através de nossos resultados. No entanto, aqueles parâmetros cujos resultados não foram estatisticamente relevantes não invalidam pesquisas anteriores e nem são também invalidados. É preciso lembrar que cada pesquisa realizada com informantes é diferenciada e, por mais que o pesquisador tente amenizar diferenças extralinguísticas, cada participante possui características próprias que não conseguimos controlar nas pesquisas. Acreditamos que toda e qualquer comparação entre pesquisas feitas com seres humanos podem apresentar características comuns, no entanto também podem apresentar diferenças não atribuídas aos parâmetros controlados pelo pesquisador, mas sim devido às características intrínsecas a cada um dos participantes avaliados.

## 5.2 LIMITAÇÕES

Um dos percalços encontrados nesta pesquisa foi a grande quantidade de tempo investido em etiquetagem de dados. Nossa pesquisa avaliou 2227 dados que receberam 8 etiquetagens cada, totalizando mais de 17800 etiquetas marcadas manualmente. Esse, com certeza, foi um ponto limitador para esta pesquisa que, na sua forma inicial, pretendia avaliar também alguns aspectos da percepção dos segmentos plosivos, mas que devido ao tempo disponível, não puderam ser avaliados.

Além disso, outra dificuldade foi encontrar participantes jovens, que não falam outra língua, que moram na mesma cidade ou região onde adquiriram a linguagem e cujos pais também sejam oriundos dessa mesma região, para que pudéssemos controlar a influência desses fatores extralinguísticos na análise. A demanda por profissionais qualificados, que tenham experiência e formação adequada faz com que haja cada vez mais mobilidade entre as pessoas, não só de uma cidade

ou estado para outro, mas também de país para país. O pesquisador que objetiva minimizar as influências extralinguísticas que possam influenciar ou até mesmo enviesar seus resultados pode encontrar certa dificuldade para avaliar participantes, especialmente jovens, que estejam dispostos a participar de forma voluntária de pesquisas acadêmicas, pelos motivos acima relacionados.

### 5.3 INDICAÇÕES PARA PESQUISAS FUTURAS

Por fim, acreditamos que esta pesquisa tenha ajudado a entender melhor como funcionam os parâmetros acústicos relacionados à produção dos segmentos plosivos no PB. No entanto, não encontramos respostas definitivas para todas as questões lançadas aqui. Acreditamos que é parte da pesquisa acadêmica e do processo de construção do conhecimento científico tentar encontrar respostas, mas também lançar mais perguntas para que mais e mais pesquisas possam ser realizadas. As perguntas não respondidas em pesquisas anteriores propiciaram o desenvolvimento desta pesquisa.

Dessa forma, acreditamos que alguns pontos ainda mereçam mais investigação com relação à produção e também percepção de plosivas no PB, tais como:

- Aprofundamento de questões que envolvam a utilização de testes perceptuais para se observar se algum dos parâmetros aqui avaliados tem, realmente, influência na distinção dos segmentos plosivos no PB;
- Investigação aprofundada das chamadas “especificidades acústicas”, observadas por Cristofolini (2008; 2013) e por nós, através da quantificação e classificação de cada uma delas, seja na fala infantil ou adulta;
- Ampliação dos contextos vocálicos envolvidos nas análises de duração e qualidade da vogal adjacente, de transição vocálica e de *burst*;
- Ampliação do número de dados para todas as análises que envolvem questões relacionadas à frequência como *burst* e transição, pois são aqueles que apresentam um menor número de estudos para o PB e são, de acordo com a literatura, de grande importância para a percepção dos segmentos plosivos;
- Avaliação de dados sintetizados e de fala espontânea ou semi-espontânea, para que se possa observar se os

parâmetros acústicos e contextos linguísticos se comportam de maneira similar para esses tipos de fala e se têm a mesma influência na percepção dos segmentos plosivos;

- Estender ou replicar a presente pesquisa para outros dialetos do Brasil, objetivando avaliar se os resultados aqui obtidos são observados para outros falares no PB.
- Avaliar, na percepção de fala, o caráter variável e gradiente na produção de plosivas e, assim, problematizar a questão da definição de limites de categorias, considerando a presença do que Albano (2001) denominou fronteiras vagas ou regiões “fuzzy”.

#### 5.4 CONSIDERAÇÕES FINAIS

Na presente pesquisa, observaram-se as principais propriedades acústicas, possíveis influências de contextos linguísticos e ainda a presença de especificidades acústicas relacionadas à produção das plosivas no falar cricumense de indivíduos adultos. O objetivo foi delinear, descrever e avaliar quantitativamente cada um desses aspectos para um detalhamento aprofundado e também caracterização da produção de plosivas vozeadas e não-vozeadas do PB.

Levantamos também, ainda que brevemente, a questão da gradiência e variação na língua, mostrando, numericamente, que os segmentos de fala podem apresentar variação na sua produção, divergindo e questionando o que foi previamente estabelecido como “típico”.

Em uma pesquisa anterior<sup>72</sup>, observamos um universo variado de pesquisas sobre os mais diversos parâmetros acústicos relacionados às plosivas para línguas como o inglês, mas poucos, comparativamente, para o PB. Boa parte dos estudos para as plosivas do PB investigou o VOT, por esse ser um parâmetro bastante utilizado para se observar a produção das plosivas. Poucos são os estudos, nessa língua, que se debruçaram, por exemplo, sobre parâmetros relacionados à transição de F1 e F2, especialmente na fala adulta típica no PB. Foi a partir disso que optamos por desenvolver a presente pesquisa, almejando contribuir para a área de Fonética Acústica no Brasil e, talvez, servir como base para outras pesquisas nessa e em outras áreas, que venham a se interessar pela investigação de algum aspecto aqui abordado.

---

<sup>72</sup> Ver Alves (2011).

## REFERÊNCIAS

ABRAMSON, Arthur S. Laryngeal timing in consonant distinctions. **Phonetica**, v. 34, p. 295-303, 1977.

ABRAMSON, Arthur S.; LISKER, Leigh. **Voice onset time in stop consonants: acoustical analysis and synthesis**. Proceedings of the 5th International Congress of Acoustic. Liege: Imp. G. Thone, Paper A 51, 1965.

ABRAMSON, Arthur S.; LISKER, Leigh. Voice-timing perception in Spanish word-initial stops. **Journal of Phonetics**, v. 1, p. 1-8, 1973.

ALBANO, Eleonora. **O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro**. Campinas: Mercado das Letras, 2011.

ALVES, Mariane A. **Production of English and Portuguese Voiceless Stops by Brazilian EFL Speakers**. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2011.

ALVES, Mariane A.; DIAS, Eva C. O.. Estudo da produção do VOT em plosivas não-vozeadas diante de vogal alta posterior e anterior do português brasileiro. **Anais do IX Encontro do Círculo de Estudos Linguísticos do Sul**. Palhoça: Editora da Unisul, 2010.

ALVES, Mariane A.; SEARA, Izabel C.; PACHECO, Fernando S.; KLEIN, Simone. On the voiceless aspirated stops of Brazilian Portuguese. Lecture Notes in Artificial Intelligence, LNAI 5190, Proceedings of the 8th International Conference of the Portuguese Language. **Springer**, v. 5190, p. 248-51, 2008

BARBETTA, Pedro Alberto. **Estatística Aplicada às Ciências Sociais** (7a ed). Florianópolis: Editora da UFSC, 2010.

BARBOSA, Plínio Almeida. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on automatic segmental duration generation. **Caderno Estudos Linguísticos**, v. 31, p. 33-53, 1996.

BARROCO, Mário André Lopes; DOMINGUES, Maria Teresa Pedrosa; PIRES, Maria de Fátima Marques de Oliveira; LOUSADA, Marisa; JESUS, Luis M. T. Análise temporal das oclusivas orais do português europeu: um estudo de caso de normalidade e perturbação fonológica. **Rev. CEFAC**, v. 9, n. 2, p. 154-163, 2007.

BEHLAU, Mara S. **Análise do tempo de início da sonorização na discriminação de sons plosivos do português**. Tese de Doutorado. São Paulo: Escola Paulista de Medicina, Unifesp, 1986.

BENKÍ, José R. Place of articulation and first formant transition pattern both affect perception of voicing in English. **Journal of Phonetics**, v. 29, p. 1-22, 2001.

BERTI, Larissa Cristina. **Aquisição incompleta do contraste entre /s/ e /ʃ/ em crianças falantes do português brasileiro**. Tese de Doutorado. Campinas: Universidade Estadual de Campinas, 2006.

\_\_\_\_\_. Contrastes e contrastes encobertos na produção da fala de crianças. **Pró-fono Revista de Atualização Científica**, v. 22, n. 4, p. 531-536, 2010.

BERTI, Larissa Cristina; MARINO, Viviane Cristina de Castro. Contraste fônico encoberto entre /t/ e /k/: um estudo de caso de normalidade de transtorno fonológico. **Revista CEFAC**, v. 13, n. 5, p. 866-875, 2011.

BETTONI-TECHIO, Melissa. **Production of final alveolar stops in Brazilian Portuguese/English interphonology**. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2005.

BLUMSTEIN, Sheila E.; STEVENS, Kenneth N. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. **The Journal of the Acoustical Society of America**, v. 66, n. 4, p. 1001-17, out. 1979.

\_\_\_\_\_. Perceptual invariance and onset spectra for stop consonants in different vowel environments. **Journal of the Acoustical Society of America**, v. 67, n. 2, p. 648-62, 1980.

BOERSMA, Paul; WEENINK, David. **Praat - doing phonetics by computer**. Versão 5.3.39. Amsterdã: University of Amsterdam, 2013. Disponível em [www.praat.org](http://www.praat.org). Acesso em 27/05/2013.

BONATTO, Maria Teresa Rosângela Lofredo. A produção de plosivas por crianças de três anos falantes do português brasileiro. **Revista CEFAC**, v. 9, n.2, p.199-206, 2007.

BONATTO, Maria Teresa Rosângela Lofreda; MADUREIRA, Sandra. A study on perception and production of voicing contrasts in the speech of 3-year old children. **Revista CEFAC**, v. 11, n. 1, p. 67-77, 2009.

BRITO, A. T. **Análise fonético-fonológica do estudo do vozeamento, desvozeamento e não vozeamento de obstruintes na fala infantil normal e com desvios**. Tese de Doutorado. Belo Horizonte: Universidade Federal de Minas Gerais, 2000.

BRITTO, Ana Teresa Brandão de Oliveira e. **Estudo do contraste de vozeamento em sujeitos com e sem desvio fonológico**. Tese de Doutorado. Belo Horizonte: Pontifícia Universidade Católica de Minas Gerais, 2010.

BROWMAN, Catherine P.; GOLDSTEIN, Louis. Articulatory phonology: an overview. **Phonetica**, v. 49, n. 3-4, p. 155-180, 1992.

CEMIM, Juliana. **Aquisição do contraste do vozeamento: um estudo longitudinal**. Tese de Doutorado. Florianópolis: Universidade Federal de Santa Catarina, 2014.

CHO, Taehon; LADEFOGED, Peter. Variation and universals in VOT: evidence from 18 languages. **Journal of Phonetics**, v. 27, n. 2, p. 207-229, abr. 1999.

CLARK, John; YALLOP, Colin. **An introduction to phonetics and phonology**. 2a ed. Cambridge: Blackwell Publishing Ltd., 1995.

COHEN, Gustavo Vargas. **The VOT Dimension: a bi-directional experiment with English and Brazilian-Portuguese Stops**. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2004.

COOPER, Franklin S; DELATTRE, Pierre C.; LIBERMAN, Alvin M.; BORST, John M.; GERSTMAN, Loius J. Some experiments on the perception of synthetic speech sounds. **Journal of the Acoustical Society of America**, v. 24, n. 6, p. 597-606, 1952.

CRISTOFOLINI, Carla. **Trocas ortográficas : uma interpretação a partir de análises acústicas**. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 2008.

\_\_\_\_\_. **Gradiência na fala infantil: caracterização acústica de segmentos plosivos e fricativos e evidências de um período de “refinamento articulatorio”**. Tese de Doutorado. Florianópolis: Universidade Federal de Santa Catarina, 2013.

DELATTRE, Pierre C.; LIBERMAN, Alvin M.; COOPER, Franklin S. Acoustic Loci and Transitional Cues for Consonants. **Journal of Phonetics**, v. 27, n. 4, p. 769-773, 1955.

DENES, P. Effect of Duration on the Perception of Voicing. **Journal of the Acoustical Society of America**, v. 27, n. 4, p. 761-764, 1955.

DIAS, Eva C. O.; SEARA, Izabel C. Redução e apagamento de vogais átonas finais na fala de crianças e adultos de Florianópolis: uma análise acústica. **Letrônica**, v.6, n.1., p.71-93, 2013.

FANT, Gunnar. **Acoustic theory of speech production: with Calculations based on X-Ray Studies of Russian Articulations**. Amsterdam: Mouton & Co., 1960.

\_\_\_\_\_. Stops in CV-syllables. **Speech Sounds and Features**. Cambridge: MIT Press, p.110-139, 1973.

FICKER, Luisa Barzaghi. **Produção e percepção de plosivas do português brasileiro: um estudo fonético-acústico de um sujeito com deficiência auditiva**. Tese de Doutorado. São Paulo: Pontifícia Universidade Católica de São Paulo, 2003.

FIELD, Andy. **Discovering Statistics using SPSS**. 3a ed. London: SAGE Publications Ltd., 2009.

FISCHER-JØRGENSEN, E. Acoustical analysis of stop consonants. **Miscellanea Phonetica**, v. 2, p. 42-59, 1954.

FORREST, Karen; WEISMER, Gary; HODGE, Megan. Statistical analysis of word-initial voiceless obstruents: preliminary data. **The Journal of the Acoustical Society of America**, v. 84, n. 1, p. 115-23, 1988.

GIMSON, Alfred C. **An Introduction to the Pronunciation of English**. 3a ed. London: E. Arnold, 1980.

GOLDMAN-EISLER, Frieda. **Psycholinguistics: experiments in spontaneous speech**. New York: Academic Press, 1968.

HAGGARD, Mark; AMBLER, Stephen; CALLOW, Mo. Pitch as a voicing cue. **The Journal of the Acoustical Society of America**, v. 47, n. 2, p. 613-7, 1970.

HALLE, Morris; HUGHES, George W.; RADLEY, J. P. A. Acoustic properties of stop consonants. **Journal of the Acoustical Society of America**, v. 29, n. 1, p. 107-116, 1957.

HAYS, William L. **Statistics for the Social Sciences**. New York: Holt, Rinehart and Winston, Inc., 1973.

JAKOBSON, Roman O.; FANT, Gunnar.; HALLE, Morris. **Preliminaries to Speech Analysis: the distinctive features and their correlates**. Cambridge, MA.: The MIT Press, 1963.

JOHNSON, Kenneth. **Acoustic and auditory features**. 2a. ed. Malden, MA.: Blackwell Publishing Ltd., 2003.

JONGMAN, Allard; WAYLAND, Ratre; WONG, Sereba. Acoustic characteristics of English fricatives. **The Journal of the Acoustical Society of America**, v. 108, n. 3 Pt 1, p. 1252-63, 2000.

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

KEWLEY-PORT, Diane. Measurement of formant transitions in naturally produced stop consonant-vowel syllables. **The Journal of the Acoustical Society of America**, v. 72, n. 2, p. 379-89, 1982.

KLATT, Dennis H. Interaction between two factors that influence vowel duration. **Journal of the Acoustical Society of America**, v. 54, n. 4, p. 1102-1104, 1973.

\_\_\_\_\_. Voice onset, frication, and aspiration in word-initial consonant clusters. **Journal of Speech, Language, and Hearing Research**, v. 18, p.686-706, 1975.

\_\_\_\_\_. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. **Journal of the Acoustical Society of America**, v. 59, n. 5, p. 1208-1221, 1976.

KLEIN, Simone. **Estudo do VOT no Português Brasileiro**. Dissertação de Mestrado. Florianópolis: Universidade Federal de Santa Catarina, 1999.

KLOKER, D. Vowel and sonorant lengthening as cues to phonological phrase boundaries. **Journal of the Acoustical Society of America**, v.57, 533(A), 1975.

LABOV, William. **Padrões sociolinguísticos**. São Paulo: Parábola, 2008.

LADEFOGED, Peter. **A course in phonetics**. 4a. ed. Boston, MA: Heinle & Heinle, 2001.

LADEFOGED, Peter; JOHNSON, Kenneth. **A course in phonetics**. 6. ed. Boston, MA: Wadsworth, 2011.

LADEFOGED, Peter. MADDIESON, Ian. **The sounds of the world's languages**. Oxford, UK: Blackwell Publishing Ltd., 1996.

LARSON-HALL, Jenifer. **A Guide to Doing Statistics in Second Language Research Using SPSS**. New York: Routledge, 2010.

LEHISTE, Isle. Some factors affecting the duration of syllabic nuclei in English. **Proceedings of the First Salzburg Conference on Linguistics**, p.81-104, 1975.

LEHISTE, Isle, PETERSON, Gordon. E. Transitions, glides and diphthongs. **Journal of the Acoustical Society of America**, v.33, p.268-277, 1961.

LIBERMAN, Alvin. M.; DELATTRE, Pierre. C.; COOPER, Franklin. S. The role of selected stimulus-variables in the perception of the unvoiced stop. **The American Journal of Psychology**, v. 65, p. 497-516, 1952.

\_\_\_\_\_. Some cues for the distinction between voiced and voiceless stops in initial position. **Language and Speech**, v. 1, n. 3, p. 153-167, 1958.

LIBERMAN, Alvin .; DELATTRE, Pierre C.; COOPER, Franklin S.; GERSTMAN, Louis J. The role of consonant-vowel transitions in the perception of stop and nasal consonants. **Psychological Monographs: general and applied**, v. 68, n. 8, p. 1-13, 1954.

\_\_\_\_\_. Tempo of frequency change as a cue for distinguishing classes of speech sounds. **Journal of Experimental Psychology**, v. 52, n. 2, p. 127-137, 1956.

LIBERMAN, Alvin M.; HARRIS, Katherine Safford; HOFFMAN, Howard S.; GRIFFITH, Belder C. The discrimination of speech sounds within and across phoneme boundaries. **Journal of Experimental Psychology**, v. 54, n. 5, p. 358-368, 1957.

LIEBERMAN, Philip.; BLUMSTEIN, Sheila E. **Speech physiology, speech perception and acoustic phonetics**. Cambridge: Cambridge University Press, 1988.

LISKER, Leigh. Closure duration and the intervocalic voiced-voiceless distinction in English. **Language**, v. 33, n. 1, p. 42-49, 1957.

\_\_\_\_\_. On explaining vowel duration variation.pdf. **Glossa: an international journal of linguistics**, v. 8, n. 2, p. 234-246, 1974.

\_\_\_\_\_. Is it VOT or a first-formant transition detector? **Journal of the Acoustical Society of America**, v. 57, n.6, p.1547-1551, 1975.

\_\_\_\_\_. On generalizing the rapid-rapid distinction based on silent gap duration. **Status**, v. 65, p. 251-259, 1981.

LISKER, Leigh.; ABRAMSON, Arthur. S. A cross-language study of voicing in initial stops: acoustical measurements. **Word**, v. 20, n. 3, p. 384-422, 1964a.

\_\_\_\_\_. Stop Categorization and Voice onset time. **Proceedings of the Fifth International Congress of Phonetic Sciences**, p.389-391, 1964b.

\_\_\_\_\_. Voice Onset Time in the Production and Perception of English Stops. **Proceedings of the 39th Annual Meeting of the Linguistic Society of America**, p. 3.1-3.15, 1964c.

\_\_\_\_\_. Some effects of context in voice onset time in English stops. **Language and Speech**, v. 10, p. 1-28, 1967.

\_\_\_\_\_. The voicing dimension: some experiments in comparative phonetics. **Proceedings of the Sixth International Congress of Phonetic Sciences 1967**, p.563-567, 1970.

\_\_\_\_\_. Distinctive features and laryngeal control. **Language**, v. 47, n. 4, p. 767-785, 1971.

LÖFQVIST, Anders. Interarticulator programming in stop production. **Journal of Phonetics**, v. 8, p. 475-490, 1980.

LOUSADA, Marisa; MARTINS, Paula; JESUS, Luis M. T. Estudo do pré-vozeamento, frequência do burst e locus de F2 das oclusivas orais do português europeu. **XXI Encontro Nacional da Associação Portuguesa de Linguística**. Lisboa, p.485-494, 2006.

MALMBERG, Bertil. **Manual of Phonetics**. London: North-Holland, 1968.

MARTINS, Carla. **Manual de Análise de Dados Quantitativos com Recurso ao IBM SPSS: saber decidir, fazer, interpretar e redigir**. Braga, Portugal.

McMURRAY, Bob; TANENHAUS, Michael K., ASLIN, Richard N. Gradient effects of within-category phonetic variation on lexical access. **Cognition**, v.86, B33-B42, 2002.

MELO, Roberta M.; MOTA, Helena B.; Mezzomo, Carolina L.; BRASIL, Brunah de C.; LOVATTO, Liane; ARZENO, Leonardo. Caracterização acústica da sonoridade dos fones plosivos do português brasileiro. **Revista CEFAC** v.14, n. 1, 2012.

- MENESES, Francisco de O. As vogais desvozeadas no português brasileiro: investigação acústico-articulatória. Dissertação de Mestrado. Campinas: IEL, 2012.
- NEWELL, K. M.; HANCOCK, P.A. Forgotten moments: a note on skewness and kurtosis as influential factors in inferences extrapolated from response distributions. **Journal of Motor Behavior**, v. 16, n. 3, p.320-335, 1984.
- ÖHMAN, S. Coarticulation in VCV utterances:spectrographic measurements. **Journal of the Acoustical Society of America**, v.39, p.151-168, 1966.
- OLLER, D. K. The effect of position in utterance on speech segment duration in English. **Journal of the Acoustical Society of America**, v.54, p.1235-1247, 1973.
- PEREIRA, Lílian. C. K.; MADUREIRA, Sandra. A produção das plosivas alveolares /t/ e /d/ por um sujeito com deficiência auditiva: um estudo fonético-acústico. **Revista Intercâmbio**, v. 23, p. 128-151, 2011.
- PORT, Robert. F.; ROTUNNO, Rosemaire. Relation between voice-inset time and vowel duration. **Journal of the Acoustical Society of America**, v. 66, n. 3, p. 654-662, 1979.
- RUSSO, Ieda; BEHLAU, Mara. **Percepção de fala: análise acústica**. São Paulo: Lovise, 1993.
- SEARA, Izabel. C. **Estudo acústico-perceptual da nasalidade das vogais do português brasileiro**. Tese de Doutorado. Florianópolis: Universidade Federal de Santa Catarina, 2000.
- SEARA, Izabek. C.; NUNES, Vanessa. G.; LAZZAROTTO-VOLCÃO, Cristiane. **Para conhecer: Fonética e fonologia do português brasileiro**. São Paulo: Contexto, 2015.
- SNOEREN, Natalie D. **Variations phonologiques em production et perception de la parole: le phenomene de l'assimilation**. Tese de Doutorado. Paris: Université Paris 5, 2005.

SPSS Inc. **SPSS Statistics for Windows**, Version 17.0. Chicago: SPSS Inc, 2008.

STEVENS, Kenneth. N. **Acoustic Phonetics**. Cambridge, MA.: The MIT Press, 2000.

STEVENS, Kenneth. N.; KLATT, Dennis. H. Role of formant transitions in the voiced-voiceless distinction for stops. **The Journal of the Acoustical Society of America**, v. 55, n. 3, p. 653-9, mar. 1974.

TRAINOTTI, Kamila. T.; MARGOTTI, Felício. W. Empréstimos lexicais do italiano no português do sul do Brasil e suas implicações em sala de aula. **Linguagens - Revista de Letras, Artes e Comunicação**, v. 2, n. 3, p. 263-274, 20008.

UMEDA, Noriko. Another consistency in phoneme duration. **Journal of the Acoustical Society of America**, v. 58, p. S62(A), 1975.

\_\_\_\_\_. Consonant duration in American English. **Journal of the Acoustical Society of America**, v. 61, n. 3, p. 846-858, 1977.

VELOSO, João. Vozeamento, duração e tensão nas oposições de sonoridade das oclusivas orais do português. **Revista da Faculdade de Letras: Línguas e Literaturas**, v. XIV, p. 59-80, 1997.

WOODS, Anthony; FLETCHER, Paul; HUGHES, Arthur. **Statistics in Language Studies**. Cambridge: Press Syndicate of the University of Cambridge, 1986.

YAVAS, Mehmet. **Applied English Phonology**. Hoboken, NJ: John Wiley and Sons, 2011.

ZAHORIAN, Stephen. A.; NOSSAIR, Zaki. B.; COLEMAN, R. F. Evidence against acoustic invariance in initial voiced stop consonants. **Journal of the Acoustical Society of America**, v. 81, n. S1, p. S36, 1987.

## APÊNDICE A – Questionário biográfico e Termo de Consentimento

Esse questionário tem como objetivo coletar informações sobre os seus dados biográficos para que a pesquisadora tenha acesso às informações pertinentes ao objetivo da tese. Lembramos que os nomes de todos os participantes serão preservados; os participantes serão indicados na tese através de números. Lembramos também que a sua participação é fundamental para que pesquisas sobre a nossa língua sejam viabilizadas. Obrigada!!

### Perguntas:

- 1) Quando e em que cidade você nasceu?
  - 2) Você reside na cidade onde nasceu?
  - 3) Já residiu em algum outro lugar? Por quanto tempo?
  - 4) Qual a sua idade?
  - 5) Qual o seu grau de escolaridade?
  - 6) Você já foi diagnosticado com algum problema auditivo ou possui alguma dificuldade para escutar?
  - 7) Tem alguma dificuldade com leitura?
  - 8) Já detectou, em qualquer fase da vida, algum problema com a sua voz?
  - 9) Alguma vez, durante a infância, ou ainda hoje, você trocava as letras na fala ou escrita?
  - 10) Você fala alguma outra língua além do português? Com quantos anos você começou a aprender essa língua?
- 

### Termo de Consentimento Livre e Esclarecido:

Eu, \_\_\_\_\_, declaro participar por livre e espontânea vontade como participante dos testes de produção ou de percepção para a pesquisa intitulada “Estudo acústico-articulatório da produção e percepção de plosivas surdas e sonoras do português brasileiro”, conduzida pela doutoranda Mariane Antero Alves, do PPGLg/UFSC. Estou ciente de que a pesquisa foi aprovada pelo Comitê de Ética, processo 2057, fr:434924, e que meu nome não será divulgado.

\_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de 2013.

Ass: \_\_\_\_\_



**APÊNDICE B – *Corpus* com as palavras contendo os segmentos-alvo para o teste de produção**

Plosiva	Sílabas no vocábulo	Posição na palavra	Vogal posterior	Tonicidade	
				Tônica	Átona
[p]	Dissílabo	Início	[a]	paga	patê
			[e]	pena	pegou
			[ɛ]	pele	-
			[i]	pino	pinhão
			[o]	poço	portão
			[ɔ]	posso	-
			[u]	pude	puxão
[p]	Dissílabo	Meio	[a]	capaz	napa
			[e]	ipê	-
			[ɛ]	tripé	-
			[i]	tupi	gripe
			[o]	capô	-
			[ɔ]	cipó	-
			[u]	xampu	topo
[p]	Trissílabo	Início	[a]	página	pateta
			[e]	pêssego	pecado
			[ɛ]	péssimo	-
			[i]	pícolo	pinóquio
			[o]	pômulo	polícia
			[ɔ]	pólvora	-
			[u]	público	pubica
[p]	Trissílabo	Meio	[a]	apaga	apagar
			[e]	apelo	inspeção
			[ɛ]	aperta	-
			[i]	rapina	rápido
			[o]	oposto	empossa r
			[ɔ]	-	-
			[u]	tapume	crápula
[p]	-	-	[a]	-	napa

				-	capo
				-	rapo
				-	capa
				-	tapa
				-	sapo
[t]	Dissílabo	Início	[a]	tapa	tabu
			[e]	tema	tenaz
			[ɛ]	tela	-
			[i]	-	-
			[o]	todo	total
			[ɔ]	toca	-
			[u]	tudo	tumor
[t]	Dissílabo	Meio	[a]	guatá	bata
			[e]	patê	-
			[ɛ]	pitel	-
			[i]	-	-
			[o]	chatô	-
			[ɔ]	totó	-
			[u]	itu	gato
[t]	Trissílabo	Início	[a]	tábula	tabela
			[e]	tênuê	temaki
			[ɛ]	térmico	-
			[i]	-	-
			[o]	tônico	tomilho
			[ɔ]	tórrida	-
			[u]	túmulo	tumulto
[t]	Trissílabo	Meio	[a]	batata	cítara
			[e]	antena	cátedra
			[ɛ]	pateta	-
			[i]	-	-
			[o]	contorno	átono
			[ɔ]	cartola	-
			[u]	batuque	rótulo
[t]	-	-	[a]	-	cata
				-	nata
				-	nato

				-	data
				-	gato
				-	mato
[k]	Dissílabo	Início	[a]	cata	cavar
			[e]	queijo	quebrar
			[ɛ]	quero	-
			[i]	quilo	quiser
			[o]	como	comer
			[ɔ]	colo	-
			[u]	cunha	kumon
[k]	Dissílabo	Meio	[a]	marcar	vaca
			[e]	porquê	-
			[ɛ]	sequer	-
			[i]	caqui	roque
			[o]	cocô	-
			[ɔ]	bocó	-
			[u]	pacu	saco
[k]	Trissílabo	Início	[a]	cátedra	caneta
			[e]	quênia	-
			[ɛ]	quéchua	-
			[i]	química	quimono
			[o]	cômodo	comando
			[ɔ]	cósmico	-
			[u]	cúmulo	cultura
[k]	Trissílabo	Meio	[a]	escada	ácaro
			[e]	esquema	-
			[ɛ]		-
			[i]	esquilo	líquido
			[o]	escopo	rococó
			[ɔ]	rococó	-
			[u]	escudo	cálculo
[k]	-	-	[a]	-	vaca
				-	paca
				-	saca
				-	faca
				-	saco

				-	caco
[b]	Dissílabo	Início	[a]	baga	babá
			[e]	beco	beber
			[ɛ]	bela	-
			[i]	bico	bidê
			[o]	bolo	bocó
			[ɔ]	bola	-
			[u]	bucho	bujão
[b]	Dissílabo	Meio	[a]	babá	naba
			[e]	saber	-
			[ɛ]	timbé	-
			[i]	subir	cabe
			[o]	tambor	-
			[ɔ]	timbó	-
			[u]	tabu	nabo
[b]	Trissílabo	Início	[a]	básico	bacana
			[e]	bêbado	beleza
			[ɛ]	bélico	-
			[i]	bígamo	bicudo
			[o]	-	bolacha
			[ɔ]	bócio	-
			[u]	bússola	buzina
[b]	Trissílabo	Meio	[a]	-	bárbaro
			[e]	cabelo	cobertor
			[ɛ]	aberto	-
			[i]	cabine	jubilar
			[o]	abono	abolir
			[ɔ]	aborta	-
			[u]	abutre	abusar
[b]	-	-	[a]	-	naba
				-	cabo
				-	rabo
				-	baba
				-	gaba
				-	nabo
[d]	Dissílabo	Início	[a]	data	datar

			[e]	dedo	dedar
			[ɛ]	dela	-
			[i]	-	-
			[o]	dolo	dosar
			[ɔ]	doca	-
			[u]	duna	durar
[d]	Dissílabo	Meio	[a]	nadar	cada
			[e]	cadê	-
			[ɛ]	puder	-
			[i]	-	-
			[o]	odor	
			[ɔ]	xodó	-
			[u]	vodu	todo
[d]	Trissílabo	Início	[a]	dádiva	damasco
			[e]		decote
			[ɛ]	débito	-
			[i]	-	-
			[o]	-	domador
			[ɔ]	dórico	-
			[u]	dúvida	duquesa
[d]	Trissílabo	Meio	[a]	adaga	indagar
			[e]	modelo	adesão
			[ɛ]	cadela	-
			[i]	-	-
			[o]	adorno	adoçar
			[ɔ]	desdobra	-
			[u]	caduco	deduzir
[d]	-	-	[a]	-	cada
				-	nada
				-	nado
				-	fada
				-	fado
				-	fado
[g]	Dissílabo	Início	[a]	gata	gabar
			[e]	gueto	

			[ɛ]	guerra	-
			[i]	guizo	guiar
			[o]	gota	gostar
			[ɔ]	gosta	-
			[u]	gula	guri
[g]	Dissílabo	Meio	[a]	tragar	vaga
			[e]	erguer	-
			[ɛ]	migué	-
			[i]	seguir	sangue
			[o]	xangô	-
			[ɔ]	mongol	-
			[u]	sagu	tango
[g]	Trissílabo	Início	[a]	gárgula	galinha
			[e]	-	guerrilha
			[ɛ]	-	-
			[i]	-	guitarra
			[o]	gônada	gostoso
			[ɔ]	gótico	-
			[u]	-	gutural
[g]	Trissílabo	Meio	[a]	agarra	agarrar
			[e]	portu- guês	-
			[ɛ]	aluguel	-
			[i]	esguia	enguiçar
			[o]	begônia	agonia
			[ɔ]	mongóli a	-
			[u]	agulha	gárgula
[g]	-	-	[a]	-	vaga
				-	paga
				-	saga
				-	baga
				-	lago
				-	pago

## APÊNDICE C – Script para a extração de dados relativos à transição vocálica

```

# Geracao de tabela com frequencias formantes
# Entrada:
# arquivo .wav
# arquivo .TextGrid com marcacao de segmentos a serem analisados
# Processamento:
# Analise de formantes no arquivo .wav original
# Analise de intensidade no arquivo .wav original
# Selecao de 5 pontos de analise em cada segmento
# Extracao da intensidade em cada um dos pontos
# Extracao de F1, F2 e F3 em cada um dos pontos
# Saida:
# Arquivo no formato txt (campos separados por tabulacao)
# com dados obtidos no processamento
# Nome do arquivo de saida eh igual ao de entrada, com extensao .txt
#
# Fernando S. Pacheco
# LINSE/UFSC
#
nFORMANTES=3
#a linha acima corresponde ao número de formantes que a análise vai
apresentar: F1, F2, F3
nPONTOS=5
#a linha acima corresponde ao número de pontos que ele vai pegar pra cada
segmento
select all
if numberOfSelected() > 0
    Remove
endif
form Arquivo a processar (extração de formantes)
    word Folder_(Diretório) I:\2013\MARIANE\
    word File_(Arquivo_áudio_com_extensão) DadosMari_P3.wav
endform
#folder$=""
#file$="teste1.wav"
fil$ = folder$ + file$
Read from file... 'fil$'
filename$ = selected$ ("Sound")
filegrid$ = filename$ + ".TextGrid"
filegrid$ = folder$ + filegrid$
Read from file... 'filegrid$'
select Sound 'filename$'
To Formant (burg)... 0.0 4 10000 0.025 50

```

```

select Sound 'filename$'
To Intensity... 100 0.0 no
select Sound 'filename$'
plus TextGrid 'filename$'
Extract non-empty intervals... 5 yes
#o número que aparece na linha acima corresponde ao tier, para buscar os dados
no tier 1, coloque 1 e assim por diante
nselected = numberOfSelected ("Sound")
#nao vou mais precisar do arquivo de audio
select Sound 'filename$'
plus TextGrid 'filename$'
Remove
#
#limpar janela info
clearinfo
printline N_SEG; NOME; DUR(s); INST_ANALISE(s);
INTENSID(dB); F1(Hz); F2(Hz); F3(Hz);
INST_ANALISE(s); INTENSID(dB); F1(Hz); F2(Hz);
F3(Hz); INST_ANALISE(s); INTENSID(dB); F1(Hz);
F2(Hz); F3(Hz) INST_ANALISE(s);
INTENSID(dB); F1(Hz); F2(Hz); F3(Hz);
INST_ANALISE(s); INTENSID(dB); F1(Hz); F2(Hz);
F3(Hz);

#contador de segmentos
i = 1
while i <= nselected
select all
soundID = selected ("Sound", 'i')
select 'soundID'
nomeseg$ = selected$ ("Sound")
tp_ini = Get starting time
tp_fim = Get finishing time
duracao = Get total duration

#definicao dos pontos de analise
#inicio do segmento
tp1 = tp_ini
#meio do segmento
tp2 = (tp_ini+tp_fim)/2
tp3 = (tp_ini+tp2)/2
tp4 = (tp_fim+tp2)/2
#fim do segmento
tp5 = tp_fim

```

#Os formantes são obtidos a partir do arquivo original e não de cada segmento. Assim, evita-se o problema com as bordas dos segmentos

```

for k from 1 to nPONTOS
  tp_analise = tp^k'
  if k == 1
    print 'i"tab$ 'nomeseg$' 'tab$' 'duracao' 'tab$' 'tp_analise'
  else
    print 'tab$' 'tp_analise'
  endif
for n from 1 to nFORMANTES
  select Formant 'filename$'
  fn = Get value at time... 'n' 'tp_analise' Hertz Linear
  select Intensity 'filename$'
  in = Get value at time... 'tp_analise' Cubic
  #separacao seguinte entre 1o. e outros formantes apenas para formatacao
na impressao
  if n == 1
    print 'tab$' 'in' 'tab$' 'fn'
  else
    print 'tab$' 'fn'
  endif
  #n = numero do formante
  #tp = tempo
endfor
endfor
print 'newline$'
i=i+1
endwhile
#criar arquivo de saida
#tabela no formato cvs
fileout$ = folder$ + filename$ + ".txt"
#apaga arquivo (se existente)
filedelete 'fileout$'
#copia conteudo da janela info para o arquivo
fappendinfo 'fileout$'

```





## APÊNDICE D – A Estatística aplicada na validação de parâmetros acústicos em análise de fala

Nesta seção, pretendemos refletir e relatar acerca de três pontos relacionados à estatística.

Em primeiro lugar, gostaríamos de demonstrar, através do relato (informal) da vivência da pesquisadora, como é possível aprender, entender e aplicar a estatística na área de Letras. Trazemos um breve resumo da trajetória da pesquisadora para que um leitor iniciante na área da estatística possa se sentir encorajado a seguir em frente e se aprofundar cada vez mais sobre o tema em questão.

Em seguida, discorreremos sobre a importância do uso de estatística (descritiva e inferencial) para os estudos na área de Letras, mais especificamente, na área de Fonética Acústica. No nosso ponto de vista, esta reflexão é relevante para qualquer pesquisador que queira conduzir pesquisas acadêmicas (e, portanto, fazer ciência) de maneira séria e comprometida, ao utilizar dados de ordem quantitativa, ou seja, que envolva a coleta de dados com características que possam ser medidas (LARSON-HALL, 2010). Esse processo se dá através da análise de dados e da inferência/reflexão sobre os resultados, respaldados pela estatística.

Logo após esta breve reflexão, pretendemos mostrar alguns pontos-chave na teoria estatística, a partir de espécie de roteiro, para ajudar os leitores a compreender alguns conceitos que ajudam a descrever os dados (Estatística Descritiva) e que também auxiliam no processo de escolha dos testes estatísticos para análise de dados (Estatística Inferencial), para que se possa compreender as escolhas feitas para esta pesquisa. Vale salientar, contudo, que não objetivamos que o leitor se torne um *expert* em estatística ao ler esse roteiro, mas sim que seja capaz de seguir o texto, entendendo as escolhas estatísticas e as decorrentes análises de resultados feitas pela pesquisadora.

Por fim, apresentaremos as características do conjunto de dados coletados para a presente pesquisa que foram observadas através de testes de normalidade e homogeneidade, bem como explicar a escolha dos testes estatísticos utilizados para a análise dos dados.

A) “SENTA QUE LÁ VEM HISTÓRIA”<sup>73</sup>

Nesta seção, iremos deixar de lado, brevemente, um pouco do rigor acadêmico e discorrer sobre o nosso caminho pela área da Estatística, com o intuito de auxiliar aqueles que, por ventura, possam não estar familiarizados com o tema. Deixamos um pouco de lado a pesquisadora e assumimos um papel de futura professora que gostaria de deixar algo a mais (ou, pelo menos, uma “pulgã atrás da orelha”) daqueles estudantes que irão trabalhar com análise quantitativa de dados na área de Linguística (e não só Fonética).

Ao iniciar o caminho por essa área, várias vezes nos deparamos com artigos, dissertações ou teses que se utilizavam de estatística para analisar os seus dados e, como desconhecemos muitos termos e conceitos, não conseguimos acompanhar a leitura dos textos. Isso foi algo muito frustrante, pois tinha a intenção de seguir para um mestrado e doutorado e vi que ainda havia muito a aprender. No entanto, isso não foi um fator desmotivador. Pelo contrário, foi um ponto de virada que me fez ir à busca de conhecimento para que pudesse, enfim, começar a ler as pesquisas sem ter que “pular” qualquer informação.

Para começar nossa conversa, gostaria de explicar o que me motivou a escrever este capítulo (que podia facilmente não ser incluído na tese). Minha experiência com estatística começou ainda na graduação quando comecei a me aventurar pelas leituras de artigos científicos, dissertações e teses para conduzir trabalhos de iniciação científica. Chegando à análise dos dados, tinha de fazer uma leitura rápida, pois, quando me deparava com algo como “ $U = 0,623$ ” ou “ $p = 0,003$ ”, não fazia ideia do que aqueles números poderiam significar. Corria os olhos até encontrar a informação explícita sobre qual resultado o autor tinha encontrado. Isso não tinha nenhum problema maior, visto que nenhum graduando em Letras precisa, obrigatoriamente, dominar os conceitos de estatística, até mesmo porque essa não costuma ser uma cadeira obrigatória no currículo dos cursos de graduação dessa área.

Foi somente durante a escritura da dissertação de mestrado que senti realmente que não poderia ficar me baseando apenas em média, mediana e desvio padrão, ou seja, apenas na análise descritiva dos dados, mas precisava ir além e começar a entender sobre Estatística Inferencial (análise de dados baseados em testes estatísticos).

---

<sup>73</sup> Nome de um quadro do programa infantil “Rá-Tim-Bum”, da TV Cultura, que foi ao ar nos anos 90.

Assim, comecei a buscar cursos dentro da própria universidade e em outras instituições para começar a entender como aplicar estatística aos dados de Fonética. A primeira vista, achei que aqueles universos de números e letras não faziam muito sentido juntos, mas, com a ajuda de bons professores e colegas, fui enfrentando os desafios até que um dia as coisas começaram a fazer sentido. Para mim, foi como aprender uma língua estrangeira: no começo o aluno se frustra, passa uns 2 ou 3 anos fazendo aulas, falando ou entendendo pouco, mas, de repente, como em um *insight*, as coisas começam a fazer sentido e, quando se vê, está falando razoavelmente bem e entendendo quase tudo que o professor fala. No meu caso, a relação com estatística foi um pouco mais aprofundada em um curto período de tempo, pois precisava aplicar esses conhecimentos às análises.

Ao entrar no doutorado, consultei alguns livros sobre Estatística<sup>74</sup> e, a partir deles, comecei um estudo intenso sobre o tema que culminou com a escritura deste capítulo e com a aplicação de testes estatísticos aos dados sem precisar de auxílio de um estatístico. Algo que, no início, ainda na graduação, achei que jamais seria possível.

Gostaria que esse breve relato da minha experiência com a estatística possa servir de inspiração para que o leitor que esteja começando essa jornada não desista no meio do caminho e veja que é possível que um aluno das Letras entenda de estatística e que aplique esses conhecimentos nas suas pesquisas. Na seção abaixo, vamos refletir um pouco sobre o uso de estatística em pesquisas com Linguística e Linguística Aplicada.

## B) POR QUE ESTUDAR ESTATÍSTICA QUANDO ESCOLHI ESTUDAR LETRAS?

Essa questão fazia parte do meu íntimo durante anos (e, às vezes, ainda faz, principalmente quando estamos inseridos num mar de tabelas, gráficos, análises e números). “Ora, mas se eu escolhi estudar Letras é porque eu odeio números!”. Não necessariamente. Se você escolheu Letras para estudar Literatura, é provável que você não necessite de estatística, mas acreditamos que se você se interessou por Linguística,

---

<sup>74</sup> Ver Barbetta (2010), Martins (2011), Woods *et al.* (1986), Hays (1973), Larson-Hall (2010) e Field (2009).

uma boa parte das subáreas pode (e, porque não, deve!) usar estatística em suas análises, mesmo que seja Estatística Descritiva.<sup>75</sup>

Uma das grandes contribuições da estatística para qualquer área é sem dúvida, a capacidade que ela tem de sintetizar os resultados obtidos com as análises dos dados. Tomemos, como exemplo, um gráfico de barras que apresenta a frequência com que os valores de determinada variável aparecem em um conjunto de dados. Essa é, sem dúvida, uma forma muito sucinta de apresentar os dados ao seu leitor. De acordo com Field (2009), a estatística “(...) is an intrinsic part of the research process and it gives you enormous power to answer questions that are interesting (...)” (p.28)<sup>76</sup>.

No entanto, podemos nos perguntar: “Mas porque eu devo utilizar estatística quando faço uma pesquisa?”. Podemos usar estatística quando estamos conduzindo uma pesquisa de ordem **quantitativa**, ou seja, que trabalhe com dados que possam ser quantificados e, portanto, medidos. Pesquisas de ordem **qualitativa** envolvem a observação de pessoas (e não de números), de linguagem (como observação de conversação, de artigos de revista, de textos produzidos por alunos de determinada faixa etária, dentre outros) e de tantos outros tópicos que envolvam uma análise qualitativa, isto é, que não podem ser quantificadas.

Qual a vantagem de lançar mão da estatística em pesquisas quantitativas? Várias. Se alguém quer fazer uma pesquisa quantitativa é porque, muito provavelmente, quer observar uma determinada população, porque tem uma suposição sobre alguma coisa (uma hipótese) que envolva essa população e gostaria de testá-la. A estatística vai, portanto, auxiliar o pesquisador a observar se uma hipótese pode ou não ser sustentada. Ela fornecerá a esse pesquisador subsídios sólidos para se confirmar se aquela hipótese pode ou não ter aplicação no mundo real (*população-alvo*).

Tomemos um exemplo de pesquisa muito comum na área de Aquisição de Segunda Língua. Suponha que um determinado pesquisador queira observar se uma determinada metodologia “x” de aprendizagem é mais eficaz para se aprender inglês como língua estrangeira (ILE) do que uma metodologia “y”. Como ele pesquisaria isso? Aplicando cada metodologia com um grupo de alunos (de uma mesma turma, por exemplo) e testando-os antes e depois da aplicação

---

<sup>75</sup> Parte da estatística que descreve os dados, como por exemplo, reporta média, desvio padrão, intervalo e apresenta os dados em gráficos.

<sup>76</sup> “(...) é uma parte intrínseca do processo de pesquisa e ela dá a você um poder enorme para responder perguntas que são interessantes” (FIELD, 2009, p.28). Minha tradução.

para ver qual grupo conseguiu se sair melhor após a aplicação do teste, se o que passou pela metodologia “x” ou pela “y”. Como ele poderia observar a melhora? Através dos *scores* obtidos por cada grupo após os testes aplicados antes da inserção da metodologia e após o seu uso. Note que esse pesquisador coletou dados com apenas um grupo de alunos de ILE, e não com todos os alunos de ILE do mundo (o que seria impossível!). O pesquisador selecionou então uma **amostra** da população-alvo para verificar qual metodologia é a mais eficaz. Como relata Field (2009), “(...) we collect data from a small subset of the population (known as *sample*) and use these data to infer things about the population as a whole” (p.35)<sup>77</sup>. Usando a estatística, ele vai poder observar, dentre outras coisas, se os resultados obtidos podem apoiar a sua hipótese de que a metodologia “x” é mais eficaz que a “y” (usando um teste estatístico e observando se o resultado é estatisticamente significativo ou não).

Em suma, a estatística nos dá respaldo matemático sobre os nossos resultados, permitindo-nos, assim, fazer inferências e chegar a conclusões mais precisas e, porque não, mais cuidadosas a partir de uma amostra da população investigada. Saímos do mundo dos “achismos” e chegamos ao mundo do “pode ser que” ou é “provável que”.

#### D) ESTATÍSTICA: pressupostos<sup>78</sup>

Nesta subseção, apresentaremos alguns pressupostos teóricos sobre estatística que poderão auxiliar o leitor iniciante em Estatística a acompanhar o desenvolvimento do raciocínio sobre a tomada de decisão com relação aos testes estatísticos que foram aplicados nesta pesquisa. Não esperamos que o leitor iniciante se torne um exímio estatístico após a leitura, mas sim que ele seja capaz de acompanhar e compreender “os porquês” por trás de cada escolha (e também para que ele não faça uma leitura superficial das partes em que são reportados os resultados!). A subseção foi desenvolvida por tópicos, quase como um glossário, para que, em caso de necessidade posterior, o leitor possa rapidamente acessar a informação que deseja de maneira rápida e clara.

---

<sup>77</sup> “(...) nós coletamos dados a partir de um pequeno subgrupo da população (conhecido como **amostra**) e usa esses dados para inferir coisas sobre a população como um todo” (FIELD, 2009, p.36). Minha tradução.

<sup>78</sup> Basicamente, as informações contidas nessa subseção são provenientes do livro “Doing Statistics using SPSS”, escrito pelo professor Dr. Andy Field em 2009 e também de palestras atendidas pela pesquisadora na Escola de Inverno da UCPEL 2011, ministradas pelas professoras Dra. Andréia Schurt Rauber e Dra. Denise Cristina Kluge.

## Estatística Descritiva: pressupostos teóricos

- Existem dois tipos de natureza de pesquisa:
  - a. **Qualitativa:** pesquisa que observa as características relacionadas aos sujeitos, que não podem ser mensuradas/medidas.
  - b. **Quantitativa:** tipo de pesquisa que observa características que podem ser mensuradas/medidas.
- **Hipótese:** uma suposição feita a partir da observação de um fenômeno ou de uma teoria. Ao rodarmos os testes estatísticos, nunca se prova uma hipótese de pesquisa, mas sim se verifica a aceitação ou rejeição do que se chama de **hipótese nula**.
  - a. Hipótese nula ( $H_0$ ) é a que prevê que não haverá efeitos (por exemplo, de uma variável sobre outra) ou correlação entre variáveis.
  - b. Hipótese alternativa ( $H_1$ ) ou experimental é a que prevê que vai haver efeitos ou correlações.

Assim, se a hipótese nula for rejeitada, podemos afirmar que o efeito que estávamos buscando ao testar a nossa amostra poderá também ser encontrado na população-alvo. De acordo com Field (2009),

If our data give us confidence to reject the null hypothesis then this provides support for our experimental hypothesis. However, be aware that even if we can reject the null hypothesis, this doesn't mean we prove the experimental hypothesis – it merely supports it. (p. 27)<sup>79</sup>

- c. **Hipótese direcional:** é aquela que aponta a direção do efeito. Também é conhecida como hipótese unicaudal. Ex.: A

---

<sup>79</sup> “(...) Se nossos dados nos dão confiança para rejeitar a hipótese nula, então isso apoia a hipótese experimental. No entanto, tenha em mente que mesmo que nós possamos rejeitar a hipótese nula, isso não significa que nós provamos a hipótese experimental – isso apenas a apoia.” (FIELD, 2009, p.27). Minha tradução.

variável peso será correlacionada com a variável altura no sentido de que quanto maior o peso, maior será a altura.

d. **Hipótese não-direcional:** aquela que aponta que um efeito vai ocorrer, mas não aponta a direção desse efeito. Também é conhecida como hipótese bicaudal. Ex.: Plosivas surdas diante de vogais altas tenderão a apresentar maior valor de VOT.

- **Variável:** algo que pode ser medido/observado. É tudo aquilo que pode variar.
  - a. Variável **independente:** é a razão porque algo varia; ela é independente porque seu valor não depende de outra variável e pode ser controlada pelo pesquisador. Pode ser também chamada de variável preditora.
  - b. Variável **dependente:** é o efeito causado por outra variável (independente). Seu valor é apenas obtido/observado e não é controlada pelo pesquisador. Seu valor depende do que a variável independente vai causar. Pode ser também chamada de variável resultado.
  - c. Variável **categórica:** tipo de variável que está relacionada às características de um sujeito, que não podem ser medidas (através de uma escala, por exemplo) e que podem ser incluídas em uma categoria. Ex.: sexo, estado civil, escolaridade, ter ou não filhos, profissão, religião, cor dos olhos, nacionalidade, gravidez, etc.
    - i. Variável **dicotômica** ou **binária:** quando existem apenas duas categorias e o pertencimento a uma categoria automaticamente exclui a outra. Ex.: sexo, estar ou não grávida, resposta sim/não, etc.
    - ii. Variável **nominal:** variável categórica com mais de duas possibilidades. Esse tipo de variável só pode ser analisado através de frequências. Ex.: cor dos olhos, nacionalidade, estado civil, língua materna.

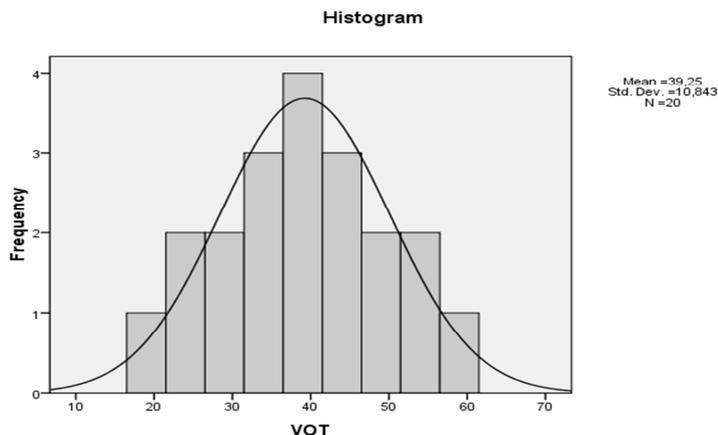
- iii. Variável **ordinal**: variável categórica que pode ser ordenada ou ranqueada. Ex.: rankings de seleções, posição final numa corrida, nível socioeconômico, grau de escolaridade, trimestres de gravidez, escala de Likert, nível de segunda língua/língua estrangeira, grau de satisfação, etc. Observação: Sempre que ranqueamos algo baseado em nossas preferências pessoais, isso deve ser considerado como um dado ordinal (ex.: grau de beleza; conforto auditivo; o quanto um determinado som é similar a uma plosiva bilabial vozeada).
  - d. Variável **contínua**: pode assumir qualquer valor na escala e atribui valores a cada item/participante.
    - i. Variável **intervalar**: variável numérica que assume intervalos iguais na escala de medida. Ex.: temperatura em Celsius, latitude, longitude, etc. Além disso, os intervalos na escala são iguais. Ex.: a diferença entre 20° e 15° C é a mesma que entre 45° e 40°, ou seja, 5° C.
    - ii. Variável de **razão**: mesmo que a variável intervalar, mas que pode assumir um ponto zero na escala, o que quer dizer que quando a variável assume o valor 0 (zero), ela não existe. Ex.: peso, altura, idade, número de filhos, quantidade de dinheiro, tempo de reação para uma resposta, etc.
    - iii. Variável **discreta**: aquela variável que só pode assumir valores inteiros na escala: Ex.: idade, número de filhos, anos de estudo, etc.
- Tipos de pesquisa:
    - a. **pesquisa correlacional** ou **transversal**: pesquisa em que são feitas apenas observações sobre o fenômeno em investigação sem a manipulação direta do pesquisador.
    - b. **pesquisa experimental**, é aquela em que há manipulação de uma ou mais variáveis pelo pesquisador.
  - Tipos de *design* de pesquisa:

- a. **Design intergrupos, intersujeitos ou independente:** diferentes grupos ou pessoas passam por experimentos diferentes; diferença entre dois grupos independentes.
  - b. **Design intragrupos, intrassujeito, de medidas-repetidas:** a mesma pessoa ou o mesmo grupo é avaliado em diferentes momentos; diferenças entre o mesmo grupo. Esse tipo de design também pode ser chamado de design de medidas repetidas.
- **Varição assistemática:** aquela que ocorre devido a fatores desconhecidos e não pela manipulação do pesquisador.
  - **Varição sistemática:** mudanças em uma variável que ocorre pela manipulação do pesquisador.

**Observação:** há mais probabilidade de se encontrar variação assistemática em *designs* intersujeitos, uma vez que os participantes apresentam diferenças entre si. Num *design* intrassujeito, esses efeitos são minimizados porque se usa os mesmos participantes nas duas coletas. Assim, um *design* intrassujeito tende a detectar com mais precisão os efeitos de uma variável sobre outra do que o *design* intersujeitos.

- **Distribuição de frequências:** gráfico que plota a frequência com que os dados ocorrem. Ele apresenta o valor de observações no eixo x (eixo horizontal) e a frequência no eixo y (eixo vertical), como pode ser visto na Figura 32.

Figura 32 – Histograma mostrando os valores da variável no eixo x (abscissas) e a distribuição de frequências no eixo y (coordenadas)



Fonte: Nossos dados

Uma das propriedades de um histograma é mostrar como se comporta a distribuição dos dados, isto é, se essa distribuição pode ser considerada normal ou não. Uma distribuição é considerada normal quando, entre outras características, ela apresenta um histograma similar ao apresentado na Figura 32. Esse gráfico apresenta uma curva em forma de sino, sinalizando que a distribuição é normal. Cada “ponta” do gráfico é chamada de cauda (*tail*) e, no gráfico normal, elas apresentam pouca concentração de dados.

Quando a distribuição não é normal, ela apresenta determinados “desvios”, chamados de **assimetria** e **curtose**.

- **Assimetria:** indica o quão assimétrica é a distribuição. Assim, um gráfico com assimetria positiva apresenta bastante concentração de dados para a esquerda do gráfico, enquanto que a assimetria negativa aponta concentração de dados na região direita do gráfico (veja Figura 22).
- **Curtose:** grau de achatamento da curva. Um histograma com dados que apresentam algum grau de curtose podem ser mais pontudos ou mais achatados que o gráfico da distribuição normal. Uma curtose positiva (chamada também de leptocúrtica) indica que há muitos dados distribuídos pelas caudas, tendendo a ser mais pontudo do

que o gráfico da distribuição normal. Já um gráfico que representa uma distribuição com curtose negativa (platicúrtica) é mais achatado, apresentando menos dados nas caudas. Para comparação entre os gráficos da curtose positiva e negativa e a distribuição normal, veja Figura 23. O gráfico da distribuição normal apresenta curtose e assimetria igual ou perto de zero.

- Medidas de tendência central: medidas que mostram o centro de uma distribuição.
  - a. **Moda:** valor mais frequente na amostra. Ex.: Na amostra ‘1 3 4 4 5 6 7 7 7 8 11’, a moda é 7. Uma amostra pode ser bimodal (quando apresenta duas modas, ou multimodal, quando apresenta três ou mais modas).
  - b. **Mediana (M):** após a ordenação da amostra, a mediana é valor que se encontra no meio da amostra. Calcula-se a mediana somando o número de itens da amostra a um e dividindo por dois  $M = \frac{N+1}{2}$ . O resultado indica a **posição** em que se encontra o número representativo da amostra. Ex.: na amostra acima, a mediana seria o número que se encontra na posição  $(11+1)/2 = 6$ , portanto a mediana seria 7. Numa amostra que tem um número par de valores, a mediana será calculada fazendo-se a média entre os dois valores mediais, podendo assumir valores não inteiros. Na amostra “1 3 4 4 5 6 7 7 7 8”, os valores mediais seriam 5 e 6. Calculando-se a média entre esses valores, teríamos uma mediana de 5,5.
  - c. **Média ( $\bar{X}$ ):** soma de todos os valores da amostra divididos pelo número de itens da amostra. Ex.: Na amostra ‘1 3 4 4 5 6 7 7 7 8 11’, a média é 8,45.

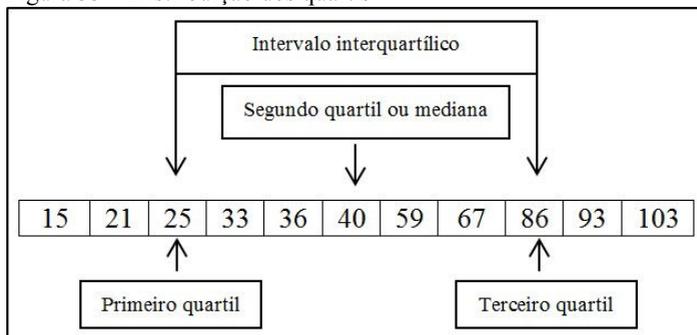
**Observação:** Na amostra ‘1 3 4 4 5 6 7 7 7 8 11 43’, o número 43 seria considerado um **dado espúrio**, por se distanciar demasiadamente dos demais. A moda e a mediana são valores que pouco são afetados por esse dado, mas a média sofre uma grande alteração se ele for incluído nos cálculos. Exemplo: Para as amostras ‘1 3 4 4 5 6 7 7 7 8 11 12’ e ‘1 3 4 4 5 6 7 7 7 8 11 43’, a moda e a mediana continuariam as mesmas: moda=7 e mediana=  $6+7/2=6,5$ . No entanto, a média da primeira

amostra seria 6,25, enquanto a média da segunda amostra seria 8,83. Sugere-se excluir esse tipo de dado, para não enviesar a análise.

- **Medidas de dispersão:** mostram o quão dispersos os dados são entre si. Podemos ver a dispersão através de:
  - a. Intervalo entre o valor mínimo e máximo (*range*): o único problema é que esse indicador é diretamente afetado por dados espúrios.
  - b. Quartis: três valores que dividem os dados em quatro partes iguais
    - i. **Segundo quartil** ou mediana: divide os dados em duas metades iguais ou em 50%.
    - ii. **Primeiro quartil:** divide a metade inferior da mediana na metade, ou seja, os primeiros 25% dos dados.
    - iii. **Terceiro quartil:** divide a metade superior dos dados na metade, ou seja, os 25% finais dos dados.
    - iv. **Intervalo interquartilico:** excluindo-se os 25% iniciais e finais dos dados, sobram os 50% “mediais” da amostra. O valor que representa o intervalo interquartilico é a diferença entre o maior e o menor valor observados.

Na Figura 33 podemos observar o posicionamento dos quartis.

Figura 33 – Distribuição dos quartis



Fonte: Adaptado de Field, 2009, p. 24.

A mediana ou segundo quartil se encontra no sexto valor, ou seja, 40. O primeiro quartil é calculado excluindo-se o valor da mediana,.

Aplicando a fórmula  $M = \frac{N+1}{2}$ , temos  $M = \frac{5+1}{2} = 3$ , que corresponde ao terceiro valor entre 15, 21, 25, 33 e 36 que é 25. Mesmo procedimento deve ser adotado para o terceiro quartil, que também é o terceiro número entre 59, 67, 86, 93 e 103, ou seja, o número 86. Por fim, calcula-se o intervalo interquartilístico diminuindo-se o terceiro do primeiro quartil. No exemplo da Figura 33, temos  $86-25=61$ .

- **Desvio padrão (DP):** mostra o quão dispersos os valores estão da média. Quanto menor o DP, mais próximos da média estão os dados, isto é, há menos variância entre eles. Do contrário, um DP grande mostra que os dados estão mais afastados da média e que há variância considerável entre eles.
- **Intervalo de confiança:** intervalo de valores entre os quais poderá estar (com, por exemplo, uma probabilidade de 95%) o valor real (da média, por exemplo) que é representativo da população-alvo. Um intervalo de confiança pode ser de 95% (de probabilidade de conter o valor real a partir da nossa amostra), de 99% ou de qualquer outro valor que possa ter alta probabilidade de conter o valor real da população-alvo. Isso quer dizer que, ao fazermos a média de 100 amostras e calcularmos um intervalo de confiança de 95% para cada uma delas, em 95 amostras poderemos encontrar o valor real da média da população-alvo entre os valores mínimos e máximos de cada intervalo de confiança.
- **Teste estatístico:** “(...) amount of variance explained by the model we’ve fitted to the data compared to the variance that can’t be explained by the model (...)” (FIELD, 2009, p. 52)<sup>80</sup>. Se o teste for capaz de explicar essa variância, ele será maior que 1. Quanto maior o valor do teste estatístico, mais variância ele consegue explicar. Para ser significativo, um teste estatístico deve apresentar uma probabilidade (de se obter esse resultado ao acaso) menor que 0,05 (i.e., menor que 5%). Se isso acontece, podemos aceitar a hipótese de investigação que previu que haveria determinado efeito na população e rejeitar a nula, que previu que não haveria tal efeito.

---

<sup>80</sup> “(...) a quantidade de variância explicada pelo modelo que aplicados aos dados, comparada com a variância que não pode ser explicada pelo mesmo modelo (...)” (FIELD, 2009, p.52). Minha tradução.

- **Testes paramétricos:** testes realizados com dados que apresentem distribuição normal, que sejam homogêneos e de natureza intervalar.
  - a. **Normalidade:** podemos atestar normalidade visualmente, através do histograma. Caso esse gráfico apresente a curvatura em formato de sino (como na Figura 32 mostrada anteriormente), os dados tendem a ser normais.

Outra forma de se verificar a normalidade é através do cálculo da assimetria e curtose. Amostras assimétricas ou que apresentem curtose não apresentam distribuição normal e apresentam valores diferentes de 0.

Existem ainda testes que são aplicados para se verificar a normalidade dos dados: **Kolmogorov-Smirnov** (doravante, KS) e **Shapiro-Wilk** (SW). Quando os resultados desses testes forem significativos ( $p < 0,05$ ), os dados não apresentam distribuição normal. Caso contrário, um  $p > 0,05$  indica que os dados são normais. Isso acontece porque a hipótese testada aqui ( $H_0$ ) é a de que a distribuição dos dados em análise não é diferente de uma distribuição normal. Portanto, se  $p < 0,05$ , rejeita-se  $H_0$ , o que indica que a distribuição em questão é diferente da distribuição normal. No entanto, se  $p > 0,05$ , aceita-se  $H_0$  e a distribuição pode ser considerada normal.

De acordo com Field (2009), para se ter confiança no julgamento da normalidade, devemos utilizar o máximo de técnicas possíveis para tal fim, pois essa informação será decisiva na escolha de um teste estatístico.

- b. **Homogeneidade:** testa se há homogeneidade entre os dados.

Para se testar a homogeneidade de dados em *designs* correlacionais usam-se gráficos e, em *designs* intergrupos, usa-se o teste de Levene. Esse teste tem como hipótese nula a homogeneidade entre os dados. Assim, como para a normalidade, o resultado significativo ( $p < 0,05$ ) rejeita a hipótese nula, indicando, portanto, que os dados não são homogêneos. Caso contrário, quando  $p > 0,05$ , aceita-se a  $H_0$ , concluindo-se que os dados apresentam homogeneidade entre si.

## Estatística Inferencial: Testes Estatísticos<sup>81</sup>

Testes de correlação<sup>82</sup> :

- **Coefficiente de correlação de Pearson:** relação linear entre duas variáveis intervalares (ou entre uma intervalar e uma categórica binária) em uma amostra que apresente distribuição normal.
- **Coefficiente de Correlação de Spearman ( $r_s$ ):** aplicado para se verificar se há correlação entre variáveis provenientes de dados não-paramétricos. Para aplicá-lo, ao menos uma das variáveis deve ser ordinal.
- **Qui-quadrado ( $\chi^2$ ):** relação entre duas variáveis categóricas através da análise das frequências.

Testes de diferença

### *Testes Paramétricos*

Estes testes são rodados quando se atesta que os dados apresentam distribuição normal e quando são homogêneos entre si.

- **Teste-*t*:**

Existem dois tipos de teste-*t*. Quando se quer investigar a diferença entre dois grupos ou variáveis, usamos o teste-*t* de amostras independentes. Quando se quer avaliar o mesmo grupo, mas avaliados em diferentes momentos, usamos o teste-*t* de amostras pareadas.

Ambos os tipos desse teste estatísticos são realizados apenas em amostras que apresentam distribuição normal e que tenham como variável dependente uma variável intervalar (MARTINS, 2011). No caso do teste-*t* de amostras independentes, os dados ainda deverão ser homogêneos e independentes.

---

<sup>81</sup> É possível encontrar uma tabela no Apêndice D com o resumo das informações pertinentes a vários outros tipo de testes. Incluímos aqui nesta subseção aqueles mais comumente usados em estudos Linguísticos e aqueles usados nesta tese. As informações que estão dispostas na tabela do Apêndice D baseiam-se em Field (2009), Larson-Hall (2010), Hays (1973), Martins (2011), Barbetta (2010) e Woods *et al.*(1986).

<sup>82</sup> Existem outros testes de correlação, tais como Kendall's tau, bisseral, ponto-bisseral e correlação parcial, mas que não serão abordados nesta tese.

Em testes-*t* pareados, o resultado do teste é negativo quando a média do primeiro experimento é menor que a do segundo, lembrando que nesse teste, as diferenças individuais são eliminadas porque o *score 1* e o *score 2* pertencem ao mesmo participante e, portanto, a diferença entre eles é atribuída somente à manipulação feita entre o momento 1 e 2. Assim, a média é calculada a partir das diferenças entre os pares de *scores* (*score 1* e 2).

Já no caso de amostras independentes, temos dois grupos distintos que participaram da mesma coleta. Assim, existem diferenças entre participantes que não podemos atribuir somente à manipulação ou experimentação. Nesse caso, devemos não mais olhar para a média das diferenças entre os pares de *scores* dos participantes, mas sim entre as médias dos dois grupos.

Os correlatos não-paramétricos desses dois tipos de teste-*t* são: **Mann-Whitney**, para amostras independentes e **Wilcoxon signed-rank** para amostras pareadas, que serão detalhados abaixo.

- **ANOVA:**

Esse teste é utilizado quando se quer investigar diferenças entre mais de três grupos, variáveis ou momentos.

O teste ANOVA unifatorial analisa diferenças entre três grupos ou mais, enquanto que o ANOVA de medidas repetidas analisa diferenças entre o mesmo grupo com coletas em três momentos diferentes ou mais.

Alguns dos pressupostos para se rodar quaisquer dos ANOVAs são:

- ele parte do princípio de que as médias dos grupos são iguais (hipótese nula).

- a variável dependente deve ser intervalar e a distribuição deve ser normal. É, portanto, um teste paramétrico (MARTINS, 2011).

Para se descobrir onde estão as diferenças entre os grupos é preciso rodar análises *post hoc*. Essas análises investigam as diferenças em pares, ou seja, comparam os grupos, variáveis ou momentos entre si, com análises dois a dois. Alguns exemplos de testes *post hoc* são **Scheffé**, **Gabriel**, **Turkey**, **Bonferroni**, dentre outros.

Contudo, após rodar um teste *post hoc* é preciso utilizar a **correção de Bonferroni**, que é um método para se corrigir o nível de significância *p*, dividindo o valor especificado pelo número de

comparações feitas. Por exemplo, se o valor de  $p$  for especificado em 0,05 e o número de comparações for 3, divide-se 0,05 por 3, o que nos dá um valor de  $p = 0,17$ . Assim, as diferenças entre grupos só serão significativas caso o novo nível de  $p$  seja alcançado.

Por fim, quando os dados não apresentam normalidade, pode-se aplicar a eles o teste não-paramétrico equivalente ao ANOVA, que é denominado **Kruskal-Wallis**.

**Observação:** Existem outros tipos de testes paramétricos como ANOVA fatorial, ANCOVA e MANOVA que, na verdade, são variações do ANOVA Unifatorial e, ainda, o teste ANOVA para medidas repetidas. Não entraremos em detalhes sobre esses testes neste trabalho. Para mais informações, sobre eles, consulte Field (2009).

### *Testes não-paramétricos*

São testes aplicados aos dados que não apresentam distribuição normal. Para rodar esses testes, os dados são primeiramente ranqueados e, então, os cálculos são feitos em cima desses *rankings* ao invés dos dados em si.

- **Mann-Whitney:** esse teste é o correlato não-paramétrico do teste- $t$  de amostras independentes, pois se constitui de uma análise entre duas variáveis, condições ou grupos, mas aplicados a dados que não apresentam distribuição normal.
- **Wilcoxon signed-rank (amostras pareadas):** teste utilizado quando se quer observar a diferença entre dois conjuntos de dados oriundos de um mesmo grupo de participantes testados, por exemplo, em dois momentos diferentes. Ele é o equivalente não-paramétrico do teste- $t$  de amostras pareadas/dependentes.
- **Kruskal-Wallis:** avalia a diferença entre mais de três grupos ou variáveis que sejam independentes. Assim como o Anova unifatorial, esse teste mostra que há diferença entre os grupos, mas não aponta onde estão essas diferenças. Para poder saber onde elas estão, precisamos rodar testes *post hoc*, através do teste Mann-Whitney, comparando pares de grupos, usando novamente a correção Bonferroni (dividir o valor de  $p$  pelo número de grupos avaliados) para observar se as diferenças são significativas ou não.

**Observação:** Existem outros testes não-paramétricos, como Wilcoxon *rank-sum*. Porém esses testes não serão detalhados neste trabalho. Para mais informações, consulte Field (2009).

- **Friedman:** da mesma forma que o teste de Kruskal-Wallis, o teste de Friedman deve ser aplicado a dados não-normais. Eles diferem no tipo de *design*, pois o Kruskal-Wallis avalia amostras independentes (intergrupos), enquanto que Friedman avalia amostras pareadas (intragrupos). Geralmente conjuntos de dados são coletados em três ou mais momentos com os mesmos participantes, por isso a necessidade de se usar esse teste.

Em suma, a estatística dispõe de variados testes que vão englobar diferentes formas de *design* de pesquisa e coleta de dados. Assim, precisamos seguir alguns passos para poder escolher com convicção o tipo de teste a ser aplicado em nossos dados, a saber:

- 1) Passo 1 – identificar o tipo de *design* de coleta: se independente/intergrupos ou se dependente/pareado/intragrupo;
- 2) Passo 2 – identificar as variáveis: se nominal, ordinal ou intervalar; se dependente ou independente;
- 3) Passo 3 – identificar as hipóteses nulas para cada hipótese do experimento, lembrando sempre que os testes avaliam  $H_0$  e não  $H_1$ . Dependendo do resultado do teste (se significativo ou não-significativo),  $H_0$  poderá ser rejeitada ou confirmada.
- 4) Passo 4 – avaliar se os dados são normais e homogêneos, tanto visualmente através de gráficos como o histograma, como através de testes de normalidade e homogeneidade;
- 5) Passo 5 – baseando-se nas informações avaliadas anteriormente, observar qual o tipo de teste mais se adequa ao que pretendemos investigar (o Quadro 6 apresenta um resumo dos testes de correlação e o Quadro 7 dos testes de diferença);
- 6) Passo 6 – avaliar os resultados do teste e entender o que cada um dos números gerados pelo programa estatístico quer nos dizer: se rejeitamos ou aceitamos  $H_0$  (através, principalmente do nível de significância  $p$  ( $\alpha$  – *level*)).

Quadro 6 – Resumo dos testes de correlação e suas respectivas variáveis

Teste	Tipo de Variáveis	Observação
Correlação de Pearson ( $r$ )	duas variáveis (ao menos uma intervalar)	dados paramétricos
Correlação de Spearman ( $r_s$ )	duas variáveis (ao menos uma ordinal)	dados não-paramétricos
Kendall's tau ( $\tau$ )	duas variáveis (ao menos uma ordinal)	dados não-paramétricos (amostras pequenas)
Chi quadrado ( $\chi^2$ )	duas variáveis categóricas (nominais)	

Quadro 7 – Resumo dos testes de diferenças e suas respectivas variáveis

Teste	Variável Dependente	Design	Pressupostos
Teste-t amostras independentes ( $t$ )	intervalar	intergrupos (até 2 grupos)	dados paramétricos
Teste-t de amostras pareadas ( $t$ )	intervalar	intragrupo (até 2 momentos)	dados paramétricos
One-way Anova ( $F$ )	intervalar	intergrupos (três ou mais)	dados paramétricos
Anova de medidas repetidas ( $F$ )	intervalar	intragrupos (3 ou mais momentos)	dados paramétricos
Mann-Whitney ( $U$ )	ordinal/intervalar (ranqueamento dos <i>scores</i> )	intergrupos (até dois grupos)	dados não-paramétricos
Wilcoxon signed-rank ( $T$ )	ordinal/intervalar (ranqueamento dos <i>scores</i> )	intragrupo (até 2 momentos)	dados não-paramétricos
Kruskal-Wallis ( $H$ )	ordinal/intervalar (ranqueamento dos <i>scores</i> )	intergrupos (3 ou mais)	dados não-paramétricos
Friedman ( $F_r$ )	ordinal/intervalar (ranqueamento dos <i>scores</i> )	intragrupo (3 ou mais momentos)	dados não-paramétricos

Gostaríamos de finalizar essa breve introdução à Estatística Descritiva e Inferencial reforçando que existe uma ampla gama de testes estatísticos, utilizados em variados casos, mas que por motivos de espaço, foram deixados de lado nesta introdução. Aqui relatamos os testes mais utilizados na literatura da área, bem como aqueles utilizados para análise dos dados desta pesquisa. Como citamos anteriormente, existem livros dedicados exclusivamente à Estatística aplicada às Ciências Sociais ou até mesmo à Linguística, que servirão como base para que o leitor possa se aprofundar no tema, tais como: Hays (1973), Woods *et al.* (1986), Field (2009), Barbetta (2010), Larson-Hall (2010), Martins (2011), dentre outros.

## E) ESTATÍSTICA UTILIZADA NA TESE

“Data are often unfriendly and don’t always turn up in nice normally distributed packages!” (FIELD, 2009, p. 540)<sup>83</sup>

Relataremos a seguir o processo de escolha dos testes estatísticos aplicados a cada uma das perguntas de pesquisa presentes nesta pesquisa. Para cada uma delas, relataremos quais os tipos de variáveis utilizadas, se foram utilizados paramétricos ou não (a partir dos resultados de normalidade e homogeneidade) e qual o teste estatístico utilizado.

Como já relatado nos capítulos anteriores, nossa pesquisa pretende investigar a influência dos vários parâmetros relacionados à produção das plosivas vozeadas e não-vozeadas do PB. Coletamos dados em um único momento com os participantes para avaliar como cada parâmetro pode auxiliar na distinção entre uma ou outra plosiva. Assim, podemos afirmar que nossa pesquisa é de natureza quantitativa e se caracteriza por apresentar um *design* intersujeito. Como queremos avaliar a diferença entre uma ou outra plosiva a partir de um determinado parâmetro, iremos utilizar testes que atestem a diferença entre os grupos. Em suma, temos uma pesquisa quantitativa, intersujeito, que quer investigar a diferença entre os grupos. Abaixo relacionamos cada uma das perguntas e hipóteses de pesquisa, com as respectivas informações sobre as escolhas estatísticas.

---

<sup>83</sup> “Os dados são quase sempre não-amigáveis e nem sempre aparecem em pacotes com distribuição normal!” (FIELD, 2009, p.540). Minha tradução.

**a) Pergunta 1:**

**Pergunta 1:** Há diferença entre a duração das consoantes plosivas vozeadas e não-vozeadas?

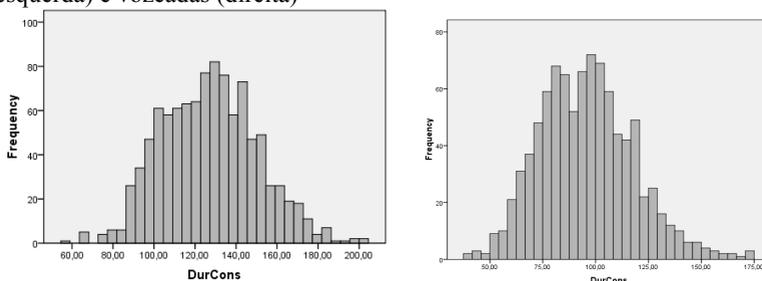
**Hipótese 1:** Haverá diferença entre as consoantes plosivas vozeadas e não-vozeadas, sendo que as últimas apresentarão valores de duração de consoante maiores do que as primeiras. Essa tendência se mantém na comparação entre os pares de plosivas homorgânicas.

Quadro 8 – Informações estatísticas referentes à P1 para a avaliação entre plosivas vozeadas e não-vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da consoante	Plosiva voz e não-voz	KS: $p < 0,05$ SW: $p < 0,05$	Mann-Whitney
Intervalar	Nominal, binária	Levene: $p < 0,05$	
		Não-paramétrico	

A Pergunta 1 almeja investigar a diferença entre as plosivas vozeadas ou não, com relação ao parâmetro duração da consoante. Esse parâmetro foi considerado como a variável dependente, pois é aquela que não foi controlada pela pesquisadora, ou seja, é a variável que foi observada. Ela é de natureza intervalar, por permitir que ela seja medida (neste caso, em milissegundos). Já a variável independente é o tipo de plosiva, se vozeada ou não-vozeada, que é uma variável nominal (não permite medição) e binária (a plosiva apresenta vozeamento ou não e, portanto, pertence a uma única categoria). O testes de normalidade e homogeneidade foram não-significativos, mostrando que os dados não apresenta distribuição normal e portanto não paramétricos. Isso é confirmado pela avaliação dos histogramas de cada grupo, dispostos na Figura 34.

Figura 34 – Histograma da duração da consoante para as plosivas não-vozeadas (esquerda) e vozeadas (direita)



Fonte: Nossos dados

Os dois histogramas possuem certo grau de assimetria (ambos positiva), por ter mais concentração de dados à esquerda e maior extensão da cauda para a direita. Ambos apresentam também múltiplos picos e poucos dados nas caudas, o que pode indicar curtose leptocúrtica ou positiva.

Baseados nessas informações, optamos pelo teste Mann-Whitney, por ser aquele que avalia duas variáveis e é não-paramétrico.

Para a avaliação entre plosivas homorgânicas, o resultado dos testes de normalidade e homogeneidade mostraram-se não-normais para a duração absoluta de [p] x [b] e de [k] x [g]. Utilizamos o teste não-paramétrico Mann-Whitney para essas comparações (comparações entre dois grupos). Para a comparação da duração absoluta de [t] e [d], os resultados apontaram para dados normais e, por essa razão, utilizamos o Teste-t de amostras independentes por ser um teste paramétrico entre dois grupos.

As escolhas estatísticas relacionadas a outras perguntas irão se resumir a um quadro tal qual o apresentado para a Pergunta 1. A linha de raciocínio desenvolvida para cada uma das perguntas foi a mesma que acabamos de relatar.

#### b) Pergunta 2.1 e 2.2:

**Pergunta 2.1:** Há influência do ponto de articulação na duração da consoante entre as plosivas não-vozeadas?

**Hipótese 2.1:** Não haverá diferença entre os três pontos de articulação com relação à duração das plosivas não-vozeadas.

Quadro 9 – Informações estatísticas referentes à P2.1, para a duração absoluta das não-vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da consoante	[p] x [t] x [k]	KS: $p > 0,05$ SW: $p > 0,05$	Anova Unifatorial com <i>post hoc</i> de Gabriel
Intervalar	Nominal, três grupos	Levene: $p > 0,05$ Paramétrico	

**Pergunta 2.2:** Há influência do ponto de articulação na duração da consoante entre as plosivas vozeadas?

**Hipótese 2.2:** O ponto de articulação influenciará a duração das consoantes vozeadas, pois, na medida em que o ponto vai se tornando mais posterior, diminui-se a duração da consoante.

Quadro 10 - Informações estatísticas referentes à P2.2, para a duração absoluta das vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da consoante	[b] x [d] x [g]	KS: $p < 0,05$ SW: $p < 0,05$	Kruskal-Wallis seguido de Mann-Whitney
Intervalar	Nominal, três grupos	Levene: $p < 0,05$ Não-paramétrico	

Para a avaliação da duração relativa de não-vozeadas e vozeadas, utilizamos o mesmo procedimento apresentado no Quadro 10.

### c) Pergunta 3:

**Pergunta 3:** Há diferenças entre a duração da oclusão de plosivas vozeadas e não-vozeadas?

**Hipótese 3:** A duração da oclusão em plosivas não-vozeadas tende a ser maior do que o mesmo momento em plosivas vozeadas, mesmo na comparação entre plosivas homorgânicas.

Para a avaliação das diferenças na duração absoluta e relativa da oclusão entre plosivas vozeadas e não-vozeadas e entre os pares de plosivas homorgânicas, utilizamos o procedimento descrito no Quadro 11.

Quadro 11 – Informações estatísticas referentes à P3

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da oclusão	Plosivas voz e n-voz	KS: $p < 0,05$ SW: $p < 0,05$	Mann-Whitney
Intervalar	Nominal, binária	Levene: $p < 0,05$ Não-paramétrico	

**d) Perguntas 4.1 e 4.2:**

**Pergunta 4.1:** Há influência do ponto de articulação na duração da oclusão entre as plosivas não-vozeadas?

**Hipótese 4.1:** Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.

Quadro 12 – Informações estatísticas referentes à P4.1, para as plosivas não-vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da oclusão	[p] x [t] x [k]	KS: $p < 0,05$ SW: $p < 0,05$	Kruskal-Wallis seguido de Mann-Whitney
Intervalar	Nominal, três grupos	Levene: $p < 0,05$ Não-paramétrico	

**Pergunta 4.2:** Há influência do ponto de articulação na duração da oclusão entre as plosivas vozeadas?

**Hipótese 4.2:** Quanto mais posterior o ponto de articulação, menor tenderá a ser a duração da oclusão.

Quadro 13 – Informações estatísticas referentes à P4.2, para as plosivas vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da oclusão	[b] x [d] x [g]	KS: $p > 0,05$ SW: $p < 0,05$	Kruskal-Wallis
Intervalar	Nominal, três grupos	Levene: $p > 0,05$ Não-paramétrico	

**e) Pergunta 5:**

**Pergunta 5:** Os dados de falantes adultos também apresentarão especificidades acústicas, como as apresentadas em dados de estudos com crianças falantes nativas do PB?

**Hipótese 5:** Os dados de falantes adultos apresentarão especificidades acústicas, sendo o tipo mais comum a presença de *edge vibrations* nas plosivas não-vozeadas e interrupções de vozeamento nas vozeadas.

Para essa pergunta, realizamos uma inspeção espectrográfica, observando e anotando quais os dados que apresentavam essas e. Para a análise, utilizamos apenas Estatística Descritiva. Não foi necessário o uso de testes estatísticos, uma vez que não medimos nenhum dos dados, apenas os observamos. Essa foi, portanto, uma análise qualitativa.

**f) Pergunta 6:**

**Pergunta 6:** Os valores das características espectrais do *burst* diferem entre os pontos de articulação?

**Hipótese 6:** Os valores de características espectrais de *burst* se diferenciarão com relação ao ponto de articulação.

Nessa análise, rodamos testes estatísticos para cada uma das medidas de *burst* (pico espectral, centroide, variância, assimetria e curtose) para os diferentes pontos de articulação de cada categoria de plosiva. O Quadro 14 traz o exemplo da análise de pico espectral aplicado entre as plosivas não-vozeadas.

Quadro 14 – Informações estatísticas referentes à P6 para a avaliação de pico espectral

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Pico espectral	[p] x [t] x [k]	KS: $p < 0,05$ SW: $p < 0,05$	Kruskal-Wallis seguido de Mann-Whitney
Intervalar	Nominal, três grupos	Levene: $p < 0,05$ Não-paramétrico	

O mesmo procedimento foi aplicado para a comparação de pico espectral das vozeadas e também para a comparação de variância, assimetria e curtose. A análise de centroide mostrou dados normais para

as plosivas não-vozeadas e, por isso, utilizamos o teste Anova Unifatorial seguido de Testes *Post Hoc* de Gabriel. A avaliação de centroide para as plosivas vozeadas seguiu o proposto no Quadro 14.

### g) Pergunta 7:

**Pergunta 7:** As medidas de *burst* são diferentes para as diferentes vogais que seguem as plosivas?

**Hipótese 7:** Acredita-se que existam diferenças no *burst* das plosivas seguidas por diferentes vogais para o PB, sendo que plosivas seguidas por vogais posteriores podem apresentar maiores valores de frequência do que aquelas seguidas por vogais anteriores.

Nesta pergunta, avaliamos se cada uma das medidas espectrais se mostrou diferente de acordo com a vogal que seguia cada um dos tipos de plosiva (bilabial, alveolar ou velar). O Quadro 15 mostra a análise de pico espectral para as plosivas alveolares, seguidas ora por [ɐ] ora por [ʊ].

Quadro 15 – Informações estatísticas referentes à P7, para a comparação de pico espectral das alveolares entre os contextos vocálicos adjacentes [ɐ] e [ʊ]

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Pico espectral	[ɐ] x [ʊ]	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$ Paramétrico	

Em todas as comparações entre plosivas alveolares foram utilizados os procedimentos descritos no Quadro 15. Nas demais, utilizamos testes não-paramétricos de Mann-Whitney.

### h) Pergunta 8:

**Pergunta 8:** Há diferenças entre o VOT de plosivas vozeadas e não-vozeadas?

**Hipótese 8:** O VOT das plosivas não-vozeadas apresentará valores positivos e as plosivas não-vozeadas poderão ser incluídas na categoria de não-aspiradas ou levemente aspiradas, além da categoria de

retardo curto. O VOT de plosivas vozeadas do PB apresentará valores negativos, indicando pré-sonorização.

Tanto a comparação entre plosivas vozeadas e não-vozeadas, quanto entre os pares homorgânicos utilizaram-se os procedimentos descritos nos Quadro 16.

Quadro 16 – Informações estatísticas referentes à P8

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
VOT	Plosiva voz x n-voz	KS: $p < 0,05$ SW: $p < 0,05$	Mann-Whitney
Intervalar	Nominal, binária	Levene: $p < 0,05$	
		Não-paramétrico	

### i) Perguntas 9.1 e 9.2:

**Pergunta 9.1:** Há diferenças entre os valores de VOT entre as plosivas não-vozeadas?

**Hipótese 9.1:** O valor da plosiva velar não-vozeada se diferenciará das demais plosivas, pois apresentará maiores valores de VOT, enquanto os valores de VOT poderão ser similares para as plosivas bilabiais e alveolares. As plosivas velares serão aquelas com mais chance de apresentar valores de VOT dentro da categoria de leve aspiração.

Para a avaliação do VOT dentre os três grupos de plosivas não-vozeadas e vozeadas, utilizou-se o teste Kruskal-Wallis por não serem dados paramétricos, conforme o Quadro 17.

Quadro 17 – Informações estatísticas referentes à P9.1 para a análise do VOT entre as plosivas não-vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
VOT	[p] x [t] x [k]	KS: $p < 0,05$ SW: $p < 0,05$	Kruskal-Wallis
Intervalar	Nominal, três grupos	Levene: $p < 0,05$	
		Não-paramétrico	

Utilizamos o mesmo procedimento que a pergunta anterior para esta análise, uma vez que os testes de normalidade e homogeneidade

indicaram dados não-paramétricos. Os testes utilizados foram Kruskal-Wallis, seguido por Mann-Whitney para cada uma das comparações entre as plosivas vozeadas.

#### j) Pergunta 10:

**Pergunta 10:** Há diferenças entre o VOT das plosivas que apresentaram aspiração e o VOT das que não apresentaram?

**Hipótese 10:** Existirá diferença significativa entre o VOT das plosivas aspiradas e não-aspiradas.

Aqui foram analisados apenas os dados referentes às plosivas não-vozeadas, que são aquelas que podem apresentar (leve) aspiração no PB. Fizemos, primeiramente, uma categorização de quais dados eram não-aspirados ( $VOT < 35$  ms) e que eram levemente aspirados ( $35 \text{ ms} < VOT < 55$  ms). O Quadro 18 exemplifica uma das análises.

Quadro 18 – Informações estatísticas referentes à P10, para a plosiva bilabial

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
VOT	[p] x [p <sup>h</sup> ]	KS: $p < 0,05$ SW: $p < 0,05$	Mann-Whitney
Intervalar	Nominal, binária	Levene: $p < 0,05$	
		Não-paramétrico	

#### k) Pergunta 11:

**Pergunta 11:** Há diferenças em F1 entre os pares de plosivas homorgânicos?

**Hipótese 11:** A região de F1 para plosivas vozeadas apresentará valores mais baixos do que para as não-vozeadas. Espera-se encontrar transições mais acentuadas para as plosivas vozeadas.

Para a análise desta pergunta, foram rodados testes entre os valores de F1 para as plosivas homorgânicas, conforme o Quadro 19.

Quadro 19 – Informações estatísticas referentes à P11, para as plosivas bilabiais

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
F1	[p] x [b]	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$	
		Paramétrico	

### l) Perguntas 12.1 e 12.2:

**Pergunta 12.1:** Há diferenças em F2 entre as plosivas não-vozeadas?

**Hipótese 12.1:** Se encontrará diferenças significativas entre os valores de transição de F2 para os diferentes pontos de articulação. Ademais, haverá influência da vogal que segue a plosiva na transição de F2, sobretudo no caso das plosivas velares.

Quadro 20 – Informações estatísticas referentes à P12.1, para não-vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
F2	[p] x [t] x [k]	KS: $p > 0,05$ SW: $p > 0,05$	Kruskal-Wallis seguido de Mann-Whitney
Intervalar	Nominal, binária	Levene: $p < 0,05$	
		Não-paramétrico	

Na análise de transição de F2 para as vozeadas (Pergunta 12.2) foram utilizados os mesmos procedimentos do Quadro 20. O resultado do teste Kruskal-Wallis foi não-significativo.

### m) Pergunta 13:

**Pergunta 13:** Há diferença na duração das vogais que antecedem as plosivas não-vozeadas e plosivas vozeadas?

**Hipótese 13:** A duração das vogais que antecedem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.

Quadro 21 – Informações estatísticas referentes à P13, para a duração da vogal [a] em frente a não-vozeadas x vozeadas

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
----	----	-----------------------------	-------------------

Duração da vogal	Plosiva voz e não-voz	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$ Paramétrico	

Mesmo procedimento foi aplicado na comparação entre as plosivas homorgânicas.

**n) Pergunta 14:**

**Pergunta 14:** Há diferença na duração das vogais que seguem as plosivas não-vozeadas e plosivas vozeadas?

**Hipótese 14:** A duração das vogais que seguem as plosivas vozeadas será maior do que nas plosivas não-vozeadas.

Para essa análise, aplicamos o mesmo procedimento da Pergunta 13, apresentada no Quadro 21.

**o) Pergunta 15:**

**Pergunta 15:** Existe diferença na duração da consoante e de seu VOT dependendo da altura da vogal que a segue?

**Hipótese 15:** A duração da consoante e de seu VOT serão maiores quando a vogal adjacente é uma vogal alta (em detrimento da vogal baixa).

Realizamos testes estatísticos para essa pergunta da seguinte forma: avaliamos a duração da consoante e do VOT para cada uma das plosivas do PB, levando em consideração a qualidade da vogal que a seguia. O Quadro 22 apresenta as informações estatísticas referentes a essa análise, para a duração da consoante [p] em frente a [e] e [u]. Mesmo procedimento foi aplicado para as demais plosivas não-vozeadas e também para todas as vozeadas.

Quadro 22 – Informações estatísticas referentes a P15, para a duração da consoante [p] em frente as vogais [e] e [u]

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da consoante	_[e] x _[u]	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$	
		Paramétrico	

**p) Perguntas 16.1 e 16.2:**

**Pergunta 16.1:** O contexto posição na palavra exerce influência sobre a duração total da consoante e VOT das plosivas não-vozeadas do PB?

**Hipótese 16.1:** A posição na palavra poderá influenciar a duração total da consoante e VOT de plosivas não-vozeadas.

Para algumas avaliações, os dados mostraram-se normais e, então, aplicamos o Teste-t de amostras independentes, como mostra o Quadro 23.

Quadro 23 – Informações estatísticas referentes a P16.1, para a duração da consoante [p] no contexto início e meio de palavra

VD	VI	Normalidade e Homogeneidade	Teste Estatístico
Duração da consoante	inicial x medial	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$	
		Paramétrico	

Para as comparações em que os dados foram não-normais, testes Mann-Whitney foram aplicados, por ser um teste não-paramétrico.

O mesmo procedimento foi aplicado à Pergunta 16.2.

**q) Perguntas 17.1 e 17.2:**

**Pergunta 17.1:** O contexto posição na palavra exerce influência sobre a duração total e o VOT plosivas não-vozeadas do PB?

**Hipótese 17.1:** Se o PB seguir a mesma tendência para outras línguas (KLATT, 1976), as consoantes em posição átona terão maior

duração do que em posição tônica. No que tange o VOT, o contexto átono propiciará maiores valores de VOT no caso de bilabiais e alveolares e oônico propiciará a maior duração do VOT de consoantes velares.

Quadro 24 – Informações estatísticas referentes a P17.1, para a duração da consoante [p] no contextoônico e átono

<b>VD</b>	<b>VI</b>	<b>Normalidade e Homogeneidade</b>	<b>Teste Estatístico</b>
Duração da consoante	tônico x átono	KS: $p > 0,05$ SW: $p > 0,05$	Teste-t de amostras independentes
Intervalar	Nominal, binária	Levene: $p > 0,05$ Paramétrico	

O mesmo procedimento foi utilizado para a Pergunta 17.2.

## ANEXO I – Certificado de aprovação no Comitê de Ética

29/05/12

Certificado



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
 Pró-Reitoria de Pesquisa e Extensão  
 Comitê de Ética em Pesquisa com Seres Humanos

**CERTIFICADO** Nº 2057

O Comitê de Ética em Pesquisa com Seres Humanos (CEPSH) da Pró-Reitoria de Pesquisa e Extensão da Universidade Federal de Santa Catarina, instituído pela PORTARIA N.º 0584/GR/99 de 04 de novembro de 1999, com base nas normas para a constituição e funcionamento do CEPSH, considerando o conteúdo no Regimento Interno do CEPSH, **CERTIFICA** que os procedimentos que envolvem seres humanos no projeto de pesquisa abaixo especificado estão de acordo com os princípios éticos estabelecidos pela Comissão Nacional de Ética em Pesquisa – CONEP.

**APROVADO**

**PROCESSO:** 2057 **FR:** 434924

**TÍTULO:** O detalhe fonético: análise acústica exploratória de segmentos de fala

**AUTOR:** Isabel Christine Seara, Isabel Christine Seara

FLORIANÓPOLIS, 28 de Maio de 2012.

\_\_\_\_\_  
 Coordenador do CEPSH/UFSC