



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE HUMANIDADES**  
**DEPARTAMENTO DE LETRAS VERNÁCULAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA**

**PEDRO HENRIQUE SOUSA DA SILVA**

**ANÁLISE ACÚSTICA DE *CLUSTERS* COM *TAP* NA FALA DE CRIANÇAS  
COM E SEM DESVIO FONOLÓGICO**

**FORTALEZA**

**2018**

PEDRO HENRIQUE SOUSA DA SILVA

ANÁLISE ACÚSTICA DE *CLUSTERS* COM *TAP* NA FALA DE CRIANÇAS  
COM E SEM DESVIO FONOLÓGICO

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutor em Linguística.

Orientadora: Prof<sup>ª</sup>. Dra. Maria Elias Soares.

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

S582a

Silva, Pedro Henrique Sousa da.

Análise acústica de clusters com tap na fala de crianças com e sem desvio fonológico /  
Pedro Henrique Sousa da Silva. – 2018.

104 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de  
Pós-Graduação em Linguística, Fortaleza, 2018.

Orientação: Profa. Dra. Maria Elias Soares.

1. Desvio fonológico. 2. Fonética acústica. 3. Tap. I. Título.

CDD 410

---

PEDRO HENRIQUE SOUSA DA SILVA

ANÁLISE ACÚSTICA DE *CLUSTERS* COM *TAP* NA FALA DE CRIANÇAS  
COM E SEM DESVIO FONOLÓGICO

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutor em Linguística.

Aprovada em: 09 / 02 / 2018.

**BANCA EXAMINADORA**

---

Prof<sup>a</sup>. Dra. Maria Elias Soares (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Plínio Almeida Barbosa  
Universidade de Campinas (UNICAMP)

---

Prof. Dr. Wilson Júnior Araújo Carvalho  
Universidade Estadual do Ceará (UECE)

---

Prof<sup>a</sup>. Dra. Vlândia Maria Cabral Borges  
Universidade Federal do Ceará (UFC)

---

Prof<sup>a</sup>. Dra. Ana Célia Clementino Moura  
Universidade Federal do Ceará (UFC)

Às crianças com dificuldades especiais na fala.

À minha esposa Renata Rodrigues, às minhas filhas Emilie Nelise e Sofia Germana, e aos meus sobrinhos Ana Beatriz, Kaio Átila, Ícaro Levi e Lorena Évelyn.

À minha avó Germana Coelho (*in memoriam*)

Ao meu tio José Nilton (*in memoriam*).

Ao meu amigo Paulo Mosânio (*in memoriam*).

## AGRADECIMENTOS

Agradeço, inicialmente, aos meus pais, Antônio Moreira e Ana Maria, por todo o empenho e esforço que fizeram para me ajudar tanto no empreendimento deste trabalho quanto na minha vida como um todo.

À minha esposa Renata Rodrigues, por ter lutado ao meu lado, equilibrando delicadeza e força, frente aos desafios internos e externos desta tese.

Às minhas filhas Emilie Nelise e Sofia Germana, eternas fontes de inspiração para a construção deste e de outros trabalhos.

À CAPES, pelo apoio financeiro com a manutenção da bolsa de doutorado.

À minha orientadora Prof<sup>ª</sup>. Dra. Maria Elias Soares, pelas sugestões relevantes para o esmero desta tese.

Ao Prof. Dr. Plínio Barbosa, da Universidade de Campinas, pelas sugestões e contribuições relevantes para o aperfeiçoamento deste trabalho. Diante disso, devo ressaltar que, caso algumas falhas e imperfeições persistam nesta tese, elas são de minha inteira responsabilidade. Também agradeço à Prof<sup>ª</sup>. Dra. Sandra Madureira, da Pontifícia Universidade Católica de São Paulo e, novamente, ao Prof. Dr. Plínio Barbosa, pela publicação do Manual de Fonética Acústica Experimental, o qual ajudou bastante no desenvolvimento desta tese.

Aos demais professores membros da banca de defesa desta tese, Prof. Dr. Wilson Carvalho, da Universidade Estadual do Ceará, Prof<sup>ª</sup>. Dra. Vlândia Borges e Prof<sup>ª</sup>. Ana Célia, ambas da Universidade Federal do Ceará, pelas sugestões relevantes apontadas na banca.

À Prof<sup>ª</sup>. Dra. Vlândia Borges, mais uma vez, e à Prof<sup>ª</sup>. Dra. Márcia Teixeira, pelas sugestões importantes na banca de qualificação, as quais ajudaram a afunilar o objeto desta tese, num momento em que ela era apenas um projeto.

Ao Prof. Dr. Ronaldo Lima, pelas sugestões relevantes numa fase inicial do desenvolvimento desta tese e também por suas contribuições, já em outra fase, nos seminários de pesquisa.

À Prof<sup>ª</sup>. Dra. Adelaide Silva, da Universidade Federal do Paraná, pelas importantes contribuições para o desenvolvimento de algumas frações da metodologia e das análises.

À Prof<sup>ª</sup>. Dra. Vanessa Giacchini - da Universidade Federal do Rio Grande do Norte, pelas sugestões relevantes sobre questões ligadas ao desvio fonológico.

Às bibliotecárias da Biblioteca do Centro de Humanidade da Universidade Federal do Ceará, entre as quais Eliene Moura e Josineide Góis e a servidora Márcia Bezerra, por terem me ajudado com questões de ordem técnica ligadas à ABNT, além de outras.

Às fonoaudiólogas do Nami, da Universidade de Fortaleza, Anna Lídia, Vanessa, Mércia, Denise e Lília, as quais nos ajudaram a angariar determinados participantes da pesquisa, sobretudo, as crianças com diagnóstico de desvio fonológico. Sem a generosidade dessas profissionais teria sido muito mais difícil.

À diretora Roberta Pereira e às professoras da Escola Mattos Dourado, por terem nos ajudado também a angariar alguns participantes da pesquisa, oriundos da reportada escola.

Às crianças que participaram da pesquisa e a seus pais/responsáveis - sem os quais este trabalho não faria sentido. Não poderia me furtar de dizer que eles me ajudaram de vários modos, a superar dificuldades técnicas e práticas no momento da coleta de dados.

Aos servidores do Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará: Eduardo, Antônia e Valdirene, por terem solucionado determinados problemas burocráticos cujas soluções facilitaram a entrega deste trabalho.

Aos meus irmãos, primos, tios, tias, parentes em geral e amigos, pelo apoio e por torcerem para que tudo desse certo, entre os quais, cumpre destacar, aqui, Marcos Antônio, meu irmão, João Márcio, meu primo e a colega Débora Giffoni.

Esta tese pertence, por afeição e gratidão, um pouco a cada uma dessas pessoas.

“Quem não quer ver o que há de elevado num homem, olha tanto mais agudamente para o que nele é baixo e superficial – e com isso se revela.” (NIETZSCHE, 2005, p. 171).



## RESUMO

Esta tese tem como objetivo central explicar/descrever padrões acústicos do *tap* em onset complexo na fala de crianças com e sem desvio fonológico à luz da Fonética Acústica. Também visa investigar determinados detalhes fonéticos presentes em *clusters* com *tap*, nos quais esta consoante é precedida por consoantes oclusivas, surdas e sonoras, e pela fricativa labiodental surda, de acordo com a sintaxe fônica do português brasileiro. A pergunta norteadora desta pesquisa é esta: em que medida as produções acústicas do *tap* em onset complexo se distinguem na fala de crianças com e sem desvio fonológico? Os procedimentos metodológicos consistiram na aplicação de teste de nomeação de figuras, visando eliciar palavras-alvo dissílabas paroxítonas, com *tap* em onset complexo seguido por vogal tônica [a]. Os testes se aplicaram a dois grupos distintos de participante: um constituído de crianças com desvio fonológico (N=4) e outro, de crianças sem o desvio (N=4). Para a gravação dos dados, utilizaram-se um gravador *Zoom* modelo H4n (*Handy Recorder*), com taxa de amostragem de 44100 Hz, e um microfone de lapela. Todos os dados dos participantes foram gravados em cabine acusticamente tratada. A hipótese básica foi a de que as produções acústicas do *tap* em onset complexo do grupo de crianças com desvio fonológico (GDF) se distinguem significativamente das do grupo de crianças sem o desvio (GSDF). A análise do *tap* se pautou pela análise espectrográfica e pela forma de onda. Quanto às hipóteses específicas, afirma-se que GDF e GSDF se distinguem quanto: (i) às oclusivas surdas, sendo o parâmetro acústico o VOT relativo; (ii) às oclusivas sonoras, sendo o parâmetro o VOT, vozeamento/(des-)vozeamento; (iii) à fricativa labiodental surda, sendo os parâmetros a duração relativa da fricativa surda e sua composição espectral. Os resultados corroboraram a hipótese básica. Já em relação à primeira hipótese específica, constatou-se que diferenças há nas produções de GDF e GSDF, mas não são significativas. Entretanto, considerando-se os efeitos da variável sexo no VOT, encontrou-se diferença significativa ( $p = 0,03571$ ) em relação à média dos valores do VOT da velar [k], sendo a média das meninas mais alta que a dos meninos. Quanto às demais hipóteses, não foram encontradas diferenças estatisticamente significativas.

**Palavras-chave:** Desvio fonológico. Fonética acústica. *Tap*.

## ABSTRACT

The purpose of this dissertation is to explain/describe acoustic patterns of tap in complex onset in the speech production of children with and without phonological disorder with respect to Acoustic Phonetics. It also aims to investigate certain phonetic details of clusters with tap, such as those involving voiced and voiceless stops, and the voiceless labiodental fricative, which precede the onset reported tap with respect to the Brazilian Portuguese phonological system. The guiding question for this research is: to what extent do acoustic productions of tap in complex onset differ in the oral production of children with and without a phonological disorder? The methodological procedure consisted of the application of a figure naming test to elicit disyllabic target words with tap on complex onset followed by the tonic vowel [a]. Tests were applied to two distinct groups of participants: one group consisting of children with a phonological disorder (N = 4) and the other consisting of children with no phonological disorder (N = 4). Data were recorded with a model H4n Zoom Recorder (Handy Recorder), with a sample rate of 44100 Hz, and a lapel microphone. All participants' data were recorded in an acoustically treated booth. The basic hypothesis was that the acoustic productions of tap in complex onset of the group of children with a phonological disorder (GDF) are significantly different from those in the group of children without the disorder (GSDF). The tap was analyzed from the waveform and a spectrographic analysis. As for the specific hypotheses, it was stated that GDF and GSDF are distinguished as: (i) voiceless stops, the acoustic parameter being the relative VOT; (ii) voiced stops, the parameters being the VOT, voicing/(un)voicing; (iii) the voiceless labiodental fricative, the parameters being the relative duration of the voiceless fricative, and the spectral composition. The results corroborated the basic hypothesis. Regarding the first specific hypothesis, it was found that there are differences in GDF and GSDF productions, but they are not significant. However, considering the effects of sex on the VOT, a significant difference ( $p = 0.03571$ ) was found in relation to the mean VOT values of velar [k], with the mean of girls being higher than that of boys. As for the other hypotheses, no significant differences were found.

**Keywords:** Acoustic phonetics. Phonological disorder. Tap.

## LISTA DE FIGURAS

Figura 1 – Modelo de primeira ordem de tubo acústico de produção de oclusiva não vozeada em três fases .....	43
Figura 2 – Modelo de primeira ordem de tubo acústico da produção de fricativa não vozeada .....	48
Figura 3 – Espectrogramas das palavras <i>pero</i> (porém) e <i>perro</i> (cachorro) do Espanhol .....	53
Figura 4 – Espectrograma de banda larga ilustrando contraste acústico entre as palavras <i>pata</i> (à esquerda) e <i>prato</i> (à direita) produzidas por um falante de Minas Gerais.....	54
Figura 5 – Espectrograma de banda larga (janela de 5 ms) de trecho da palavra <i>tribuna</i>	55
Figura 6 – Espectrograma de banda larga (janela de 5 ms) de trecho da palavra <i>tribuna</i>	56
Figura 7 – Forma de onda da sílaba tônica da palavra-alvo <i>craque</i> produzida por uma menina. ....	71
Figura 8 – Forma de onda da sílaba tônica da palavra-alvo <i>craque</i> produzida por um menino .....	71
Figura 9 – Espectrograma de banda larga da sílaba tônica <i>pra</i> da palavra <i>prato</i> produzida pelo participante 3 do GSDF.....	81
Figura 10 – Espectrograma de banda larga da sílaba tônica <i>pra</i> da palavra <i>prato</i> produzida pelo participante 2 do GSDF.....	81
Figura 11 – Espectrograma de banda larga da sílaba tônica da palavra <i>prato</i> produzida pelo participante 1 do GDF.....	82
Figura 12 – Espectrograma de banda larga da sílaba tônica da palavra <i>prato</i> produzida pela participante 3 do GDF.....	83
Figura 13 – Trajetória de F2 e F3 referentes à vogal tônica da palavra <i>prato</i> produzida pela participante 3 do GDF.....	83
Figura 14 – Forma de onda e espectrograma do <i>tap</i> no onset complexo produzido na palavra-alvo <i>trave</i> pelo participante 4 do GSDF. ....	84
Figura 15 – Forma de onda dos “ <i>taps</i> ” no onset complexo produzido na palavra-alvo <i>trave</i> pelo participante 4 do GSDF.....	85
Figura 16 – Espectrograma do <i>tap</i> no onset complexo produzido na palavra-alvo <i>trave</i> pelo participante 4 do GSDF. ....	85

Figura 17 – Espectrograma do <i>tap</i> no onset complexo produzido na palavra-alvo <i>trave</i> pelo participante 1 do GSDF. ....	87
Figura 18 – Forma de onda da palavra-alvo <i>trave</i> produzida pelo participante 1 do GDF	88
Figura 19 – Forma de onda da palavra-alvo <i>trave</i> produzida pela participante 4 do GDF	88
Figura 20 – Forma de onda da sílaba tônica <i>bra</i> da palavra-alvo <i>braço</i> produzida pelo participante 4 do GSDF.....	89
Figura 21 – Forma de onda da sílaba tônica da palavra-alvo <i>braço</i> produzida pelo participante 4 do GDF .....	89
Figura 22 – Forma de onda da sílaba <i>dra</i> da palavra-alvo <i>dragão</i> produzida pelo participante 4 do GSDF .....	90
Figura 23 – Forma de onda da sílaba pretônica da palavra-alvo <i>dragão</i> produzida pelo participante 1 do GDF, sem realização do <i>tap</i> .....	90
Figura 24 – Forma de onda da sílaba tônica <i>gra</i> da palavra-alvo <i>grade</i> produzida pelo participante 4 do GSDF .....	91
Figura 25 – Forma de onda da sílaba tônica da palavra-alvo <i>grade</i> produzida pelo participante 1 do GDF .....	92
Figura 26 – Forma de onda da sílaba tônica <i>gra</i> da palavra-alvo <i>grade</i> produzida pelo participante 3 do GSDF .....	92
Figura 27 – Forma de onda sílaba tônica da palavra-alvo <i>fraco</i> produzida pelo participante 4 do GSDF .....	93
Figura 28 – Forma de onda da sílaba tônica da palavra-alvo <i>fraco</i> produzida pelo participante 4 do GDF .....	93

## LISTA DE TABELAS

Tabela 1 – Média, desvio padrão (DesvP), coeficiente de variância (CoV) do VOT de [p], [t], [k] dos grupos GDF e GSDF. ....	69
Tabela 2 – Média, desvio padrão (DesvP), coeficiente de variância (CoV) dos valores de VOT de [p], [t], [k] produzidos pelos grupos de meninos (M) e meninas (F). ....	73

## LISTA DE GRÁFICOS

Gráfico 1 – Média de valores do VOT relativo de [p], [t], [k] dos grupos Desvio Fonológico (DF) e Controle (GSDF).....	66
Gráfico 2 – Desvio Padrão dos valores de VOT relativo de [p], [t], [k] dos Grupos Desvio Fonológico (DF) e Controle .....	67
Gráfico 3 – Cov para os valores de VOT relativo de [p], [t], [k] dos Grupos Desvio Fonológico (DF) e Controle (GSDF).....	68
Gráfico 4 – Médias dos valores de VOT relativo de [p], [t], [k] dos Grupos Masculino e Feminino.....	70
Gráfico 5 – Desvio Padrão (DP) dos valores de VOT para [p], [t], [k] em relação à variável sexo .....	72
Gráfico 6 – CoV para os valores de VOT de [p], [t], [k] em relação à variável sexo. ....	73
Gráfico 7 – Percentual de Vozeamento e Desvozeamento das Oclusivas Sonoras nos Grupos Desvio Fonológico e Controle (GSDF) .....	74
Gráfico 8 – Durações relativas da fricativa produzidas pelos participantes dos dois grupos.	76
Gráfico 9 – Espectros, com janela de 25 ms, de três fricativas labiodentais surdas produzidas pelo participante 1 do grupo sem desvio fonológico .....	78
Gráfico 10 – Espectros, com janela de 30 ms, de três fricativas labiodentais surdas produzidas pela participante 3 do grupo com desvio fonológico .....	79

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	15
<b>2</b>	<b>INCURSÕES SOBRE O DESVIO FONOLÓGICO</b> .....	22
<b>2.1</b>	<b>Algumas considerações sobre o desvio fonológico</b> .....	22
<b>2.2</b>	<b>Estudos sobre o desvio fonológico</b> .....	28
<b>3</b>	<b>TEORIA FONTE-FILTRO</b> .....	38
<b>4</b>	<b>SONS CONSONANTAIS À LUZ DA FONÉTICA ACÚSTICA</b> .....	42
<b>4.1</b>	<b>Consoantes oclusivas</b> .....	42
<b>4.2</b>	<b>VOT das oclusivas</b> .....	45
<i>4.2.1</i>	<i>Estudos sobre efeitos do sexo no VOT</i> .....	47
<b>4.3</b>	<b>Consoantes fricativas</b> .....	48
<b>4.4</b>	<b>Róticos</b> .....	50
<i>4.4.1</i>	<i>O tap, o flap e o trill</i> .....	52
<i>4.4.2</i>	<i>Características do tap em onset complexo</i> .....	53
<b>5</b>	<b>METODOLOGIA</b> .....	57
<b>5.1</b>	<b>Caracterização da pesquisa</b> .....	57
<b>5.2</b>	<b>Locus da pesquisa</b> .....	57
<b>5.3</b>	<b>Participantes da pesquisa</b> .....	58
<b>5.4</b>	<b>Procedimentos de coleta</b> .....	60
<i>5.4.1</i>	<i>Instrumentos de coleta</i> .....	60
<i>5.4.2</i>	<i>Corpus</i> .....	61
<b>5.5</b>	<b>Hipótese básica</b> .....	62
<b>5.6</b>	<b>Variáveis</b> .....	63
<b>5.7</b>	<b>Procedimentos de análise acústica dos dados</b> .....	63
<b>6</b>	<b>ANÁLISES E DISCUSSÕES</b> .....	65
<b>6.1</b>	<b>Do VOT das oclusivas surdas</b> .....	66
<i>6.1.1</i>	<i>Efeitos do sexo no VOT das oclusivas surdas</i> .....	69
<i>6.1.2</i>	<i>Análise do VOT das oclusivas sonoras</i> .....	74
<i>6.1.3</i>	<i>Análise da fricativa labiodental surda</i> .....	75

<i>6.1.4 Análise espectral da fricativa labiodental surda</i> .....	77
<i>6.2 Análise do tap no onset complexo</i> .....	79
<i>6.2.1 Análise final</i> .....	80
<b>CONCLUSÃO</b> .....	95
<b>REFERÊNCIAS</b> .....	100



## 1 INTRODUÇÃO

O que parece claro e conhecido, de familiar que é, pode não sê-lo. O avanço da ciência natural mostrou que os fenômenos do mundo externo não são na realidade como nos parecem ser. A Terra não é o centro imóvel do universo; as espécies vegetais e animais não são imutáveis [...] e a fissão de um punhado de átomos libera energia suficiente para iluminar - ou eliminar - uma populosa cidade (ao contrário do que supunham os atomistas gregos, os átomos não são indivisíveis). (GIANNETTI, 2009, p.150).

A estrutura sonora da língua se constitui a partir de sinais acústicos da fala. Esses sinais, ou sons da fala, são analógicos, isto é, variam infinitamente no instante em que se realizam, quer em seu aspecto físico vibracional, quer em seu aspecto temporal. Ora, somente esses dois aspectos já seriam suficientes, aqui, para se perceber o quão complexo pode ser o processo pelo qual se adquire tal estrutura.

Ante a complexidade desse processo, pode-se dizer que orquestrar os sons da fala de modo satisfatório nem sempre é fácil. Observa-se isso até mesmo em falantes experientes os quais podem, circunstancialmente, deslizar na produção de determinados sons, realizando, em vez destes, outros que não se esperam em dados contextos e situações, seja com ou sem consequências prejudiciais para o circuito da comunicação.

Obviamente, também se pode observar isso em falantes inexperientes, ou em aprendizes, tais quais as crianças, cujos deslizos na produção destes sons se dão com muito mais frequência, amiúde. Nas crianças, aliás, se observa, ainda, maior grau de dificuldade, uma vez que elas não adquirem a estrutura sonora de sua língua sem esforços imitativos e repetitivos.

Por outro lado, não se pode negar que, de maneira geral, os falantes conseguem usar, sim, os sons da fala de modo satisfatório; inclusive, as crianças, embora estas apresentem limitações que se lhes impõem pela idade, e das quais resultam esforços e erros relativamente normais na aquisição e no uso desses sons durante a infância. Contudo, existem aquelas crianças que parecem não se enquadrar aí, nesse perfil majoritário, por apresentarem maiores dificuldades nesse sentido. Entre as quais, encontra-se a criança com desvio fonológico.

Acredita-se que a criança com desvio fonológico esboça dificuldades especiais na aquisição e/ou no uso da estrutura sonora da língua, e que os sons de suas palavras não são percebidos da forma como se espera para sua faixa etária. Daí, dependendo da gravidade do desvio, sua fala é considerada, em menor ou maior grau, ininteligível. Segundo Ribas (2007, p.146), “O diagnóstico de desvio fonológico é obtido através da análise do sistema fonológico

e fonético da criança e dos dados da história pregressa [...]” e é caracterizado pela troca ou não-realização de sons na fala.

Todavia, os sons produzidos por essas crianças, bem como os que se produzem por outras sem desvio fonológico, ainda não foram suficientemente explorados no âmbito da Fonética Acústica. Não obstante, constam na literatura alguns estudos importantes nos quais se investigam tais sons – os quais serão abordados mais adiante.

Por ora, convém dizer que, entre os sons mais difíceis de serem adquiridos pelas crianças no português brasileiro, afigura-se o *tap*, sobretudo, quando este tem de se realizar em onset complexo<sup>1</sup>. Aqui vale lembrar que o *tap* também é comumente chamado de “r-fraco” e pode ser exemplificado como o ‘r’ que se encontra em *caro*, *prato*, *brejo*, *trigo* etc. Frise-se que em *caro*, o *tap* está em onset simples e é intervocálico; já nas outras três palavras, ele está em onset complexo, visto que é precedido por uma consoante.

Ainda no que se refere a essas três últimas palavras em itálico, interessa ressaltar nelas a presença de grupos consonantais, ou *clusters*, nos quais o *tap* aparece em onset complexo. Em *prato*, tem-se o *cluster pr*; em *brejo*, *br*; em *trigo*, *tr*. Diante disso, já podemos dizer que é nesta direção que se dá o enfoque de nossa tese a qual traz contribuições relevantes a respeito da produção acústica de *clusters* emanados da fala de crianças com e sem desvio fonológico. Eis aí um dos motivos por que se lhe propõe o título “Análise acústica de *clusters* com *tap* na fala de crianças com e sem desvio fonológico”.

Na literatura, se encontram estudos que, de certo modo, nos precederam e, com base neles, pode-se dizer que a aquisição fonológica do português brasileiro se completa por volta da idade de 5 anos, momento em que se adquirem as consoantes líquidas coronais - entre as quais o *tap* – ocupando o onset complexo (HERNADORENA, 1990; LAMPRECHT, 1990; MEZZOMO, 1999; MIRANDA, 1996; RIBAS, 2002). Destaque-se que “[...] o onset complexo é a última estrutura silábica a ser adquirida pelas crianças com desenvolvimento fonológico normal.” (RIBAS, 2007, p. 149). Mas - frise-se - esses estudos referem-se à aquisição fonológica “normal”, ou típica. Já no que se refere à aquisição fonológica atípica, como a que envolve o desvio fonológico, é diferente.

---

<sup>1</sup> O onset complexo pode ser entendido como um grupo consonantal de uma estrutura silábica CCV, formada por uma consoante (na primeira posição) C1 + outra consoante (na segunda posição) C2 + uma vogal V. Exemplo: na sílaba tônica *pra* da palavra *prato*, tem-se uma estrutura CCV, que é *pra*, e, nesta sílaba, o onset complexo corresponde às posições CC, que são ocupadas, respectivamente, pelas consoantes *pr*. Quando se diz que o *tap* está em onset complexo, significa dizer que ele está na segunda posição desse onset, ou seja, em C2. Ressalte-se, ainda, que outros termos também podem ser usados para se referir ao onset complexo, entre os quais, onset silábico, ataque complexo, encontro consonantal, grupo consonântico e/ou *cluster*.

Quanto a esta aquisição, estudos sinalizam que a maioria das produções “incorretas” das crianças com desvio fonológico são aquelas que envolvem como alvo o onset complexo (CASARIN, 2006; RAMOS, 1996; RIZZOTTO, 1997; VIDOR, 2000). Nesse diapasão, ainda, um estudo de Ribas (2009, p.74) aponta que 92% dos sujeitos com desvio fonológico, com idades entre 5 e 10 anos, não têm a estrutura silábica com onset complexo adquirida.

No que se refere à aquisição de segmentos, convém destacar duas coisas: a primeira, relativa a um fato já conhecido, ou seja, o de que na sintaxe fônica do português brasileiro somente o *tap* e a líquida lateral podem ocupar a posição da segunda consoante do onset complexo na estrutura silábica CCV; a segunda, referente a um estudo de Ribas (2011), conforme o qual o maior percentual de produções “incorretas” de consoantes é correspondente ao *tap*, que atinge a marca de 63,9%, ou seja, aproximadamente o dobro do percentual atingido pela líquida lateral, 31,3%.

Todos esses estudos aqui reportados - embora estejam metodologicamente pautados em análise de oitiva - trouxeram importantes contribuições para a aquisição fonológica do português brasileiro. No entanto, convém lembrar o que há muito já dizia, com ampla pertinência, Câmara Jr. (1954, p. 36), sobre a análise rigorosa dos sons linguísticos: “A nossa consciência linguística não toma conhecimento de certas diferenças articulatórias e acústicas existentes, que a observação fonética rigorosa apreende.”

Hoje, porém, com os avanços e recursos tecnológicos de que dispomos, a análise acústica, em termos metodológicos, está pelo menos um degrau acima da rigorosa observação fonética à qual Câmara Jr. alude.

Ora, mas os estudos que se reportaram até aqui, como já se aludiu, não utilizaram análise acústica. Basearam-se apenas em análises de oitiva, ou seja, em impressões sensoriais auditivas das quais determinadas diferenças acústicas escapam facilmente. Desse modo não se tem acesso a determinados detalhes fonéticos relevantes na produção oral, muitos dos quais podem ter passado despercebidos nesses estudos ancorados em oitiva, os quais se mostram lacunosos quanto a isto. Tais detalhes, vale destacar, só são apreendidos mediante análise acústica.

Também é importante dizer que esses estudos mesmos se concentram todos no Sul e no Sudeste do país, embora também se possam encontrar outros estudos sobre aquisição fonológica, quer típica ou atípica, em outras regiões, como o Nordeste, especialmente, na cidade de Salvador, em que se encontram precursores como Teixeira (1986, 1991, 2009). Entretanto, estudos assim ainda são escassos nessas regiões, ou seja, naquelas que estão para além da eixo Sul-Sudeste. A este respeito, Vogeley afirma que:

Apesar dos grandes avanços e pesquisas sobre os desvios fonológicos, no Brasil, ainda são necessários mais estudos que proporcionem não apenas um panorama da aquisição fonológica, considerando os diversos dialetos e especificidades [...] A isso, soma-se a necessidade de investigações nas regiões Norte e Nordeste do país, dada a carência de descrições, inclusive sobre a aquisição fonológica normal. (VOGELEY, 2011, p. 162).

Diante disso, concordamos com Vogeley (2011) e reafirmamos a necessidade de se desenvolverem mais trabalhos nessa direção, porém, não só no sentido de preencher a lacuna geográfica que se identificou anteriormente, mas também para se obter uma visão que se compatibilize de fato com a realidade da aquisição fonético-fonológica do português brasileiro, levando-se em conta, aí, suas variações diatópicas e outras, bem como sua heterogeneidade linguística como um todo.

Foi nessa esteira que os estudos e as lacunas aqui reportados nos inspiraram em direção ao desenvolvimento desta tese a qual, diferentemente deles, se pauta em análise acústica. Além disso, vale frisar, situa-se no Nordeste, especificamente, na cidade de Fortaleza.

Nesse sentido, ainda, convém destacar o pioneirismo desta tese frente ao Nordeste como um todo, na medida em que se considera seu enfoque sobre a fala de crianças com e sem desvio fonológico, seus procedimentos metodológicos, dos quais se obtiveram dados que emanaram de gravações em cabine acusticamente tratada, e seus procedimentos analíticos, os quais englobam análises acústicas de *clusters* com *tap* na fala dessas crianças, trazendo evidências, por meio de espectrogramas e formas de onda; evidências estas que, aliás, nunca foram antes expostas por trabalhos anteriores, nem no Nordeste, nem em quaisquer outras regiões do país.

Cumpre, entretanto, fazer, aqui, a seguinte ressalva, segundo a qual se deve evitar a suposição ingênua de que a análise acústica consistira apenas em testar hipóteses impressionísticas do sinal de fala, como as que decorrem dos estudos ancorados em análise de oitiva. Neste particular, concordamos com o argumento de Silva (2010), conforme o qual:

[...] o estatuto da análise acústica não é o de confirmar hipóteses impressionísticas do sinal de fala. Ao contrário, assumindo-se a perspectiva de comensurabilidade entre fonética e fonologia, a análise acústica tem precedência sobre a análise de oitiva, dado que permite observar o detalhe que, muitas vezes, é essencial para se compreenderem fatos fônicos que, de outro modo, permaneceriam obscuros. (SILVA, 2010, p. 215-216).

Agora interessa reportar alguns estudos que, diferentemente dos que foram apresentados anteriormente, utilizam em suas abordagens recursos fonético-acústicos, como a análise acústica. Começamos com os trabalhos de Miranda (2007) e Cristóforo-Silva e Miranda (2011), os quais se ancoram teoricamente em modelos multirrepresentacionais como a Fonologia de Uso e a Teoria de Exemplos.

Miranda (2007), por seu turno, busca avaliar a relevância do detalhe fonético na aquisição da estrutura CCV, enquanto Cristóforo-Silva e Miranda (2011) buscam discutir a natureza das representações fonológicas ao avaliar estratégia de construção de linguagem. Aliás, essas autoras (2011), vale frisar, também investigam a hipótese do alongamento compensatório da vogal nuclear em estruturas silábicas CCV e CV.

Ambos os estudos trazem contribuições importantes nessa direção. No entanto, deixam de explorar uma série de detalhes fonéticos que poderiam ser investigados nas estruturas silábicas enfocadas em suas abordagens, tais como o vozeamento das oclusivas, entre outros.

Tais lacunas também inspiraram, em certo aspecto, o desenvolvimento de uma parte de nossa pesquisa. Neste sentido, investigamos os reportados detalhes - quer nas produções de crianças sem desvio fonológico, como Miranda (2007) e Cristóforo e Miranda (2011) o fizeram - quer nas de crianças com o desvio fonológico, as quais não foram contempladas por nenhuma das mencionadas autoras.

Já os estudos de Mezzomo *et al.* (2008), Mezzomo, Mota e Dias (2011) e Giacchini, Mota e Mezzomo (2011) englobam as produções de crianças com desvio fonológico além de convergirem com Miranda e Cristóforo e Miranda, no sentido de lançarem mão de recursos fonético-acústicos e de investigarem o alongamento compensatório da vogal em estruturas CCV e CV. Todavia, não analisam, com certas especificidades, *clusters* com *tap*, tal como desenvolvido nesta tese.

Outros trabalhos também lançaram mão de recursos fonético-acústicos, alguns dos quais se mostram comprometidos com a Fonologia Acústico-Articulatória, de Albano (2001), quais sejam: Berti (2006, 2010, 2011), Rodrigues (2007), Freitas (2007). Em linhas gerais, todos eles exploram contrastes fônicos, contrastes encobertos (*covert contrast*), entre outros, nas produções de crianças com e/ou sem desvio fonológico, estabelecendo comparações de cunho quantitativo e qualitativo. Algumas de suas lacunas serão discutidas, mais adiante, em outro capítulo desta tese.

Já no âmbito exclusivo, por assim dizer, da Fonética Acústica, e no que concerne, especificamente, aos sons produzidos por esses grupos de crianças, consta apenas o trabalho

seminal de Panhoca-Levy (1993), o qual inspirou, em certa medida, alguns dos estudos reportados anteriormente, entre os quais o de Berti (2006).

É com base em todas as considerações acima que nossa abordagem se destina a investigar a produção de *clusters* com *tap* na fala de crianças com e sem desvio fonológico, considerando a Fonética Acústica, não levando em conta apenas seu instrumental metodológico, mas também seu arcabouço teórico. Isto, de certo modo, confere a esta tese uma originalidade peculiar que a distingue de todas as outras que a antecederam, o que, todavia, não as desabona nem de longe.

Assim, esta tese se situa num quadro teórico da Fonética Acústica, o qual se construiu aqui a partir de autores como Lisker e Abramson (1964), Fry (1976), Cho e Ladefoged (1999), Stevens (1999), Barbosa e Madureira (2015), Kent e Read (2015), entre outros. E é a partir deste quadro teórico que analisamos dois grupos distintos de crianças, sendo um deles constituído de crianças com desvio fonológico (GDF) e outro de crianças sem o desvio (GSDF).

A pergunta norteadora desta pesquisa é esta: em que medida GDF se distingue de GSDF na produção acústica de *clusters* com *tap*?

Já o objetivo central da tese consistiu em interpretar/descrever padrões acústicos do *tap* em onset complexo na fala de crianças com e sem desvio fonológico à luz da Fonética Acústica. Os objetivos específicos consistiram em investigar e comparar *clusters* com *tap*, especialmente, as consoantes que aí o precediam. As análises desses *clusters* se pautaram nos seguintes parâmetros acústicos: VOT relativo das oclusivas surdas que precediam o *tap*; VOT das oclusivas sonoras que precediam o *tap* – tendo em vista, neste particular, o vozamento e o desvozeamento destas consoantes; a duração relativa da fricativa labiodental surda e os espectros de Fourier desta fricativa.

A hipótese básica desta tese foi a de que o grupo de crianças com desvio fonológico (GDF) se distingue significativamente do grupo de crianças sem o desvio (GSDF) na produção acústica de *clusters* com *tap*, especialmente, na produção do *tap* em onset complexo.

Decorreram daí as seguintes hipóteses específicas e previsões:

- O padrão acústico das produções do GDF referentes ao *tap* em onset complexo se caracteriza pela não-realização do *tap*, isto é, pela ausência de vestígios do sinal acústico referentes ao *tap* mesmo.
- O padrão acústico das produções do GSDF referentes ao *tap* em onset complexo se caracterizará como uma estrutura acústica CCV, pois os participantes deste grupo

tenderão a realizar o *tap*, porém na posição desta consoante poderão, ainda, produzir em seu lugar variantes como *tap* fricativo alveolar surdo ou sonoro, vibrante múltipla alveolar, entre outras.

- GDF e GSDF se distinguem quanto às medidas do VOT das oclusivas surdas, sendo seus valores significativamente menores no GDF.
- GDF e GSDF se distinguem quanto ao desvozeamento das oclusivas sonoras, sendo tal processo mais frequente no GDF.
- GDF e GSDF se distinguem quanto à duração relativa das fricativas labiodentais, sendo maior no GDF.
- GDF e GSD se distinguem quanto à composicionalidade espectral das fricativas labiodentais.
- Nos encontros consonantais em que o *tap* é precedido por consoantes oclusivas alveolares [t] e [d] poderão ocorrer, além do *tap*, variantes como vibrantes múltiplas alveolares no GSDF.

Tais hipóteses específicas se construíram, em certa medida, a partir de resultados de determinados trabalhos, alguns dos quais já reportados nesta introdução, e de outros que são resenhados e discutidos, mais adiante, em outros capítulos.

Feitas essas considerações, resta dizer que esta tese se estrutura em 7 capítulos. No capítulo segundo, que é o próximo, têm-se algumas considerações sobre o desvio fonológico, além de resenhas sobre trabalhos que lançaram mão de análise acústica, quer sobre produções orais de crianças sem desvio fonológico, como os de Miranda (2007) e Cristófaros-Silva e Miranda (2011), quer sobre produções de crianças com o desvio, quer em ambas os casos, como os de Panhoca-Levy (1993), Berti (2006, 2010, 2011), Rodrigues (2007), Freitas (2007) e Vassoler (2017). No terceiro capítulo aborda-se a Teoria Fonte-Filtro. O quarto capítulo trata dos sons consonantais à luz da Fonética Acústica. No quinto capítulo consta a Metodologia empregada na pesquisa. No sexto capítulo encontram-se as análises e discussões e, por fim, no capítulo sétimo, tem-se a conclusão.

## 2 INCURSÕES SOBRE O DESVIO FONOLÓGICO

Apresentam-se neste capítulo estudos sobre o desvio fonológico, alguns dos quais nos permitem entrever questões de cunho teórico-conceitual em relação a este tema, além de outros que nos permitem focar questões metodológicas e práticas atinentes a este respeito, uma vez que lançam mão de recursos instrumentais e teóricos advindos, em parte, da Fonética Acústica combinados com alguma teoria fonológica

Para tanto, este capítulo se estrutura em dois tópicos. No primeiro, aborda-se o desvio fonológico, com o objetivo de discutir problemas conceituais a seu respeito, além de buscar interpretar e desvelar sua natureza e causa, as quais têm sido perspectivadas como desconhecidas ao longo de sua história. Já, no tópico segundo, o enfoque se dá sobre problemas metodológicos e práticos encontrados nos mencionados estudos.

### 2.1 Algumas considerações sobre o desvio fonológico

Na aquisição do sistema fonético-fonológico, a criança apresenta normalmente dificuldade em lidar com a produção dos sons das palavras. Falantes mais experientes, especialistas ou não no assunto, também percebem isto na criança. Com efeito, tornou-se lugar comum relacionar tal dificuldade com a troca e/ou com a omissão de determinados sons da língua durante o processo de aquisição fonético-fonológica da criança.

Acontece que à proporção que a idade, as experiências e os “ensaios” da criança (com a língua) vão aumentando, a dificuldade com tais sons vai, comumente, se reduzindo. Daí, suas produções, isto é, os sons de sua fala, vão se ajustando aos padrões da língua-alvo. Aqui vale lembrar o que já dizia Câmara Jr. (1954, p. 230) a este respeito: “A criança aprende os fonemas da língua, empiricamente, pela perspicácia da audição e reprodu-los depois de esforços contínuos, pois raramente acerta de início.”

Entretanto, quando isso não ocorre normalmente, de modo que a dificuldade com os sons se mostra persistente, a ponto de as produções da criança não se enquadrarem em padrões previstos para sua faixa etária, é provável que nela se caracterize um desvio fonológico.

Como se vê, o desvio fonológico não se confunde com aquela dificuldade relativamente normal das crianças na aquisição do sistema fonético-fonológico, à qual se aludiu anteriormente. Na verdade, trata-se de uma dificuldade especial na produção dos sons das palavras, razão por que tais crianças não se enquadram nos padrões previstos para sua



faixa etária; pelo menos é assim que o desvio fonológico vem se desenhando no imaginário do vigente panorama científico.

Na literatura, podem se encontrar trabalhos como o de Grunwell (1990), segundo o qual existem determinadas características que ajudam a identificar crianças com desvio fonológico, quais sejam: i) fala espontânea com erros de produção de sons consonantais; ii) idade superior a 4 anos; iii) limiares auditivos dentro dos padrões de normalidade; iv) ausência de alterações neurológicas evidentes; v) habilidades cognitivas normais; vi) capacidade de compreender a fala; vii) linguagem expressiva sem alterações com relação ao léxico e à sintaxe.

Em (i), como se pode notar, destaca-se a fala espontânea com erros. Essa característica, a meu ver, também pode ser perspectivada como fala com inteligibilidade alterada, situando-se numa escala que vai da fala ininteligível à pouco inteligível, ou vice-versa. Isso tem relação com a gravidade do desvio fonológico. Quanto menos inteligível for a fala, mais grave será o desvio. Em (ii), pressupõe-se que, antes da idade de 5 anos, é praticamente impossível de se caracterizar na criança o desvio fonológico. Já a partir de (iii), (iv), (v), (vi) e (vii), pode-se dizer, em suma, que a criança com desvio fonológico - embora apresente condições relativamente normais para que produza os sons das palavras dentro dos padrões previstos para sua faixa etária - ela não os produz.

Cumprе ressaltar que a gravidade, ou grau, do desvio fonológico é geralmente se subdividido hierarquicamente em níveis, tais como leve, moderado, grave e gravíssimo (ou severo); e que a população com desvio fonológico é, segundo Gierut (1998) e Wertzner (2001), heterogênea em relação ao modo de manifestação do problema e ao grau de severidade.

Albano (2009), por seu turno, afirma, com base na Fonologia Gestual, que:

[...] crianças portadoras dos chamados desvios fonológicos fazem distinções sistemáticas, porém às margens do padrão convencional, entre os pares de consoantes que parecem confundir. Tais contrastes - que denominamos encobertos [...] são, entretanto, imperceptíveis para os seus interlocutores, que, em geral, os assimilam a um único membro do par contrastante. Com isso, a criança cai presa do estigma de “falar errado”, torna-se ansiosa e acaba por permanecer numa fase intermediária da aquisição do contraste, em que a distinção é realizada, mas de maneira ainda não ótima. (ALBANO, 2009, p. 234-235).

Com base nessas considerações, percebe-se que Albano (2009) não diverge de Grunwell (1990) no sentido de reconhecer nas crianças com desvio fonológico produções fônicas que não se compatibilizam com o padrão convencional.

Vale ressaltar, ainda, dois aspectos importantes observados por Albano (2009). Primeiro, a criança com desvio fonológico apresenta contrastes encobertos em suas produções fônicas os quais são imperceptíveis para os seus interlocutores. Logo, abordagens tradicionais, restritas a impressões de oitiva, podem chegar a equívocos diversos nesses casos, uma vez que, sem análise acústica, tais contrastes não são percebidos com certa nitidez. Segundo, essa criança mesma pode sofrer pelo estigma de “falar errado” e, daí, seguir para um quadro de ansiedade. Assim, tem-se aí uma espécie de influência mútua entre desvio fonológico e ansiedade, que se fomenta, em certa medida, através de tal estigma.

Outro aspecto não menos importante em relação ao desvio fonológico parece estar associado a uma lacuna etiológica, qual seja: sua causa é desconhecida. A meu ver, essa lacuna etiológica impactou, ao longo do tempo, o conceito de desvio fonológico, sobre o qual, cumpre ressaltar, ainda hoje não há consenso entre pesquisadores.

Tanto é que o conceito de desvio fonológico já teve outros nomes, ou, por assim dizer, denominações. Estas, por sua vez, foram mudando ao longo do tempo como reflexo de mudanças teórico-epistemológicas que se davam no âmbito da Linguística e, especialmente, no da Fonética e no da Fonologia.

Nesse sentido, Panhoca-Levy (1993), autora que aborda com pioneirismo no Brasil o desvio fonológico à luz da Fonética Acústica, afirma o seguinte:

Teixeira (1990) chama de “Fonologia Clínica” e Yavas, Hernandorena e Lamprecht (1991) de Linguística Clínica essa vertente da linguística contemporânea preocupada com as patologias de linguagem que propõe que os desvios articulatórios funcionais passem a ser considerados como “desvios fonológicos evolutivos”, mudando, então, o enfoque, que passa de articulatório e motor a cognitivo-organizacional. (PANHOCA-LEVY, 1993, p.20).

Como visto, o desvio fonológico nos anos 1990 tivera outras denominações: foi primeiramente denominado como desvio articulatório funcional, sendo mudado depois para desvio fonológico evolutivo. Para tanto, nos bastidores dessas mudanças, segundo Panhoca-Levy, havia também uma mudança de enfoque que ora se adotava por pesquisadoras como Teixeira (1990), Lamprecht, Yavas e Hernandorena (1991). Essas, por seu turno, abandonavam naquela década a perspectiva articulatória e motora que se atribuía à natureza do fenômeno em pauta, passando, em seguida, a perspectivá-lo a partir de uma suposta natureza cognitivo-organizacional dissociada ontologicamente de suas bases articulatório-motoras.

Esta separação ontológica traz implicações lógicas sobre o estudo e o tratamento daquilo que se concebe como desvio fonológico, como se pode ver na citação a seguir.

[...] as pesquisas em Fonologia Clínica têm viabilizado a caracterização dos desvios, enquanto uma alteração ou uma demora na organização do sistema fonológico, e não mais como uma alteração articulatória, inserindo o estudo e o tratamento dos desvios fonológicos numa perspectiva muito mais abstrata ou fonológica. (VOGELEY, 2011, p. 152).

Ora, o que está por trás disto não é outra coisa senão a assunção de um duplo ontológico, ancorado no dualismo gerativista, no qual se separam esquemas articulatório-motor e cognitivo-organizacional, não apenas para fins metodológicos (ou terapêuticos, como almejava a Fonologia Clínica), mas, antes, para afirmá-los em seus bastidores epistemológicos. Desse modo, tem-se aí, em certo aspecto, uma adesão dos estudos de Fonologia Clínica ao psicologismo do gerativista Noam Chomsky, para o qual a linguagem está ontologicamente dissociada de suas bases motoras, articulatórias e outras, além de ser idealizada como ‘puramente’ mental, abstrata, homogênea, perfeita e universal.

Sob minha ótica, a qual diverge de Chomsky e de seus seguidores, entendo que os esquemas articulatório-motor e cognitivo-organizacional são distintos; porém não estão separados. Eles pertencem a uma mesma unidade da qual não podem se separar. E esta unidade é o falante real, de carne e osso, com todas as suas imperfeições, o qual se distingue radicalmente do falante ideal de Chomsky, que está alienadamente isolado da realidade. Quanto a isto, resta, ainda, lembrar que, nos anos 1990, quase todos os estudos que trabalharam com os mencionados esquemas como se fossem separados se filiam ao psicologismo chomskyano.

Agora voltando um pouco mais no tempo, consta que foi a partir dos anos 1970 que surgiram as primeiras denominações atinentes a atual denominação desvio fonológico. A primeira, de acordo com Rossi-Barbosa *et al.* (2011) e Wertzner (2011), foi a dislalia. Depois, foram emergindo sucessivamente as seguintes: atraso de fala; atraso no desenvolvimento da fala; distúrbio articulatório; distúrbio articulatório funcional; desvio de articulação e distúrbio fonológico do desenvolvimento.

Apesar dessas mudanças, ainda não há consenso sobre tais denominações, como já dito antes. Tanto é que Wertzner não usa a denominação desvio fonológico. Em seu lugar, opta por “transtorno fonológico”, termo este que não aparece sequer na lista acima, mas que se originou do inglês *phonological disorder*, muito utilizado por autores gerativistas, como, por exemplo, Gierut. Vale frisar que, neste particular, Wertzner se alinha à Gierut (1990), para

quem o transtorno fonológico se define como aquele que afeta a produção da fala e/ou a representação mental dos sons de uma língua.

A denominação transtorno fonológico, embora possa aludir ao mesmo fenômeno aludido pelo desvio fonológico, também se caracteriza tecnicamente no âmbito da Psiquiatria enquadrando-se nos moldes dos transtornos mentais descritos pela Associação Americana de Psiquiatria (APA) no Manual Diagnóstico e Estatístico de Doenças Mentais (DSM-IV)<sup>2</sup>.

Frise-se que, além de Gierut e Wertzner, outros pesquisadores, principalmente os da área médica, também optam pela terminologia descrita no DSM.

Por outro lado, há autores, em Portugal, como Barroco *et al.* (2007), os quais optam pela denominação perturbação fonológica. No Brasil, vale frisar, Barbosa e Madureira (2015) também usam essa denominação ao reportarem o trabalho de Barroco *et al.*

Há ainda aqueles que optam pela denominação distúrbio fonológico, como Berti que o define da seguinte forma:

O distúrbio fonológico é descrito como uma desordem linguística em que se observa uma dificuldade – por parte da criança - quanto ao conhecimento dos segmentos fonéticos, das regras fonológicas, ou na maneira em que utilizam esse conhecimento. (BERTI, 2011, p. 154).

Vale ainda mencionar a opção terminológica de Rodrigues (2007), que se opõe a todas as outras, optando por “queixa fonoaudiológica relacionada ao sistema fônico”. Para defender sua opção, a autora alega que em seu estudo:

[...] as dificuldades de pronúncia das crianças não são consideradas nem como ‘distúrbio articulatorio’ nem como ‘desvio fonológico’ ou ‘distúrbio fonológico’ (terminologias consagradas na Fonoaudiologia para o diagnóstico de crianças com dificuldade de pronúncia. No primeiro caso (distúrbio articulatorio), as dificuldades de fala das crianças são entendidas como decorrentes de alterações articulatorias – na estrutura e/ou atividade motora do aparelho fonador. Nos outros dois (desvio fonológico e distúrbio fonológico), é atribuída à criança uma dificuldade no conhecimento dos segmentos fonológicos e das regras fonológicas ou na maneira como utiliza esse conhecimento, não sendo, portanto, do nível articulatorio. [...]. No presente trabalho, os aspectos motor/articulatorio e simbólico não são desvinculados, sendo que as produções das crianças são permeadas sempre por uma relação entre esses aspectos. (RODRIGUES, 2007, p. 1).

Como se viu, foi possível evidenciar aqui o quão diferentes são as denominações que se atribuíram ao desvio fonológico, quer em décadas anteriores, quer no momento atual.

---

<sup>2</sup> A última versão deste Manual, no caso, o DSM-V, vem sendo amplamente criticada e até, em muitos casos, rejeitada, por pesquisadores diversos, desde a sua publicação.

Ante a essa instabilidade conceitual da qual se impõe a falta de consenso entre os pesquisadores mencionados aqui, a impressão que se tem é a de que o desvio fonológico não está fundamentado linguisticamente no falante, nem na fala, tampouco na língua; mas na história mesma em que ele se origina. Desse modo, a não observação do fator histórico através do qual este conceito se constitui pode ser, talvez, um dos motivos por que sua etiologia ainda é desconhecida.

Diante disso, esse conceito, nos parece, é ainda um “pré-conceito”, quer no sentido pejorativo do termo, quer no sentido do incipiente momento histórico em que a ciência linguística ora o emprega.

Trata-se de um conceito que, além de se constituir a partir de idealizações e generalizações arbitrárias, está assentado numa noção de normalidade a qual não foi suficientemente discutida. Como sustentar cientificamente, por exemplo, que um falante se desvia da “normalidade fonológica” se a categoria falante não está linguisticamente vinculada a pressupostos de normalidade ou anormalidade? Se tais coisas estiverem supostamente pautadas no falante ideal de Chomsky, cumpre dizer que elas também só farão sentido no imaginário teórico deste, e não na realidade mesma, tampouco no falante real.

Essa noção de normalidade/anormalidade fonológica só encontraria apoio no idealismo dualista de Chomsky, especialmente, em seu falante ideal, a partir do qual se torna possível pressupor, fabricar e separar, ainda que absurdamente, “falantes sadios” de “falantes doentes”. Isso, entretanto, é velado e mascarado de várias formas em estudos que dividem os falantes nesses moldes, os quais se ancoram no idealismo chomskyano.

Ainda nesse diapasão, interessa dizer que os estudos existentes sobre o conceito de desvio fonológico - e mesmo os de aquisição fonético-fonológica do português brasileiro - não são suficientes para se propor a partir deles um padrão “ideal” que distinga crianças com desvio fonológico de crianças sem o desvio.

Tais estudos, em geral, não atentam para a heterogeneidade linguística e, assim, muitas vezes, não observam variações diatópicas, diastráticas, entre outras, que podem implicar o processo de aquisição, quer seja típica ou atípica. Aliás, quase que a totalidade dos estudos em aquisição fonológica, como já aludido anteriormente, se encontra no Sul e no Sudeste do país. Daí se impõe a necessidade de mais estudos em outras regiões brasileiras, para se ter uma visão mais condizente em relação à aquisição fonético-fonológica vinculada à heterogeneidade linguística, bem como em relação ao conceito de desvio fonológico e suas implicações.

Registradas nossas considerações e críticas em relação, entre outras coisas, ao conceito, ou pré-conceito, de desvio fonológico, cumpre dizer que, não obstante, teremos de lançar mão deste ou de outro conceito, ou denominação, ao longo desta tese.

Reitere-se, entretanto, que, sob nossa ótica, particularmente, e aí se considerando a forma como tais conceitos vêm sendo perspectivados, linguística e historicamente, entendemos que eles, isto é, esses conceitos mesmos de que ora estamos tratando - desvio fonológico, transtorno fonológico etc. - não passam de meras ficções convencionais, as quais servem apenas para fins de designação, de descrição. Ou seja, não fazem sentido para fins de explicação.

Para se aspirar a um conceito que de fato se compatibilize com o fenômeno ao qual alude, é preciso, antes de tudo, poder pensar a razão de ser desta própria relação (conceito-fenômeno), no seio da qual o conceito é apreendido. E a razão de ser desta relação, como se viu, aqui, tem sido pautada mais na história mesma em que estes conceitos se constroem do que em supostos fundamentos linguísticos. Daí a história desses conceitos, em última análise, não deve ser ignorada, sobretudo, quando se busca utilizá-los para fins explicativos.

Considerando toda a discussão acima, admitir-se-á, nesta tese, portanto, o uso de um destes conceitos apenas para fins de designação. Nesse sentido, escolhemos o conceito de desvio fonológico, denominação a partir da qual se distinguem os grupos de crianças abordados nesta pesquisa. O que justifica tal escolha se deve, basicamente, ao fato de sua caracterização, tal como proposta por Grunwell (1990), se revelar, relativamente, com mais nitidez que as outras aqui reportadas, permitindo identificá-lo na criança de forma mais prática e objetiva, não obstante suas lacunas.

## **2.2 Estudos sobre o desvio fonológico**

Um estudo pioneiro na literatura nacional sobre a aquisição fonético-fonológica atípica à luz da Fonética Acústica é a tese de doutorado de Panhoca-Levy (1993). Essa autora investiga o contraste de vozeamento em crianças (N=3), com distúrbio articulatorio, denominação esta que se compatibilizaria hoje com a de desvio fonológico, como já discutido antes.

Na tese de Panhoca-Levy, realizaram-se as análises acústicas sobre o mencionado contraste de vozeamento, a partir do espectrógrafo digital *Kay Elemetrics DSP Sona-Graph Model 5500*. As conclusões indicam haver nessas crianças a presença de comprometimentos

neuro-motores, em graus diversos. Nesse sentido, a autora afirma serem suas dificuldades, na verdade, fundamentalmente, fonéticas e não fonológicas. Ressalta ainda que:

As alterações encontradas não são cognitivas. Ao contrário, as compensações motoras e articatórias e as buscas e tentativas de aproximação da produção correta observadas em todas as crianças aqui analisadas, indicam claramente, que elas sabem o que deveriam estar produzindo se tivessem as condições neuro-motoras de órgãos fonoarticulatórios. (PANHOCA-LEVY, 1993, p. 203).

Outro aspecto que interessa destacar a partir de Panhoca-Levy (1993) diz respeito ao fenômeno do contraste encoberto, o qual, segundo a autora, aparece nas produções das crianças investigadas em sua tese.

Ainda em relação a esse fenômeno, Corrêa e Berti (2015), por sua vez, afirmam o seguinte:

A expressão ‘contraste fônico encoberto’ (*covert contrast*) foi utilizado por Hewlett (1988) para descrever o que é categorizado como contrastes fônicos imperceptíveis auditivamente, mas detectáveis acústica e/ou articulatoriamente. Dessa forma, uma produção julgada auditivamente como omissão ou substituição categórica pode revelar, a partir da análise acústica e/ou articulatória, que o falante está produzindo sistematicamente diferenças para distinguir sons. [...]. (CORRÊA; BERTI, 2015, p. 365).

Nos estudos sobre tais contrastes encobertos, destaca-se, entre outros, o trabalho de Scobbie (1998 *apud* CORRÊA; BERTI, 2015), que realiza uma ampla revisão bibliográfica abrangendo diversos autores que até então haviam descrito a presença de contraste encoberto, quer em crianças com desenvolvimento típico, quer em crianças com desvio fonológico.

Assim, Scobbie constata que os contrastes encobertos se identificam na estrutura silábica, no modo de articulação, no vozeamento e no ponto de articulação. Daí sugere que, em princípio, nenhum parâmetro envolvido na aquisição de um determinado contraste fonológico é imune aos contrastes encobertos.

O fenômeno do contraste encoberto, vale ressaltar, não deve ser confundido com o processo de neutralização, conceito este proposto por Trubetzkoy (1949), segundo o qual a oposição entre dois fonemas, em determinados contextos, pode ser suspensa, resultando daí um arquifonema. E este, por sua vez, se define pelos traços que correspondem à interseção dos dois fonemas neutralizados, podendo, assim, se realizar ora como um fonema ora como outro. Repare-se, aí, que na neutralização o que se tem é uma ausência de contrastes entre fonemas, da qual resulta a categoria arquifonema; enquanto que no fenômeno do contraste encoberto o que está implicado são oposições entre a realização acústica (e/ou articulatória) do contraste mesmo e, *pari passu*, a sua não percepção auditiva.

O fenômeno do contraste encoberto foi também analisado por Berti (2006). Nesse trabalho, a autora investiga o estabelecimento do contraste entre as fricativas /s/ e /ʃ/ em crianças com e sem queixas fonoaudiológicas falantes do Português Brasileiro. Assume a hipótese de que as crianças experienciam diferentes formas de marcar uma distinção fônica percebida por elas, realizando isso, muitas vezes, por meio de produções gradientes (intermediárias). Para tanto, filia-se à Fonologia Acústico-Articulatória, de Albano (2001). Realiza dois estudos: um relativo à produção da fala; outro, à percepção.

O seu estudo de produção abrange uma análise acústica de crianças (N=6), com queixas fonoaudiológicas, de ambos os sexos entre cinco e sete anos de idade. Foram selecionadas seis palavras dissílabas paroxítonas do PB contendo as fricativas /s/ e /ʃ/ acompanhadas das vogais /a/, /i/ e /u/ e inseridas numa frase veículo. As crianças tinham de realizar dez repetições de cada frase veículo. Adotaram-se parâmetros acústicos relativos às características acústicas do ruído das fricativas e parâmetros relativos ao padrão temporal.

Estes indicaram que as crianças com queixas fonaudiológicas utilizam um tempo muito maior para produzir um trecho de fala. A partir daí a autora diz que tais crianças apresentam uma velocidade de movimentação dos articuladores bem inferior à velocidade apresentada pelo grupo sem queixa fonoaudiológica.

A nosso ver, essas relações podem até ser depreendidas em alguns estudos, porém de forma aleatória; ou seja, podem ocorrer, mas não necessariamente. Note-se que Berti não se apóia em teste estatístico inferencial para assegurar a validade dessas diferenças.

Por outro lado, a autora ressalta que um dos principais resultados obtidos em seu estudo foi a constatação de uma produção gradiente (produções intermediárias) nas crianças com queixas fonoaudiológicas entre as duas categorias fônicas estudadas.

Já no que concerne ao seu estudo de percepção, busca a analisar a consequência perceptual dos gradientes fônicos. Para isso, realiza um experimento de identificação, envolvendo estímulos típicos (produções categóricas das fricativas /s/ e /ʃ/ e estímulos intermediários (produções gradientes de tais fricativas), com 35 sujeitos, os quais não participaram do estudo de produção da fala e foram divididos em três grupos distintos: adultos, crianças com problemas de produção de fala e crianças sem problemas de produção de fala. Afirma que os resultados obtidos apontam para uma relação intrínseca entre produção e percepção de fala para todos os grupos de sujeitos. Sugere, ainda, que se considere a natureza dinâmica da produção da fala, contemplando-se produção e percepção em uma mesma unidade, além da adoção de perspectivas dinâmicas do desenvolvimento nos estudos referentes à aquisição dos contrastes fônicos.



Tendo em vista tais contrastes fônicos em crianças com padrões não esperados para sua faixa etária, Freitas (2007), por seu turno, busca caracterizar aspectos sincrônicos e diacrônicos atinentes ao processo de estabelecimento desses contrastes. Lança enfoque sobre a gradiência e a natureza dinâmica de tal processo.

A autora realiza um estudo longitudinal com duas crianças, tendo como aporte teórico a Fonologia Acústico-Articulatória, de Albano (2001). Em uma dessas crianças, analisa o estabelecimento de contraste entre as fricativas e plosivas coronais surdas; na outra, analisa o estabelecimento de contraste entre as fricativas coronais surdas alveolar e palatal. Cada criança foi gravada separadamente em sala acusticamente tratada. O corpus foi composto por palavras dissílabas paroxítonas, nas quais os sons obstruentes coronais apareciam em posição inicial de palavra seguidos, preferencialmente, pelas vogais [a], [i] e [u]. As palavras foram inseridas em frase-veículo (fala – *palavra* – de novo). Nas análises, a autora utiliza variados parâmetros acústicos. Daí foram achados contrastes encobertos, produções gradientes, não-linearidade, hesitações e coexistência de diferentes tentativas de marcar uma distinção fônica. Por fim, a autora propõe no quadro da Fonologia Acústico-Articulatória, a existência de deslizamentos temporais e/ou variações da magnitude de determinados gestos articulatórios nas diferentes tentativas das crianças em marcar contraste fônico em aquisição.

Nessa perspectiva, ainda, encontra-se um trabalho de Rodrigues (2007), porém com vistas à aquisição dos róticos. Nele, a autora investiga o processo de aquisição dos róticos em posições de onset simples e coda por crianças com queixa fonoaudiológica relacionada ao sistema fônico. Adota como aporte teórico a Fonologia Acústico-Articulatória. A pesquisa é longitudinal e nela se coletaram dados da fala de dois sujeitos do sexo masculino com dificuldades de pronúncia de sons róticos, por um período de cinco meses. Rodrigues enfoca a maneira particular como cada um dos dois sujeitos utiliza diferentes estratégias de reparo no processo de aquisição desses sons. Por meio da análise acústica dos dados, observa fenômenos tanto categóricos como gradientes relacionados a tais estratégias, com o intuito de descrever contrastes fônicos já percebidos auditivamente (categóricos e/ou gradientes), bem como possíveis contrastes encobertos, ainda não perceptíveis de oitiva. Os parâmetros acústicos utilizados foram: duração, trajetória formântica dos três primeiros formantes, medidas espectrais (no caso do *tap*, especialmente), e frequência fundamental. Em relação a tais medidas, destaca que a trajetória formântica mostrou ser muito relevante para o estudo dos contrastes encobertos. Na sequência, observa características do detalhe fonético na aquisição dos sons da classe dos róticos pelos dois sujeitos de sua pesquisa. Os achados confirmam a existência de contrastes fônicos encobertos durante o processo de aquisição do

sistema fônico, ou seja, contrastes já em andamento na fala dos sujeitos que ainda não são perceptíveis ao ouvinte.

Berti (2010), por seu turno, analisa acusticamente as substituições envolvendo o contraste entre /t/ e /k/ na fala de crianças em aquisição típica e desviante do mencionado contraste. Faz isso com o objetivo de identificar e quantificar a existência de contrastes encobertos. Elabora um experimento de produção de fala no qual se trabalha com a repetição de palavras, que combinavam /t/ e /k/, com /a/ e /u/ na posição acentuada. Seus sujeitos (N=9) se distribuem em três grupos: crianças em processo de aquisição do contraste investigado (G1); crianças com transtorno fonológico (G2); e crianças com produções típicas (G3). Analisa por meio do software *Praat* os seguintes parâmetros acústicos: características espectrais do *burst*; transição CV e características temporais. Encontra nos resultados, tanto nas produções do G2 como nas dos G1 a presença de contrastes encobertos nos erros de substituição das oclusivas investigadas. Conclui que muitas das substituições presentes da produção de fala de crianças em processo de aquisição típico e desviante tratam-se na verdade de contrastes encobertos.

Em outro estudo, porém ainda nessa mesma direção, Berti (2011) busca: i) descrever o conjunto de pistas fonético-acústicas necessárias para o estabelecimento do contraste entre /t/ e /k/ no PB; ii) verificar como as crianças com desvio fonológico (DF) utilizam essas pistas, em termos de produção, para estabelecer o contraste entre /t/ e /k/ e (iii) investigar o desempenho perceptual das crianças com DF.

Diferentemente do estudo anterior, o de Berti (2010), no qual se tinham três grupos de crianças, neste, isto é, em Berti (2011), participaram um adulto típico do PB do sexo feminino e três crianças com DF, de ambos os sexos, com idade média de cinco anos e quatro meses. O estímulo utilizado no experimento de produção foi constituído de palavras dissílabas paroxítonas que combinavam as oclusivas /t/ e /k/ com as vogais /a/ e /u/ na posição acentuada (taco, caco, tuba e cuba) pareadas com uma gravura colorida correspondente. O procedimento consistiu em solicitar aos sujeitos cinco repetições de cada uma das palavras no interior de uma frase veículo.

A análise dos dados se realizou no software *Praat*. Os parâmetros acústicos adotados foram: pico espectral; centróide (primeiro momento espectral); variância (segundo momento espectral) e, em relação às características espectrais do *burst*, elencaram-se o parâmetro assimetria (terceiro momento espectral); a curtose (quarto momento espectral). Além desses, também se consideraram na análise os parâmetros transição formântica CV (frequência de Locus de F2) e as características temporais (medida de duração absoluta e relativa da clusura

e medida da duração absoluta e relativa do burst). O experimento de percepção auditiva, por seu turno, envolveu uma tarefa de identificação, o qual foi conduzido com as mesmas crianças do experimento de produção, usando o *software* PERCEVAL. Nos resultados, mostra que apesar de o adulto falante típico utilizar todos os parâmetros acústicos para distinguir o contraste fônico investigado, a magnitude da significância dos oito parâmetros adotados na análise não foi equiparável, indiciando a presença de parâmetros acústicos mais ou menos robustos.

Daí a autora infere que se constatam a presença de pistas acústicas primárias e secundárias para marcar o contraste fônico entre /t/ e /k/ no PB. Outro resultado diz respeito à maneira com que as crianças com DF manipulam as pistas acústicas para estabelecer o contraste entre /t/ e /k/. A partir da análise acústica dos erros de produção de fala julgados como substituições fônicas, a autora observou que 60% dessas substituições constituem-se, na verdade, de contrastes encobertos.

Observou ainda que as crianças com DF iniciam a distinção /t/ e /k/ de duas maneiras preferencialmente: (1) pelo uso de pistas acústicas secundárias e/ou (2) pelo uso de pistas acústicas primárias, mas com magnitude insuficiente.

No entanto, como dissemos em outro tópico, a população com desvio fonológico é, segundo Gierut (1998) e Wertzener (2001), muito heterogênea em relação ao modo de manifestação do problema e ao grau de severidade. Por exemplo, podem-se encontrar crianças que apresentam dificuldades na produção de determinados sons consonantais ou vocálicos e não em outros. Há aquelas que realizam processos fonológicos frente a contextos lingüísticos e situacionais impróprios. Entre esses processos, destacam-se desvozeamentos de consoantes sonoras, tais como a oclusiva bilabial sonora /b/, a fricativa labiodental sonora /v/ e outras; há também crianças que palatalizam a fricativa alveolar surda /s/ e que, em linhas gerais, também mudam pontos de articulação de diversas formas. Quanto ao grau de severidade, este pode variar de criança para criança, em uma escala que vai do grau leve ao severo, como já discutido antes.

Assim, ante a esse aspecto heterogêneo da população com desvio fonológico, não se pode esperar que os resultados de Berti (2011) se estendam a todas as crianças com esse desvio. Tanto é que essa autora sugere, nas considerações finais de seu trabalho, que futuros trabalhos busquem dar continuidade à investigação empreendida por ela, porém com um número maior de crianças com desvio fonológico, para corroborar ou não a tendência dessas crianças em: i) se ancorarem em características acústico-auditivas não-padrão e ii) em privilegiarem, quer na produção, quer na percepção, pistas fonéticas secundárias.

Miranda (2007), por seu turno, investigou a aquisição de encontros consonantais tautossilábicos (na estrutura CCV) por crianças de Belo Horizonte. Na variedade linguística desta cidade, “[...] observa-se entre a população adulta a alternância entre sílabas CCV e CV. Por exemplo, li[vr]o alterna com li[v]o e [pr]ocura com [p]ocura.” (Cristófaros Silva, 2011).

Tendo isso em vista, Miranda se propôs a analisar como é que convivendo com tais alternâncias as crianças desta cidade adquiriam a sílaba CCV. Também buscou avaliar a relevância do detalhe fonético nesse processo de aquisição. Sua hipótese era a de que a vogal da sílaba CCV pronunciada como CV teria duração maior do que a de uma sílaba originariamente CV. Para testar sua hipótese, a autora analisou dados de 16 crianças de 3:0 a 5:11 anos, sendo testadas catorze do sexo masculino e duas do sexo feminino. Para analisar os dados do grupo experimental constitui-se um grupo controle com crianças de faixa etária análoga que já produziam sistematicamente a sílaba CCV. Testaram-se 20 palavras nas quais se cotejavam as estruturas CCV e CV: broa/boa, prato/pato, prego/pego, entre outras. Na análise acústica os parâmetros utilizados foram a duração da vogal nas estruturas CCV e CV, quer na produção das crianças que não haviam adquirido o encontro consonantal, quer nas que já o haviam adquirido. Os resultados indicaram que as crianças que ainda não produziam a estrutura CCV buscavam estratégias articulatórias refinadas para expressar a ocorrência de encontros consonantais em estruturas CCV.

A autora mostra que as crianças alongam a vogal para expressar esse alvo. Destaca-se aí a relevância do detalhe fonético na aquisição do encontro consonantal tautossilábico. Segundo Cristófaros Silva (2011), os resultados de Miranda (2007) se compatibilizam com modelos multirrepresentacionais, isto é, com a Teoria de Exemplares e a Fonologia de Uso, e indicam que informações redundantes que são expressas através de detalhe fonético fino são relevantes na organização do conhecimento de sonoridade linguística pela criança.

Ainda nessa perspectiva mesma, outro estudo, no caso, o de Cristófaros Silva e Miranda (2011), traz uma contribuição para o debate sobre a natureza multirrepresentacional da linguagem ao analisar a aquisição de encontros consonantais tautossilábicos no Português falado em Belo Horizonte. Fazem isso com base na Teoria de Exemplares. Desse modo, as autoras (2011) buscam discutir a natureza das representações fonológicas ao avaliar o contraste encoberto como estratégia de construção de linguagem. Investigam a hipótese do alongamento compensatório no contraste encoberto entre as estruturas silábicas CCV e CV. Analisam dois grupos distintos de participantes: um controle composto por 10 crianças que já adquiriram o encontro consonantal tautossilábico; e um experimental composto por 10 crianças que ainda não adquiriram tal encontro.

Para o grupo experimental foram selecionadas crianças só do sexo masculino, visto que as crianças do sexo feminino que participaram da pesquisa tinham, em algum grau, apresentado sílabas CCV em sua produção. As crianças do grupo experimental encontravam-se na faixa etária entre 3:3 e 4:6 anos. O grupo controle foi constituído de cinco meninos e cinco meninas, com idade entre 4:1 e 5:11, sendo que todas já produziam sistematicamente as sílabas CCV. Foram testadas 20 palavras, que além de terem significado diferente tinham como diferença as estruturas CCV e CV, como em broa/boa, prato/pato etc. Vale frisar que foram as mesmas palavras utilizadas no experimento de Miranda (2007).

Para a coleta dos dados utilizaram-se diferentes estratégias: nomeação de figuras, reconto de história, jogo de memória e, em alguns poucos casos, repetição. Os dados foram gravados com gravador digital e microfone unidirecional em uma creche de atendimento diurno. Posteriormente, foram analisados no *software* Praat. Vale frisar ainda que, para testar a hipótese de alongamento compensatório, foram extraídas as medidas de duração da vogal das estruturas silábicas CCV e CV em cada um dos itens lexicais analisados.

Os resultados indicam que: i) as crianças que ainda não adquiriram a estrutura CCV utilizam estratégias de alongamento compensatória da vogal para diferenciar acusticamente estruturas CCV de CV e ii) as crianças que já adquiriram a sílaba CCV apresentavam duração análoga para vogais em sílaba CCV ou CV. Isso indica que há sistematicidade quanto à duração das vogais tônicas, quer em sílabas CCV, quer em CV, para as crianças que já adquiriram a sílaba CCV.

“Os resultados indicam a relevância da análise acústica nos estudos fonológicos e a pertinência do detalhe fonético às representações do componente sonora da linguagem.” (CRISTÓFARO SILVA; MIRANDA, 2011, p. 14). Depreendem-se ainda desses resultados, segundo as autoras, evidências para os modelos multirrepresentacionais que sugerem que a experiência e o uso são cruciais na organização e no gerenciamento do conhecimento linguístico. Uma pergunta que emerge daí, de acordo com as autoras, é a de como se daria, portanto, a apropriação do *tap* pela criança que ainda não produz a sílaba CCV.

Vale lembrar que em nossa pesquisa buscamos trazer contribuições nessa direção, uma vez que investigamos a produção do *tap* em onset complexo em tal estrutura silábica, quer em crianças com desvio fonológico, quer em crianças sem o desvio.

Ainda no sentido da proposta de Cristófaró Silva e Miranda (2011), porém com base na Fonologia Gestual e em técnicas acústica e ultrassonográfica, encontra-se o estudo de Vassoler (2016).

Neste, considerando os padrões silábicos CCV e CV nas produções de crianças em processos de aquisição típica e atípica, a autora (2016) busca descrever e investigar o padrão de coordenação gestual imbricado nessas produções. Também busca identificar se as chamadas reduções dos encontros consonantais em crianças com desenvolvimento atípico poderiam ser consideradas como contrastes encobertos.

Suas hipóteses foram as três seguintes: i) a produção de CV e CCV das crianças em aquisição típica apresentariam diferentes medidas ultrassonográficas (distância das raízes e da área do contorno da língua e de outros gestos articulatórios no trato vocal) e acústica (duração); ii) as medidas ultrassonográficas e acústica poderiam diferenciar a condição clínica das crianças; iii) as crianças com produções atípicas poderiam apresentar diferenças nas medidas ultrassonográficas e acústica na comparação da produção das sílabas CCV e CV, ainda que CCV tenha sido julgada auditivamente como CV.

Nos resultados, corrobora-se a primeira hipótese. Já a segunda e a terceira são confirmadas apenas parcialmente. Os resultados articulatórios sugerem a presença de elevação da ponta da língua para produzir o *tap* com magnitude reduzida e os resultados acústicos sugerem a não sobreposição dos gestos na produção de CCV. Vale destacar, ainda, que, segundo a autora:

Nas chamadas simplificações/reduções de CCV, à luz da Fonologia Gestual, espera-se que nos contrastes encobertos haja mudanças gestuais devido à sobreposição dos gestos adjacentes e/ou à redução da magnitude dos gestos. (VASSOLER, 2016, p. 135).

Na sequência, Vassoler afirma, ainda, que seu trabalho confirmou parcialmente a premissa da Fonologia Gestual, já que se confirmou a presença de contrastes encobertos por conta de mudanças da magnitude gestual nas produções julgadas como simplificações de CCV. Em suma, constata-se que “crianças atípicas” produzem as sílabas CCV e CV de maneiras diferente da produção dos grupos de “crianças típicas”.

A autora, entretanto, não aborda determinadas especificidades sobre a estrutura CCV, isto é, não investiga especificamente produções acústicas do *tap* em onset complexo de crianças com desvio fonológico. Além disso, não busca investigar nem descrever padrões acústicos do *tap* em onset complexo, mas sim padrões de coordenação gestual imbricados na produção das estruturas CCV e CV à luz da Fonologia Acústico-Articulatória.

Já em nossa pesquisa, vale lembrar, buscaram-se a investigação e a descrição de tais padrões acústicos; especialmente, o CCV.

Como se viu, apresentou-se, aqui, um breve panorama dos estudos atinentes à aquisição do sistema fonético-fonológico do português brasileiro. Relembrando isso, em síntese, o primeiro deles foi o de Panhoca-Levy (1993), que se colocou como uma proposta pioneira sobre aquisição fonético-fonológica do português brasileiro à luz da Fonética Acústica. Na sequência, apresentaram-se os trabalhos de Berti (2006), Freitas (2007), Rodrigues (2007), Berti (2010) e Berti (2011), os quais se apoiam na Fonologia Acústico-Articulatória. Em seguida, destacaram-se também os trabalhos de Miranda (2007) e Cristófar-Silva e Miranda (2011), os quais, por seu turno, se embasam nos Modelos Multirrepresentacionais, quer dizer, na Teoria de Exemplares e na Fonologia de Uso. Por fim, foi resenhado o trabalho de Vassoler (2016), o qual realiza por meio da Fonologia Acústico-Articulatória, abordagem com combinação de técnicas acústicas e ultrassonográficas.

### 3 TEORIA FONTE-FILTRO

O pressuposto elementar da Teoria Fonte-Filtro é o de que na fala existe forte correlação entre fenômenos acústicos e articulatórios, de modo que mudanças no trato vocal podem causar mudanças nas frequências de ressonância<sup>3</sup> geradas nesse trato. Outro pressuposto é o de que as características do sistema de produção da fala (trato vocal) podem ser inferidas por meio de análise sobre a saída do sinal acústico desse sistema. É, basicamente, a partir dessas assunções que se desdobram nessa teoria recursos teóricos tais como conceitos, fórmulas e funções matemáticas, as quais permitem calcular, por exemplo, frequências de ressonância produzidas no reportado trato.

A Teoria Fonte-Filtro é importante para se entender as relações acústico-articulatórias, (KENT; READ, 2015), bem como para fornecer recursos teóricos necessários a uma análise acústica da fala. Nesse sentido, propõe que o sistema de produção da fala se divide em dois componentes distintos: uma fonte, que gera energia para a produção de um som da fala, e um filtro, que molda o som gerado pela fonte.

A fonte pode se situar em qualquer ponto entre a glote e os lábios, podendo ser tanto periódica, como a gerada pela vibração das pregas vocais, quanto aperiódica, como a produzida pela passagem do ar em região estreita no trato oral, por exemplo: na produção da consoante fricativa surda [s] (BARBOSA; MADUREIRA, 2015). Mas duas fontes distintas podem se combinar para a produção de um mesmo som, como ocorre, por exemplo, no caso da consoante fricativa vozeada [z], na qual se encontram periodicidade e aperiodicidade.

O filtro, por seu turno, é o próprio trato vocal e age como um corpo ressoador sobre o som originado pela fonte, potencializando faixas de frequência desse som de acordo com áreas no plano transversal ao longo do eixo sagital do trato, isto é, o eixo que vai da glote aos lábios.

Movimenta-se dinamicamente assumindo formas distintas que vão, como já dito antes, moldar o som produzido pela fonte. Para que este som se realize como uma vogal [a], por exemplo, tem-se que sua fonte geradora é a vibração laríngea enquanto o filtro o molda assumindo um formato no qual a língua abaixa e os lábios ficam abertos, sem qualquer obstrução para a passagem do ar. Já na produção de uma consoante, como a fricativa

---

<sup>3</sup> Frequência é o número de repetições de uma vibração num dado intervalo de tempo. A ressonância, por sua vez, pode ser definida como um fenômeno acústico no qual a vibração de um corpo (objeto), em sua frequência natural, age em resposta a frequências similares aplicadas por uma força, transiente (impulso) ou contínua. Esta força é uma fonte sonora externa ao mencionado corpo. É a partir desse fenômeno que se realiza a frequência de ressonância.



labiodental surda [f], tem-se uma constrição estreita entre os lábios superiores e os dentes inferiores, a qual se caracteriza como fonte de energia do som dessa fricativa, cujo filtro reside na porção anterior à constrição.

O filtro também pode ser exemplificado como um tubo reto e uniforme, com uma de suas extremidades parcialmente fechada por uma membrana elástica, que simule as pregas vocais, e a outra aberta, correspondente à abertura da boca (KENT; READ, 2015; MARUSSO, 2005). Supondo que essa membrana vibre como uma fonte de energia acústica, tal como as pregas vocais, essa energia viaja ao longo do tubo, que é um ressonador.

Esse tipo de tubo, o ressonador, possui um número infinito de ressonâncias situadas em diferentes frequências dadas por uma relação que pode ser denominada, segundo Kent e Read (2015, p. 39), “múltiplo ímpar do quarto de comprimento de onda:”

$$F_n = (2n - 1) c/4l$$

Nesta fórmula, **n** é um número inteiro; **c**, a velocidade do som (aproximadamente 35000 cm/s) e **l** é o comprimento do tubo (esse **l** vem da palavra inglesa *length*). Tal fórmula significa que um tubo, como o que se descreveu há pouco, ressoará com amplitude máxima um som cujo comprimento de onda sonora for quatro vezes maior que o seu. As ressonâncias nesse tubo ocorrem em múltiplos e é por isso que a expressão  $(2n - 1)$  é usada para gerar o conjunto de números ímpares, os quais, ao serem colocados na fórmula que contém essa expressão, gerariam uma sequência como  $c/4l$ ,  $3c/4l$ ,  $5c/4l$ ,  $7c/4l$  e assim sucessivamente, rumo ao infinito positivo.

Agora vejamos um exemplo de cálculo de frequência de ressonância com a fórmula  $F_n = (2n - 1) c/4l$ . Assumindo o valor de **l** como a medida do comprimento de um tubo, ou melhor, de um trato vocal de um homem, que é em média 17,5 cm, pode-se dizer que a primeira ressonância terá a seguinte frequência:

$$F_n = (2n - 1) c/4l$$

$$F_1 = (2 \cdot 1 - 1) c/4(17,5\text{cm})$$

$$F_1 = 1 \cdot c/4(17,5\text{cm})$$

$$F_1 = 35000\text{cm/s}/70\text{cm}$$

$$F_1 = 500\text{Hz}$$

Como se pode notar na aplicação dessa fórmula para F1, tem-se que o n é igual a 1. Portanto, obtém-se inicialmente  $c/4l$ , e o resultado final será 500 Hz. Logo, F2, que é  $3(c/4l)$  será igual a  $3(500 \text{ Hz})$ , ou seja, 1500 Hz. Esse mesmo raciocínio é aplicável em relação à sequência de números ímpares combinada com a sequência das demais frequências F3, F4, F5 etc. Assim, veja:  $F3 = 5(500 \text{ Hz}) = 2500 \text{ Hz}$ ;  $F4 = 7(500\text{Hz}) = 3500 \text{ Hz}$ ;  $F5 = 9(500 \text{ Hz}) = 4500 \text{ Hz}$  etc.

Desse modo, considerando-se as cinco primeiras ressonâncias de um tubo com 17,5cm de comprimento, vimos que as frequências que se obterão são as seguintes: 500 Hz, 1500 Hz, 2500 Hz, 3500 Hz e 4500 Hz, as quais apresentam entre si intervalos de 1000 Hz. Nesse sentido, ainda, ressonâncias mais altas ainda podem ser obtidas pela continuação desses cálculos para diferentes soluções da expressão  $(2n - 1)$ , continuando a reportada sequência, com intervalos de 1000 Hz, infinitamente.

Relacionando isso com a produção da fala humana, interessa notar os seguintes aspectos: i) o trato vocal médio masculino tem, como já aludido, um comprimento de cerca de 17,5 cm (da glote até os lábios) – mas vale frisar, aqui, que, sobre tal comprimento, podem ser encontradas outras estimativas na literatura, como reporta Stevens (1999, p.25), baseado em Goldstein (1980): comprimento do trato vocal masculino: 16,9 cm, sendo o comprimento da faringe 8,9 cm e o da cavidade oral 8,1 cm; e ii) o trato vocal tem aproximadamente as mesmas frequências de ressonância de um tubo reto de mesmo comprimento e seção transversal.

Também interessa notar que as frequências assumem valores numa proporção inversa ao comprimento do tubo. Assim, se o tubo apresenta comprimento maior que 17,5 cm, o dobro, por exemplo, 35 cm, as frequências de ressonância assumirão valores mais baixos, ou seja, 250 Hz, 750 Hz, 1250 Hz, 1750 Hz e 2250 Hz, quer dizer, assumirão valores, como se pode perceber, que serão a metade dos respectivos valores das frequências obtidos quando o comprimento é de 17,5 cm.

Por outro lado, num tubo de comprimento menor, como o trato vocal feminino, que é em média cerca de 10 a 15% menor que o masculino, medindo algo em torno de 14,1 cm, segundo Goldstein (1980 *apud* STEVENS, 1999, p. 25), os valores de suas frequências de ressonância serão maiores que as frequências do trato vocal masculino. Isso acontece porque a relação entre o comprimento do tubo e os valores das frequências de ressonância se estabelecem numa proporção relativamente inversa. Quanto maior o comprimento do trato vocal, mais baixas serão as frequências de ressonância e quanto menor o comprimento do trato, maiores serão tais frequências. Nas crianças, cujo trato vocal é ainda menor que o de

uma mulher, por exemplo, os valores de suas frequências de ressonância serão bem maiores do que as frequências de ressonância produzidas por uma mulher.

## 4 SONS CONSONANTAIS À LUZ DA FONÉTICA ACÚSTICA

Introduzindo uma parte de seu livro que abrange as consoantes à luz da Fonética Acústica, Fry (1976, p. 123) afirma que a análise acústica dos sons consonantais apresenta problemas mais complexos do que a análise acústica das vogais. Muitos desses sons são aperiódicos e, portanto, não têm estrutura harmônica. Geralmente, a duração deles é bem menor que a das vogais. Por exemplo, os sons plosivos, em particular, envolvem muitos *bursts* curtos; e frequentemente resultam de ruídos gerados em diferentes pontos ao longo do trato vocal, a partir dos quais os efeitos de filtragem nesse trato se caracterizam de forma muito complexa.

Tendo isso em vista, e visando aclarar determinados detalhes fonéticos atinentes aos sons consonantais à luz da Fonética Acústica, para posteriores empreendimentos analíticos, abordam-se, neste capítulo, tais sons plosivos, isto é, as consoantes oclusivas e, na sequência, as consoantes fricativas e os róticos. Vejamos.

### 4.1 Consoantes oclusivas

Stevens (1999) desenvolve as contribuições de Fry em sua obra sobre Fonética Acústica, trazendo grandes avanços, graças também ao desenvolvimento tecnológico e científico que se deu no período que separa sua obra da de Fry. Mas aqui, cumpre dizer, não caberia resenhar tais autores, pois, assim, fugiríamos de nossos propósitos nesta tese.

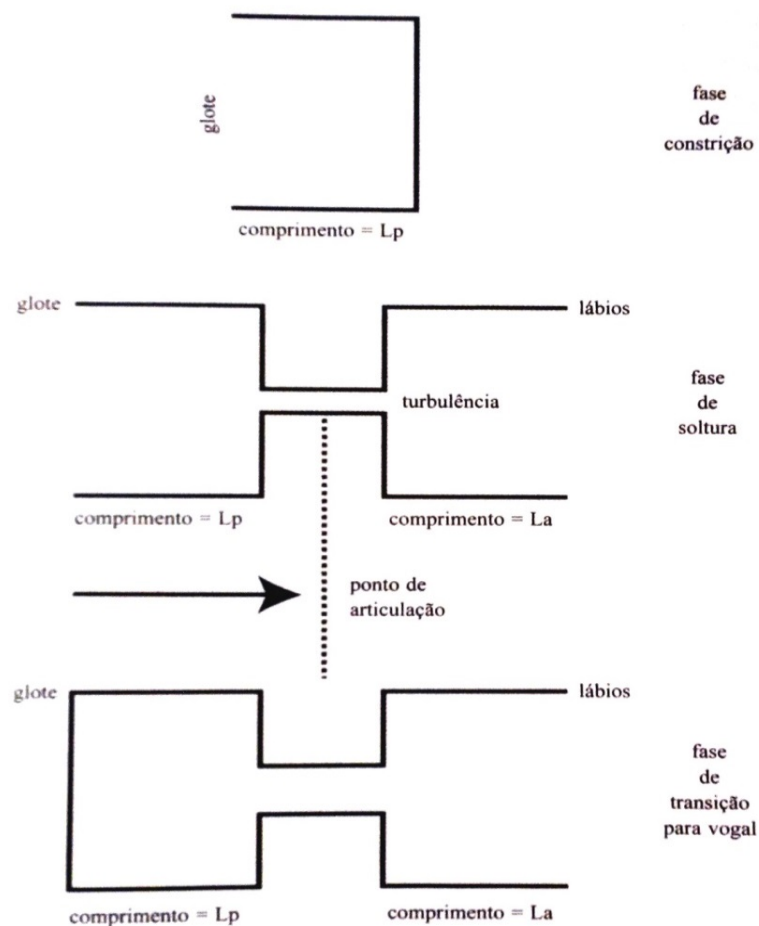
No entanto, convém reportar, em conformidade com Stevens (1999, p. 333), que as consoantes oclusivas apresentam três classes de intervalos de tempos na onda de som, as quais permitem identificar suas características fonéticas: i) intervalos correspondentes ao momento da constrição no trato vocal; ii) intervalos correspondentes ao momento da soltura da constrição; e iii) intervalo de tempo em que não há mais constrição no trato.

Barbosa e Madureira (2015, p. 115), por sua vez, perspectivam esses intervalos de tempo como fases, sendo o primeiro deles correspondente à fase de constrição; o segundo, à fase de soltura; e o terceiro, à fase de transição (para vogal).

Durante a fase de oclusão com constrição total no ponto de articulação, não há som algum numa oclusiva não vozeada e há som laríngeo filtrado pelas paredes do pescoço na oclusiva vozeada (este som pode ser capturado pelo microfone). Durante a fase de soltura, a configuração no trato satisfaz à condição de geração de turbulência porque o fluxo de ar passa com grande velocidade pela constrição que se abre, por conta da alta pressão atrás do ponto de

oclusão. Nesse momento, produz-se um ruído fricativo transiente com as mesmas características do ruído de uma fricativa homorgânica. A terceira fase desempenha um papel na produção de movimento dos formantes à margem da vogal seguinte (ou precedente, se a oclusiva sucede a vogal). Esse movimento formântico à margem da vogal permite a identificação da oclusiva. Vale ressaltar, ainda, que as duas primeiras fases são mais facilmente compreensíveis pelo exposto deixando marcas acústicas correspondentes a silêncio (oclusiva não vozeada) ou barra de vozeamento (oclusiva vozeada) seguido de um breve ruído de fricção. A seguir, vejamos, na figura 1, um modelo de primeira ordem de tubo acústico da produção de oclusiva não vozeada nas reportadas três fases.

Figura 1 – Modelo de primeira ordem de tubo acústico de produção de oclusiva não vozeada em três fases



Fonte: Barbosa e Madureira (2015, p.115).

Legenda: Esse modelo de tubo acústico de produção de oclusiva não vozeada se apresenta em três fases: oclusão, soltura e transição para a vogal seguinte. Mostram-se os comprimentos das cavidades anterior e posterior com relação à constrição na segunda e terceira fases.

Após a fase de soltura da oclusiva, a mandíbula ainda está alta, o que produz um F1 baixo de um ressoador de Helmholtz<sup>4</sup>, no início da vogal seguinte (na transição C – V) ou ao seu final na transição (V - C). O que ocorre com os movimentos de F2 e F3 depende da posição de restrição.

Nas oclusivas bilabiais ([p], [b]) pré-vocálicas, os lábios ficam arredondados no início da vogal. Essa constrição labial se dá em ponto de máxima taxa de mudança de pressão. “De acordo com a Teoria da Perturbação, essa constrição abaixa frequências de formantes.” (BARBOSA; MADUREIRA, 2015, p. 116). Desse modo, todos os formantes à margem da vogal proviriam de um valor mais baixo, caracterizando um movimento de subida de formantes.

Nas oclusivas alveolodentais ([t], [d]) pré-vocálicas, a constrição próxima a um ponto de taxa de mudança de pressão nula de F2 eleva a frequência de F2, mas não ao ponto de ultrapassar o F2 das vogais anteriores. A constrição é, ainda, suficiente para afetar também F3 com mesmo efeito elevador de formante, mas inferior ao F3 de vogais anteriores. Por esse motivo, o movimento numa transição de articulação alveolar consonântica para uma vogal depende do ponto de articulação da vogal: o movimento é de subida de F2 e F3 se a vogal é anterior, de descida se a vogal é posterior e de movimento pouco inclinado se a vogal é central.

Já nas oclusivas velares ([k], [g]) pré-vocálicas, ainda de acordo com Barbosa e Madureira (2015), o movimento de F2 e F3 se explica da seguinte forma: a constrição na região velar divide o trato oral em cavidades anterior e posterior na proporção respectiva de 1/3 e 2/3 ( $L_p = 2.L_a$ ). Assim, após a soltura e produção da oclusiva velar, a transição para a vogal produz valores de F2 e F3 que provêm das cavidades anterior e posterior. No trato oral, dois tubos produzem os mesmos valores de formantes que representam F2 e F3, logo que eles começam a diferir com a progressão da realização da vogal, uma vez que F1 é dado pelo ressoador de Helmholtz.

---

<sup>4</sup> Esta ressonância foi descoberta pelo físico alemão Helmholtz, segundo o qual, para dois cilindros unidos com áreas de seção transversal bem distintas, como uma garrafa com bojo e gargalo, o ar desse último age como pistão e cria uma ressonância própria, cuja frequência depende do volume do bojo e do comprimento e área do gargalo. As ressonâncias de Helmholtz na fala apresentam baixa frequência e ocorrem geralmente em casos especiais, sendo mais recorrentes em baixas ressonâncias de F1 de vogais altas, tais como /i/ e /u/. Para mais, ver Barbosa e Madureira (2015).

## 4.2 VOT das oclusivas

O intervalo de tempo da soltura ao início de vozeamento se constitui como um contraste de duração fundamental para as consoantes oclusivas de várias línguas do mundo. Essa duração para o vozeamento é conhecida como *Voice Onset Time* (VOT).

A criação dessa categoria de análise fonética se consolidou a partir de um trabalho de seminal Lisker e Abramson (1964). Nesse trabalho, esses autores se propuseram a demonstrar como o VOT poderia servir para distinguir consoantes oclusivas em várias línguas. Ao todo, investigaram 11, entre as quais o Inglês Americano, o Cantonês, o Holandês, o Húngaro, o Espanhol de Porto Rico, o Coreano e o Armênio, entre outras.

“Os procedimentos gerais envolveram a análise espectrográfica de dados gravados em sala acusticamente tratada.” (LISKER; ABRAMSON, 1964, p. 338). Cada um dos informantes, 17 ao todo, produziram um conjunto de palavras escolhidas para uma amostra de oclusivas iniciais pré-vocálicas encontradas em suas respectivas línguas. Para cada palavra, o informante era solicitado a produzir duas frases apresentando a palavra-alvo em posições inicial e não-inicial. Lisker e Abramson (1964) solicitavam, ainda, ao informante que enunciasse as frases com fluência e com a naturalidade de uma conversação normal. Desse modo, cada informante tinha de gravar por duas vezes cada palavra e cada frase.

Daí os dados foram analisados em espectrogramas de banda larga e, por meio destes, o VOT foi medido pela marcação do intervalo entre a soltura da oclusiva e o início da vibração glotal, isto é, o vozeamento. Nos resultados, os autores provaram que o VOT é uma medida eficiente para diferenciar pontos de articulação e modos de fonação das oclusivas.

Graças a esses autores, hoje, e como já aludido no início deste tópico, o VOT se define como a diferença de tempo entre o instante no qual as pregas vocais começam a vibrar e o instante em que se dá a soltura dos articuladores.

Considerando Lisker e Abramsom, além de outros autores, Barbosa e Madureira (2015), por seu turno, afirmam que o VOT é positivo quando as pregas vocais começam a vibrar após a fase de soltura, que é o caso das oclusivas surdas em Português. Nestas consoantes, o VOT corresponde a duração de sua segunda fase de produção. Em contrapartida, diz-se que o VOT é negativo quando as pregas vocais começam a vibrar antes da fase de soltura. É o caso das oclusivas sonoras do Português e de línguas como o Castelhana, o Italiano e o Francês.

Comparando-se os pontos de articulação das oclusivas surdas entre si, vê-se na literatura que as consoantes mais anteriores apresentam valores de VOTs menores que os

valores das consoantes posteriores. Quanto a isso, Cho e Ladefodeg (1999) asseveram que o efeito da região de articulação sobre os valores de VOT é consistente entre as línguas do mundo. E Kent e Read (2015, p.249) dizem que “A regra geral é que bilabiais possuem os VOTs mais curtos, incluindo pré-vozeamento frequente; alveolares possuem VOTs intermediários; velares possuem os mais longos VOTs”.

Aliás, cumpre dizer que poucas línguas no mundo invertem essa ordem. Para se ter uma idéia melhor disso, ressalte-se que Cho e Ladefoged (1999, p. 222) reportam apenas três exceções que quebram tal ordem, a saber: Hupa, Navajo e Tlingit. Nestas línguas, encontram-se médias de VOTs para as oclusivas velares menores que as médias para as alveolares.

Tais autores (1999, p. 209) afirmam, ainda, que muitos foneticistas, entre os quais Hardcastle e Maddieson, sugerem que um dos fatores que contribuem para as diferenças dos VOTs é o tamanho relativo da cavidade supraglotal por trás do ponto de constrição.

Há dois modos de se considerar isso: primeiro, a cavidade por trás da oclusiva velar tem um volume menor que a das oclusivas alveolar e bilabial; segundo, a cavidade à frente da oclusiva velar tem maior volume que a cavidade à frente das oclusivas alveolar e bilabial. Isto implica diferenças na pressão do ar que incide nas cordas vocais, de modo que o que determina o menor ou maior tempo para o início de vibração das pregas vocais após a fase de soltura da oclusiva é a pressão supraglotal.

Assim, quanto maior essa for, mais tempo vai demorar para a pressão supraglotal ultrapassá-la e iniciar a vibração das pregas. “Como a pressão é tanto menor quanto for o volume acima da glote, deduz-se que as oclusivas velares devem ter maior valor de VOT e que as bilabiais devem ter o menor.” (BARBOSA; MADUREIRA, 2015, p. 342).

Interessa ressaltar, ainda, que essas diferenças de VOT podem também estar associadas, ou condicionadas, a outros fatores, entre os quais, vale ressaltar, aqui, o sexo. Nesse sentido, Barbosa e Madureira (2015) sugerem que a razão para diferenças de VOT condicionadas ao sexo se deve ao volume dos tratos vocais, menores nas mulheres, o que produz pressões maiores que atrasam o início de vozeamento. Daí, portanto, deve-se esperar valores de VOT mais altos para as mulheres.

Outro fator condicionante pode ser a idade, uma vez que as diferenças anatômicas dos tratos vocais condicionadas ao sexo começam a se desenvolver muito cedo, ainda na tenra infância. A este respeito, Kent e Read (2015, p.330) com base nos trabalhos de Bennett (1981); Busby e Plant (1995); Lee et al. (1999); Whiteside e Hodgson (2000), afirmam que o dimorfismo sexual do trato vocal emerge em torno da idade de, no mínimo, 7 ou 8 anos.



#### 4.2.1 Estudos sobre efeitos do sexo no VOT

Em nossa revisão, não foram encontrados, no Português Brasileiro, trabalhos sobre os efeitos do sexo no VOT na fala de crianças. Em razão disso, apresentam-se aqui, inicialmente, trabalhos que analisam tais efeitos em dados de falantes adultos, como Celeste e Teixeira (2009) e Barbosa e Madureira (2015). Mas, logo após estes, apresenta-se um trabalho com dados de crianças, porém, realizado no Inglês Britânico, a saber: Whiteside, Henry e Dobbin (2004).

Celeste e Teixeira (2009) investigaram a variação do VOT no Português Brasileiro da consoante /k/ sob efeitos de contexto vocálico CVC, sexo e idade. Seus grupos se constituíram de mulheres adultas (N=5), homens adultos (N=5), mulheres idosas (N=5) e homens idosos (N=5). Considerando a variável sexo, encontraram-se diferenças significativas no VOT de [k]. Entretanto, diferentemente do que se esperava com base nas pesquisas com falantes do Inglês, Celeste e Teixeira (2009) encontraram valores maiores para o sexo masculino, tanto no contexto em que a consoante [k] era seguida por [a], ( $p=0,0006$ ), como no em que era seguida por [e] ( $p=0,05$ ). Vale frisar que, dada a escassez de trabalhos nesse sentido no Português Brasileiro, tais autoras buscam comparações com resultados de pesquisas no Inglês, de pesquisadores como Whiteside e Irving (1998), Morris (2009) e Robbi *et al.* (2005).

Barbosa e Madureira (2015, p. 352), por seu turno, reportam um estudo com dois falantes de sexos distintos, um paulista e uma carioca, no qual o VOT das oclusivas surdas foi maior no falante paulista e menor na falante carioca. Nesses falantes, entretanto, os valores dos VOT das oclusivas não diferem significativamente entre as taxas de elocução que se observavam no estudo.

Whiteside, Henry e Dobbin (2004), por vez, realizam um estudo no Inglês Britânico, no qual investigaram padrões de diferenças no VOT das oclusivas seguidas por /a/ e /i/ em relação ao sexo dos participantes (N=46), crianças e pré-adolescentes, com idades entre 5,8 e 13,2. Os resultados mostraram que os participantes do sexo feminino produziram valores de VOT maiores do que os participantes do sexo masculino.

Nesse estudo, os valores das médias dos VOT da consoante [k] no contexto em que ela se seguia pela vogal /a/ foram os seguintes: a) no grupo de crianças com 5 anos de idade, a média foi de 96.1 para os meninos e de 106.8 para as meninas; b) no grupo de crianças com 7 anos: 74.4 para os meninos e 87. 2 para as meninas; c) no grupo de crianças com 9 anos: 70.1 para os meninos e 80.5 para as meninas; d) no grupo de crianças com 11 anos: 72.1 para os meninos e 66.6 para as meninas; e) no grupo de crianças (pré-adolescentes) com 13 anos: 71.8

para os meninos e 90.0 para as meninas. Repare-se, aí, que, com exceção do grupo de crianças na faixa etária de 11 anos, todas as médias dos VOTs da consoante [k] produzidos pelas meninas foram maiores que a dos meninos.

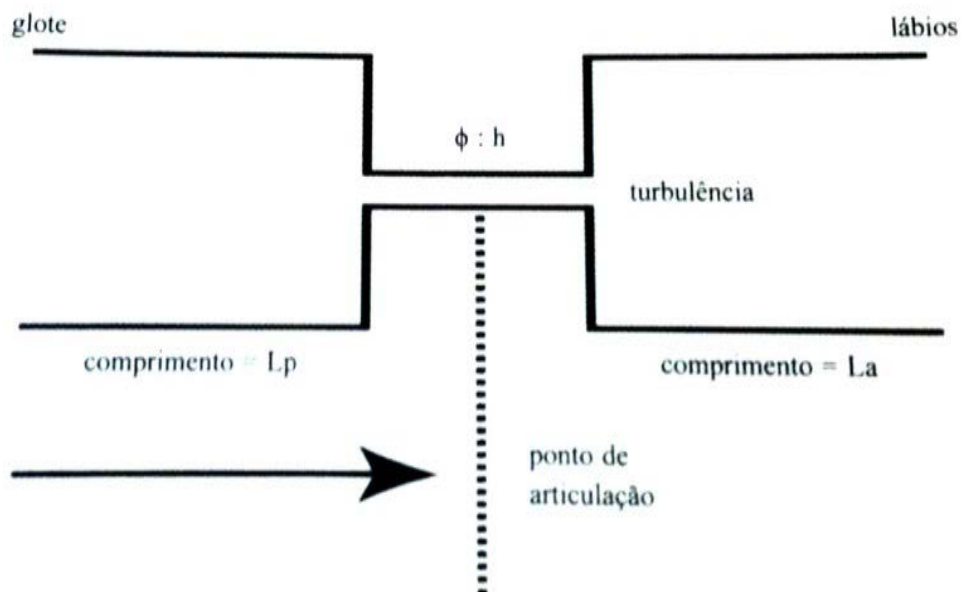
Já em relação a todas as consoantes oclusivas investigadas nesse estudo, incluindo-se a velar surda, no contexto em que se seguiam por /a/ e /i/, vale ressaltar, entretanto, que os resultados foram mais consistentes apenas no grupo de crianças mais velhas, pré-adolescentes, com 13 anos.

### 4.3 Consoantes fricativas

Neste tópico, cumpre ressaltar, parafraseamos de forma mais extensa e exclusiva os autores Barbosa e Madureira (2015), cuja abordagem é suficiente ao que nos interessa aqui.

Considerando esta ressalva, apresenta-se, logo de início, um modelo de primeira ordem de tubo acústico da produção de fricativa não vozeada indicando-se os comprimentos das cavidades posterior ( $L_p$ ) e anterior à constricção ( $L_a$ ), onde ocorre a geração e turbulência. O ponto de articulação da fricativa é definido pelo local da constricção, a qual se situa em qualquer ponto do trato oral. Para que haja fricativa, é necessário que o fluxo de ar, ao passar por uma constricção estreita, produza turbulência.

Figura 2 – Modelo de primeira ordem de tubo acústico da produção de fricativa não vozeada.



Fonte: Barbosa e Madureira (2015, p.111).

Legenda: mostram-se nesse modelo os comprimentos das cavidades posterior ( $L_p$ ) e anterior ( $L_a$ ) com relação à constricção.

A turbulência se caracteriza por uma vibração com direção aleatória das moléculas de ar gerada quando essas passam por um tubo estreito a uma determinada velocidade. Supondo que o tubo da constrição seja um cilindro de diâmetro  $h$  e a velocidade do fluxo de ar seja  $V$ , há turbulência quando o número de Reynolds ( $Re$ ), o qual se define pela equação  $Re = V.h/\nu$ , é tal que  $Re > 1200$ , em que  $\nu$  é a viscosidade cinemática do ar de cerca de  $15 \text{ cm}^2/\text{s}$ . Desse modo, em um tubo de 4 mm, por exemplo, aplicando tal equação e a condição  $Re > 1200$ , obtemos:  $V > 4,5 \text{ m/s}$ , ou seja, é preciso que o fluxo de ar ultrapasse essa velocidade para gerar a fricção de uma fricativa.

Uma vez gerada a turbulência, que é uma fonte ruidosa, o ar em vibração em diversas frequências passa pela cavidade anterior de comprimento  $L_a$  dando condições para que haja ressonância. As ondas estacionárias no tubo anterior são formadas desde que sua frequência satisfaça a fórmula  $j_{\text{ímpar}} \cdot c_{\text{som}} / (4L_a)$ , em que  $c$  é a velocidade do som e  $j_{\text{ímpar}}$  é qualquer número ímpar maior que zero. A vibração ruidosa também se propaga no tubo posterior formando ondas estacionárias nesse tubo de comprimento  $L_p$ . Entretanto, por conta do grande estreitamento na constrição essas ondas estacionárias ficam aprisionadas na parte posterior do trato (agem como antiformantes), minimizando a energia nas faixas de frequência regidas pela reportada fórmula, caso o tubo posterior esteja aberto na glote (fricativa não vozeada) e  $j_{\text{par}} \cdot c_{\text{som}} / (4L_a)$ , em que  $j_{\text{par}}$  é qualquer número par maior que zero, caso o tubo posterior esteja fechado na glote (fricativa vozeada).

Nas fricativas, devido à complementaridade entre as cavidades posterior e anterior no trato oral (quando uma é curta, a outra é longa e vice-versa), quando há um formante de frequência elevada, há também um antiformante em frequência baixa.

Aqui vale frisar que fricativas bilabiais não geram formantes: não há cavidade anterior, e o espectro do som é tal como produzido nos lábios. Isso é válido também para as fricativas labiodentais: a cavidade entre os dentes inferiores e o lábio superior é muito pequena para gerar formante audível.

A formação de ondas estacionárias não é o único fenômeno que se dá com a propagação do ruído fricativo no trato oral. A intensidade do ruído é aumentada quando o fluxo turbulento passa por um obstáculo no trato oral. Fricativas labiodentais têm seu fluxo passando pelo lábio superior, aumentando sua intensidade global. As fricativas alveolares, por sua vez, têm seu fluxo passando pelos incisivos superiores, aumentando sua intensidade global. Já o obstáculo no trato para as fricativas pós-alveolares são os incisivos inferiores, efeito acrescido por ressonâncias na cavidade que aumentam sensivelmente a intensidade desse som.

#### 4.4 Róticos

No inventário fonético-fonológico das línguas do mundo, há uma classe de sons designada róticos. O termo rótico vem do grafema grego  $\rho$ , chamado *rho* (lê-se: rô), o qual se refere aos sons de ‘r’. A maioria das classes de sons tradicionais, na teoria fonética, se define com base em propriedades articulatórias e auditivas dos sons mesmos, mas os róticos, não. (LADFODEG; MADDIESON, 1996). Estes, na verdade, se baseiam no fato de estarem associados com uma característica peculiar do sistema ortográfico derivado da tradição Greco-Romana: a letra ‘r’, ou sua contraparte grega *rho*. Assim, os róticos podem ser entendidos como uma classe fonológica cujos sons se associam com a letra ‘r’.

Em geral, os róticos se produzem como aproximante, alveolar e retroflexo, como vibrantes (*trills*) bilabial, alveolar e uvular, como *tap* alveolar, pós-alveolar e retroflexo, como fricativas de ponto posterior (velar, uvular e glotal) e mesmo como *offglides* (aproximantes à direita da vogal) rotacizantes, no caso do “R caipira”, comum no interior de São Paulo. (BARBOSA; MADUREIRA, 2015). Tais produções podem ser vozeadas e não vozeadas, dependendo de contexto e língua.

Os membros mais prototípicos da classe dos róticos são os *trills* que se realizam com a ponta ou a lâmina da língua. Esses membros partilham relações fonológicas com um conjunto heterogêneo de sons tais como *taps*, fricativos, aproximantes, além de se realizarem em pontos bilabiais, alveolares e uvulares.

Como já se pode notar, os róticos apresentam, foneticamente, grande variabilidade articulatória. Isso dificulta que se estabeleçam possíveis definições a partir de suas propriedades ou modos de articulação. E, como já dito, tais propriedades não definem essa classe (LADFODEG; MADDIESON, 1996, p. 2015). Contudo, fonologicamente, os róticos tendem a se comportar de maneira semelhante, apresentando, principalmente, características distribucionais comuns. Em particular, ocupam posições privilegiadas na estrutura silábica das línguas do mundo, sendo, ainda, em algumas destas as únicas consoantes permitidas como segundo membro de grupos silábicos, ou como primeiro membro de grupo em posição de coda. Vale ressaltar, entretanto, que, na sintaxe fônica do Português, assim como na de determinadas línguas como o Inglês, a líquida lateral [l] também é permitida como segundo membro de grupos consonantais.

No Português Brasileiro, como já aludido, os róticos abrangem várias classes de sons: aproximantes retroflexas, fricativas e *taps* com distribuição específica. Frise-se, porém, que, de acordo com Albano (2005, 2007), os róticos fricativos [x, ɣ, X, ɣ̃, h, h̃] só compartilham

com suas contrapartes vibrantes ou aproximantes [r, ɾ, ɹ, ʀ, ʁ, ʕ, ʁ] a presença de ressonâncias do trato vocal com características vocálicas.

Quanto à distribuição específica dos róticos, encontram-se, por exemplo, no falar da cidade de Fortaleza, quer em suas variedades estigmatizadas, quer em suas variedades de prestígio, realizações dos fonemas fricativos sonoros /v, z, ʒ / nas quais estes se neutralizam com o fonema vibrante múltiplo /r/, em sua variante aspirada [h], como em: “vou” – [‘hõ], “mesmo” – [‘mefmu], “a gente” – [a’hêʃi] (ALENCAR, 2007; ARAGÃO, 2009). Em estudo específico sobre variedades de prestígio em Fortaleza, ou seja, sobre o uso da norma culta nesta cidade, Macambira (1985, p. 153) diz que “[...] ordinariamente não se pronuncia o r seguido por consoante fricativa: garfo - /’gafu/ [...], curso /’kusu/.” O autor também aponta que, em contextos nos quais o r é seguido por s e este por outra consoante, o r também não é pronunciado, como, por exemplo, em: *perspectiva*, *superstição*, *perspicaz*.

Alencar (2007), por sua vez, ao analisar variantes do r no falar fortalezense, constata que, no contexto inicial, prevalece a fricativa glotal [h]; no contexto intervocálico, dá-se o contraste fonêmico entre /r/ e /r/; nos contextos pós-vocálicos medial e final, dependendo do contexto fonológico subsequente, podem ocorrer as variantes [h], [h̃], [r], [Ø]; o apagamento em posição pós-vocálica final, como em *mar*, é mais forte e recorrente do que em posição pós-vocálica medial, como, por exemplo, em *parque*.

Aqui, vale lembrar que os róticos, em início de sílaba, se realizam como fricativa glotal [h] ou *tap* [r], por exemplo: *rumo*, *Israel*, *carro* e *caro*. Em algumas localizações da região sul, entretanto, podem se encontrar vibrantes múltiplas (ou trills) alveolares nesses contextos.

Já no que diz respeito aos encontros consonantais, ou *clusters*, como em *prato*, consta sistematicamente no português brasileiro um *tap* (CRISTÓFARO; CARMAGOS, 2016). Essa sistematicidade, no entanto, se dá apenas de um ponto de vista eminentemente fonológico, pois, sob a ótica fonética, como veremos mais adiante, podem ocorrer, em vez do *tap*, nesses contextos, variantes como *tap* fricativo alveolar, vibrante múltipla alveolar, entre outras. Some-se a isto o que já dizia Seraine (1972, p. 12) a respeito de realizações de *clusters* com *tap* na região Nordeste do país: “No colóquio descuidado, até entre pessoas cultas, é comum, em todo o Nordeste, a síncope do r, pertencente a grupos consonânticos reais (br, dr, gr etc.) [...]”.

#### 4.4.1 O tap, o flap e o trill

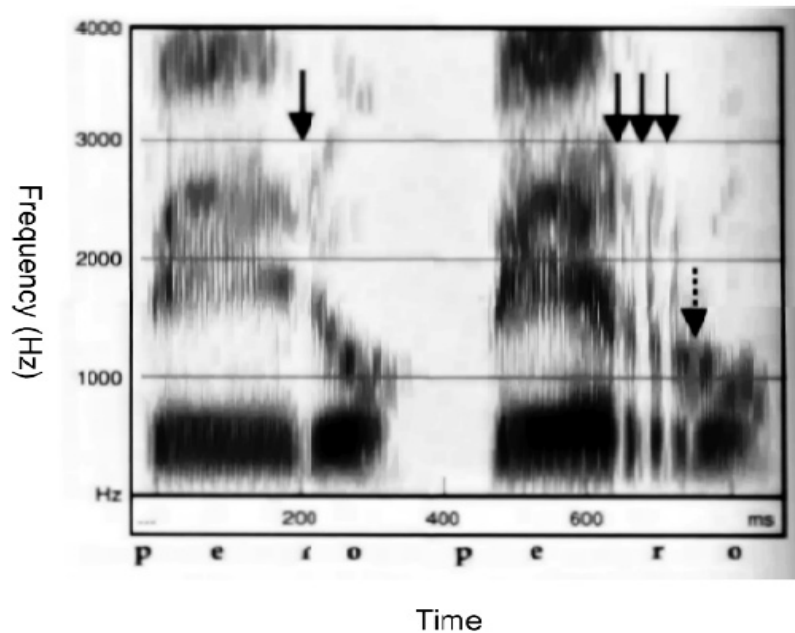
O *tap* pertence a uma classe de róticos que realizam invariavelmente apenas uma breve e curta oclusão. (LADEFOGED, MADDIESON, 1996, p. 230). Tal classe de róticos se constitui de *taps* e/ou *flaps*.

Há linguistas, como Lindau (1985), autora de *The story of /r/*, os quais não distinguem *tap* de *flap*. Por outro lado, há aqueles que fazem, sim, tal distinção: Ladefodeg e Maddieson (1996, p. 231), por exemplo, propõem que o flap é um som em que um breve contato é feito a partir de um movimento de um articulador ativo em direção tangencial à região de contato, atingindo, apenas de passagem, a superfície superior do trato vocal. Quanto ao *tap*, propõem que ele é um som no qual um breve contato entre os articuladores também é realizado, porém a partir de um movimento de um articulador ativo que se faz diretamente - e não tangencialmente como o *flap* - em direção ao céu da boca. Ambos, entretanto, são coronais.

Assim, os *flaps* se realizam tipicamente com a ponta da língua encurvada para trás, em um movimento que segue em direção ao cume alveolar, atingindo esse cume só de passagem, deslizando, em seguida, num rápido movimento para frente; já os *taps* são feitos tipicamente pelo movimento direto da ponta da língua com a região dental ou alveolar. Em síntese, pode-se dizer, *tap* e *flap* se distinguem porque este se realiza com a ponta da língua encurvada, enquanto aquele, não.

Já no que concerne a comparações entre o *tap* e o trill, vale ressaltar determinados aspectos espectrográficos destes sons. O *tap*, como já aludido antes, se realiza a partir de apenas uma breve e curta oclusão. O trill, por sua vez, se distingue dele na medida em que pode realizar duas ou mais oclusões sucessivas. Observe-se a figura 4, na qual Ladefodeg (2003), a partir de duas palavras do Espanhol, aponta diferenças e semelhanças nas características espectrográficas de tais sons. Vejamos.

Figura 3 – Espectrogramas das palavras *pero* (porém) e *perro* (cachorro) do Espanhol



Fonte: Ladefoged (2003).

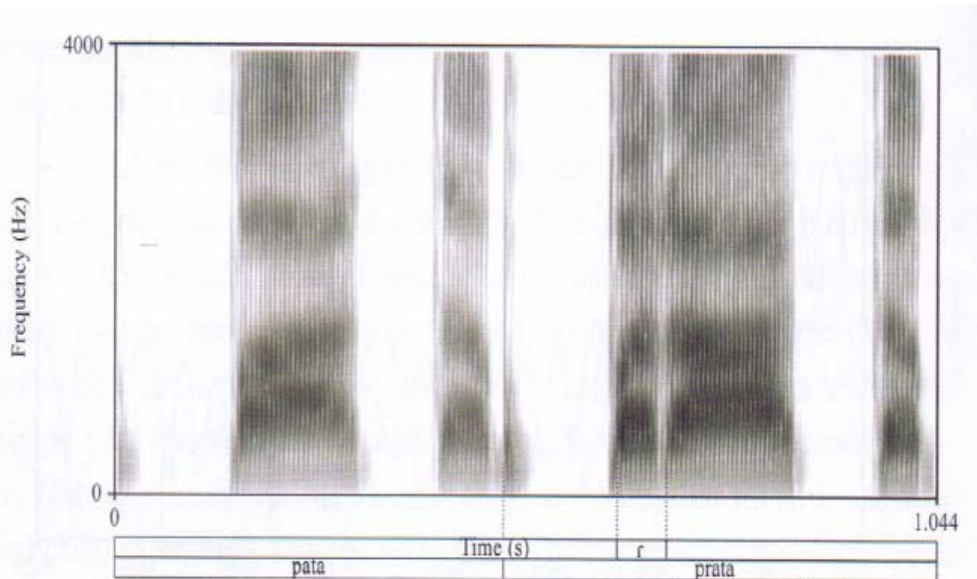
Na figura 3, na palavra *pero* (à esquerda), pode-se observar um clareamento espectrográfico o qual é correlato acústico da batida da ponta da língua referente ao *tap*. Observe-se, ainda nesta palavra, duas assinaturas acústicas que caracterizam o *tap*: i) elevação do F2 à margem direita da vogal que o antecede; e ii) queda brusca de energia de curta duração. Já na palavra *perro*, note-se que há um clareamento semelhante ao do *tap*, porém seguido de outros que se dão sucessivamente no tempo, os quais correspondem a batidas múltiplas da língua.

Como se pode notar, esses róticos se distinguem, em termos espectrográficos, a partir, basicamente, da quantidade de clareamentos espectrográficos correspondentes à quantidade de batidas da ponta da língua. Aqui, entretanto, cumpre destacar que nesta comparação com o trill, apresentou-se o *tap* em onset simples.

#### 4.4.2 Características do *tap* em onset complexo

Apresentam-se aqui determinadas características acústicas do *tap* em onset complexo, as quais, em dados momentos, são cotejadas com características de estruturas silábicas CV, isto é, sem a presença do *tap*. Também se apresentam variantes que podem ocupar a posição do *tap* no onset complexo em certas condições. Agora observemos a figura a seguir.

Figura 4 – Espectrograma de banda larga ilustrando contraste acústico entre as palavras pata (à esquerda) e prato (à direita) produzidas por um falante de Minas Gerais



Fonte: Barbosa e Madureira (2015, p.542).

Na figura 4 acima, especialmente, na palavra *prata* (à direita), observe-se que após a soltura da oclusiva e antes da queda curta e brusca de energia caracterizadora da batida do *tap*, há um segmento vocálico.

De acordo com Barbosa e Madureira (2015, p.541), o segmento acústico intercalado entre a oclusiva e a vogal seguinte caracteriza o *tap* em onset complexo em qualquer língua. Isso se dá porque no momento da soltura da oclusiva, o corpo da língua, por coarticulação antecipatória, já se encontra a caminho da posição da vogal que segue o *tap*, mas a ponta da língua ainda não tocou a região alveolar.

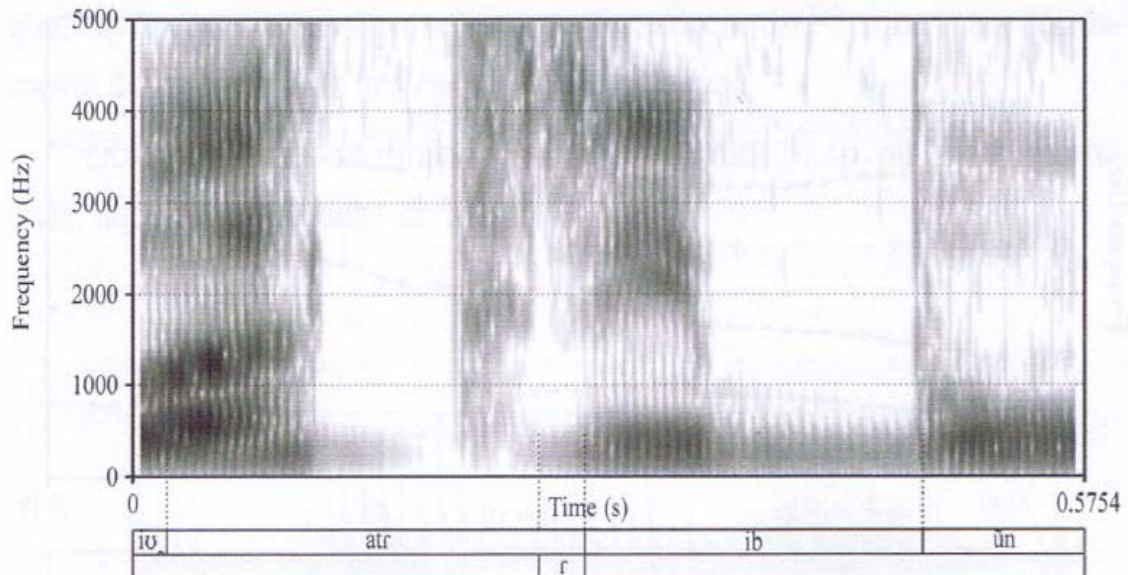
Ainda na figura 4, observe-se que nas palavras *pata* (à esquerda) e *prato* (à direita), o padrão formântico do elemento vocálico continua, como se não houvesse batida, durante a vogal tônica. O trecho todo, que envolve *tap* e vogal tônica, tem o mesmo padrão do trecho de vogal tônica da palavra *pata*, com exceção do clareamento espectrográfico que é correlato acústico da batida da ponta da língua.

Também consta no Português Brasileiro realizações do *tap* em onset complexo como fricativa alveolar não vozeada. Segundo Barbosa e Madureira (2015), isso ocorre em função do efeito de saliência (ou acento frasal), uma vez que numa situação de fala rápida, uma palavra perde a saliência que tinha numa situação de fala mais cautelosa (ou hipoarticulada), em que os segmentos acústicos não atingem seus alvos articulatorios, acarretando valores menores de durações ou modificações no modo (oclusivas realizadas como fricativas,



fricativas como aproximantes) ou ponto de articulação (vogais baixas podem se alçar, por exemplo). Vejam-se, agora, na sequência, as figuras 5 e 6, logo a seguir.

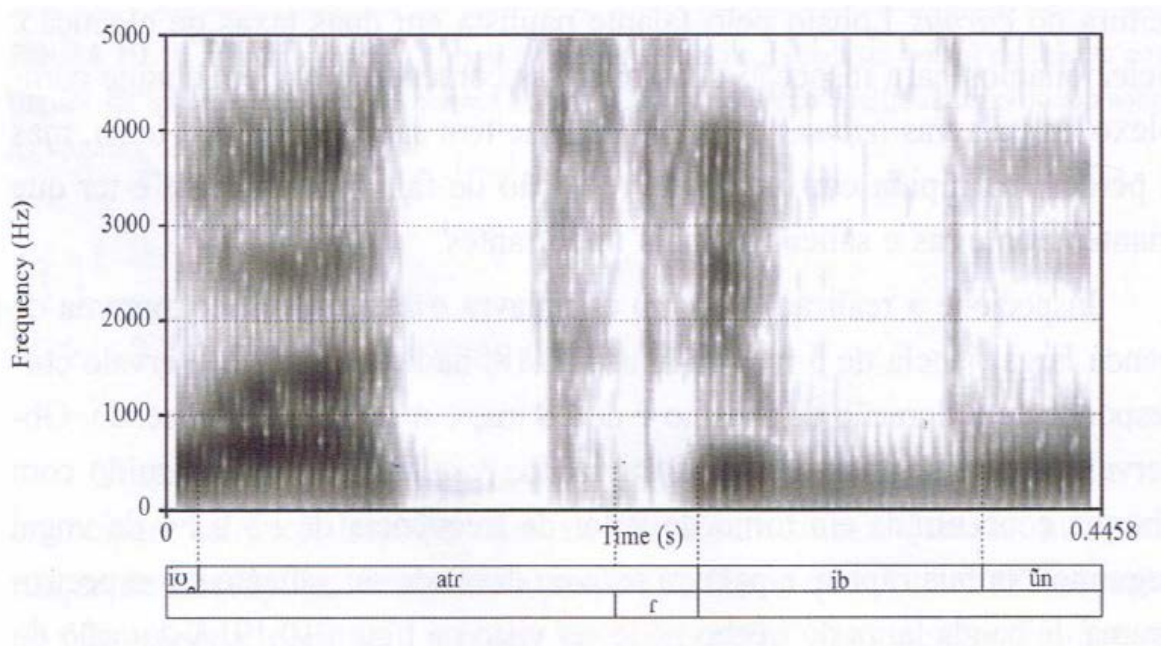
Figura 5 – Espectrograma de banda larga (janela de 5 ms) de trecho da palavra *tribuna* para ilustrar configuração acústica do *tap* em falante paulista na leitura do corpus Lobato na taxa lenta



Fonte: Barbosa e Madueira (2015, p. 558).

Como se pode observar na figura 5, na qual se ilustra a realização do *tap* na palavra *tribuna*, em taxa lenta, o *tap* se mostra como um segmento vozeado, pois apresenta barra de sonoridade e segundo Barbosa e Madueira (2015) tem duração de 28 ms. Pode-se notar que ele está fricativizado, pois há considerável ruído com energia concentrada em torno do valor de frequência F3 e F4 da vogal seguinte. Já na fala rápida, ele se realiza de forma diferente. Vejamos isso considerando a figura 6, logo a seguir.

Figura 6 – Espectrograma de banda larga (janela de 5 ms) de trecho da palavra *tribuna* para ilustrar configuração acústica do *tap* em falante paulista na leitura do *corpus* Lobato na taxa rápida.



Fonte: Barbosa e Madureira (2015, p. 558).

Na figura 6, encontra-se a palavra *tribuna*, porém, agora, produzida na taxa rápida, na qual essa palavra deixa de ser saliente. A primeira diferença em relação a outra (em taxa lenta), reside na duração do *tap*, que, antes, era de 28 ms, e, agora, de 40 ms, segundo Barbosa e Madureira (2015). No espectrograma da figura 6, podem-se observar, ainda, outras diferenças. Note-se a ausência de barra de vozeamento e que o ruído se distribui a partir de 1500 Hz. Trata-se, portanto, de uma fricativa não vozeada. Esta é percebida por conta de sua curta duração, como o fonema /r/. No Português Europeu, Jesus e Shaddle (2005) também reportam essa variante do *tap* no onset complexo, entretanto, sugerem que ela é mais comum em posições finais de palavras.

## **5 METODOLOGIA**

Este capítulo se ocupa, inicialmente, em caracterizar a pesquisa e, na sequência, apresenta os demais aspectos metodológicos. Para tanto, estrutura-se em 7 tópicos, a saber: 5.1 Caracterização da pesquisa; 5.2 Lócus da pesquisa; 5.3 Participantes; 5.4 Procedimentos de coleta de dados; 5.5 Hipóteses; 5.6 Variáveis; 5.7 Procedimentos de análise de dados.

### **5.1 Caracterização da pesquisa**

A presente pesquisa não só visa interpretar processos fônicos atinentes à produção do *tap* em onset complexo por crianças com e sem desvio fonológico à luz da Fonética Acústica, mas também quantificar esses dados analisando-os estatisticamente quando oportuno.

No que concerne aos procedimentos técnicos desta, foram realizados testes de nomeação de figuras, em cabine com tratamento acústico, com dois grupos distintos, o GDF, que é constituído de crianças com desvio fonológico (N=4), e o GSD, de crianças sem o desvio (N = 4).

Aqui, cumpre ressaltar que o projeto desta pesquisa teve aprovação do Comitê de Ética em Pesquisa (CEP) da Universidade Federal do Ceará, (número do Parecer: 1.956.879). Além disso, também recebeu autorização da Diretoria de Pesquisa, Desenvolvimento e Inovação (DPDI), do Núcleo de Atenção Médica Integrado (Nami), da Universidade de Fortaleza, para o uso de suas dependências físicas. Ressalte-se, ainda, que esta pesquisa só teve início após a aprovação do CEP bem como após a autorização do Nami.

### **5.2 Lócus da pesquisa**

O lócus da pesquisa – especialmente a coleta de dados - se deu nas instalações do Núcleo de Atenção Médica Integrado (Nami), da Universidade de Fortaleza (Unifor). O Nami oferece serviços diversos à população, tais como Psicologia, Fisioterapia, Fonoaudiologia e outros.

O setor de Fonoaudiologia, cumpre ressaltar, é referência na cidade de Fortaleza e atua na avaliação, prevenção e tratamento dos problemas de voz, fala, linguagem e audição, dentre outros. Sua infraestrutura abrange salas e cabines acusticamente tratadas, ambiente este adequado para a coleta de dados que serão analisados acusticamente.

O Nami atende, gratuitamente, a crianças e adultos com problemas diversos, entre os quais o desvio fonológico. Com isso, foi possível selecionar a maior parte dos 18 participantes lá no Nami mesmo, exceto quatro. Esses foram solicitados na Escola Municipal de Ensino Fundamental Professor Francisco Maurício de Matos Dourado, a qual se localiza próximo ao Nami.

### **5.3 Participantes da pesquisa**

Nesta pesquisa, os critérios de inclusão dos participantes foram os seguintes: i) assinar Termo de Consentimento Livre e Esclarecido (assinado por pai/mãe/responsável); ii) assinar Termo de Assentimento (assinado pelo participante); iv) ter idade entre 5 e 10 anos; iv) ter diagnóstico de desvio fonológico expedido por fonoaudiólogo (para o participante com o desvio); v) apresentar histórico sem problemas na fala e sem comprometimento intelectual (para o participante sem o desvio).

Ao todo, participaram 18 crianças, mas, desse total, apenas quatro apresentavam diagnóstico de desvio fonológico. Assim, formou-se, com estas quatro crianças, o grupo GDF, que significa grupo de crianças com desvio fonológico.

Em seguida, buscou-se, a partir das catorze crianças restantes, selecionar as crianças que iriam constituir o grupo de crianças sem o desvio, ou seja, o GSDF – grupo de crianças sem o desvio fonológico - de modo que este grupo se compatibilizasse relativamente com o outro, o GDF, em relação aos fatores idade e sexo.

Considerando essa relativa compatibilidade, e o máximo de homogeneidade possível que se podia obter na formação desses grupos, o GSDF foi formado com 4 crianças, com faixa etária e sexo relativamente compatíveis com as do outro grupo.

Com os grupos já estabelecidos, vale dizer que o GDF é formado por dois meninos e duas meninas, enquanto o GSDF tem três meninos e uma menina. Essa foi a melhor configuração possível, levando-se em conta aí todos os dezoito participantes que foram gravados na pesquisa.

Convém dizer agora que se optou por codificar os nomes dos participantes, para que tivessem suas identidades preservadas. Na codificação, tem-se, por exemplo: pe1MV, onde ‘pe’ significa ‘participante do grupo GDF’; o número 1, aí, indica a ordem de participação nos testes de nomeação de figuras; e as letras maiúsculas aludem ao nome do participante. No caso dos participantes do grupo GSDF, tem-se esse mesmo padrão de codificação, mudando-se apenas “pe” para pc – letras que aludem a participante do GSD.

Vejam, agora, informações mais específicas sobre os participantes do grupo de crianças com desvio fonológico (GDF):

i) participante pe1MV:

Sexo: masculino;

Idade: 9 anos;

Diagnóstico: desvio fonológico/gravidade: severo;

Instrumento de avaliação do diagnóstico: ABFW.

ii) participante pe2CL:

Sexo: masculino;

Idade: 5 anos e 7 meses;

Diagnóstico: desvio fonológico/gravidade: moderado;

Instrumento de avaliação do diagnóstico: ABFW – Fonologia – imitação e nomeação;

iii) participante pe3MA:

Sexo: feminino;

Idade: 5 anos;

Diagnóstico: desvio fonológico/ gravidade leve;

Instrumento de avaliação do diagnóstico: Inventário Fonético/Fonológico.

iv) participante pe4LF:

Sexo: feminino;

Idade: 7 anos e 2 meses;

Diagnóstico: desvio fonológico/gravidade leve;

Instrumento de avaliação do diagnóstico: Inventário Fonético/Fonológico.

Agora vejamos informações específicas dos participantes do grupo controle.

i) participante pc1JA:

Sexo: masculino;

Idade: 7 anos e 4 meses;

Histórico: segundo informações de seu pai, não há problemas em sua fala.

ii) participante pc2DG:

Sexo: masculino;

Idade: 5 anos;

Histórico: segundo informações de sua mãe, não há problemas em sua fala.

iii) participante pc3MN

Sexo: feminino;

Idade: 5 anos e 2 meses;

Histórico: segundo informações de sua professora, não há problemas em sua fala.

iv) participante pc4AF

Sexo: masculino;

Idade: 7 anos e 1 mês;

Histórico: segundo informações de sua mãe, não há problemas em sua fala.

#### **5.4 Procedimentos de coleta de dados**

O procedimento de coleta de dados consistiu basicamente em realizar com os participantes um teste de nomeação de figuras, o qual acontecia numa cabine acusticamente tratada. Respeitando-se o princípio *ceteris paribus*, a tarefa dos participantes foi tão somente nomear/responder perguntas atinentes às figuras temáticas apresentadas no teste de nomeação. Além da estratégia de nomeação de figuras, vale ressaltar, também foram usadas, quando oportuno, outras estratégias para eliciação de dados, como a de repetição da palavra quando o participante não sabia nomear a figura.

O objetivo do teste de nomeação era fazer com que os participantes produzissem oralmente palavras-alvo dissílabas paroxítonas com o *tap* em onset complexo. Essas palavras foram as seguintes: *prato, trave, craque, braço, dragão, grade, fraco*.

##### **5.4.1 Instrumentos de coleta**

Para a coleta dos dados, utilizou-se um teste de nomeação de figuras (temáticas) a partir das quais se buscava eliciar sete palavras-alvo, seis das quais se iniciavam com oclusivas, três surdas e três sonoras, e uma, com fricativa alveolar surda: *prato, braço, trave, dragão, craque, grade e fraco*. O pesquisador usava estratégias diversas que sugeriam a

nomeação das figuras pelos participantes. Entre tais estratégias destacaram-se perguntas a respeito das figuras, contação de história e, em alguns casos, estratégias de repetição de palavras.

Os testes de nomeação de figuras foram gravados no interior de uma cabine acusticamente tratada, na qual se utilizaram um microfone de lapela e um gravador *Zoom* modelo H4n (*Handy Recorder*), com taxa de amostragem de 44100 Hz.

Ao entrar na cabine, o participante era convidado a se sentar e em sua roupa se colocava o microfone de lapela, com, aproximadamente, 12 cm de distância da boca. Em seguida, iniciava-se a gravação do teste o qual se repetia três vezes consecutivas para cada participante.

O pai/mãe/responsável e o participante eram instruídos sobre os testes e só participavam dele após assinarem termo de consentimento livre e esclarecido (responsável) e termo de assentimento (participante). Daí o responsável pelo participante era convidado a acompanhar visualmente todos os procedimentos, porém, do lado de fora da cabine, através de um vidro transparente. Na cabine, entretanto, só entravam o experimentador e o participante.

Todos os participantes de ambos os grupos tiveram seus dados gravados em cabine acusticamente tratada no Núcleo de Atenção Médica Integrada (Nami/Unifor).

#### **5.4.2 Corpus**

Com a coleta de dados, obteve-se um corpus de 168 palavras, que correspondem a gravações das sete palavras-alvo, *prato*, *braço*, *trave*, *dragão*, *craque*, *grade*, *fraco*. Frise-se aí que o pesquisador combinou elegeu palavras nas quais os clusters com tap eram precedidos por seis oclusivas e a fricativa labiodental surda. Tais palavras-alvo foram produzidas, cada uma, três vezes por cada participante de ambos os grupos GDF e GSDF.

As palavras-alvo seguiram um padrão de controle baseado nas seguintes condições: i) ser dissílaba paroxítona; ii) apresentar o *tap* em onset complexo seguido pela vogal tônica [a].

Assim, das sete palavras-alvo, somente a palavra-alvo *dragão*, que é oxítona, não seguiu a primeira condição. A escolha desta se deu em razão de não ter sido encontrada, no léxico, outra palavra mais adequada ao teste com crianças e que se compatibilizasse com os mencionados padrões de controle.

A escolha das sete palavras-alvo deste teste se baseou também em fatores subjetivos atinentes à faixa etária dos participantes. O teste foi rodado três vezes sucessivas para cada participante.

### 5.5 Hipótese básica

A hipótese básica desta tese foi a de que o grupo de crianças com desvio fonológico (GDF) se distingue significativamente do grupo de crianças sem o desvio (GSDF) na produção acústica de *clusters* com *tap*, especialmente, na produção do *tap* em onset complexo.

Decorreram daí as seguintes hipóteses específicas e previsões:

- O padrão acústico das produções do GDF referentes ao *tap* em onset complexo se caracteriza pela não-realização do *tap*, isto é, pela ausência de vestígios do sinal acústico referentes ao *tap* mesmo.
- O padrão acústico das produções do GSDF referentes ao *tap* em onset complexo se caracterizará como uma estrutura acústica CCV, pois os participantes deste grupo tenderão a realizar o *tap*, porém na posição desta consoante poderão, ainda, produzir em seu lugar variantes como *tap* fricativo alveolar surdo ou sonoro, vibrante múltipla alveolar, entre outras.
- GDF e GSDF se distinguem quanto às medidas do VOT das oclusivas surdas, sendo seus valores significativamente menores no GDF.
- GDF e GSDF se distinguem quanto ao desvozeamento das oclusivas sonoras, sendo tal processo mais frequente no GDF.
- GDF e GSDF se distinguem quanto à duração relativa das fricativas labiodentais, sendo maior no GDF.
- GDF e GSD se distinguem quanto à composicionalidade espectral das fricativas labiodentais.
- Nos encontros consonantais em que o *tap* é precedido por consoantes oclusivas alveolares [t] e [d] poderão ocorrer, além do *tap*, variantes como vibrantes múltiplas alveolares no GSDF.



## 5.6 Variáveis

As variáveis independentes são:

- i) GDF e GSDF;
- ii) Sexo

As variáveis dependentes são:

- iii) oclusivas surdas (parâmetro – VOT relativo)
- iv) oclusivas sonoras (parâmetro – VOT relativo)
- v) fricativa labiodental surda (parâmetro: duração relativa e composição espectral)
- vi) *Tap* no onset complexo (forma de onda, duração, espectrograma e, quando necessário, transição formântica (F1, F2 e F3))

## 5.7 Procedimentos de análise acústica dos dados

Após o processamento digital dos dados, o procedimento seguinte consistiu em transferi-los do gravador H4n para o software de análise acústica *Praat*, no qual procedeu-se, inicialmente, a análise dos espectrogramas e das formas de onda das consoantes oclusivas surdas [p], [t], [k], as quais precediam o *tap* nos *clusters* das palavras-alvo *prato*, *trave* e *craque*, nas quais se mediram seus respectivos VOTs.

Os dados dos participantes de ambos os grupos, GDF e GSDF, foram analisados inicialmente no software *Praat*, a partir do qual se mediram os VOTs das oclusivas surdas. Depois, mediu-se também a duração em milissegundo de cada palavra na qual a oclusiva se encontrava. Usou-se a medida do VOT relativo para o estabelecimento de comparações entre as produções de um grupo e as de outro. Tal medida se obteve da seguinte forma: dividia-se o valor do VOT pela duração da palavra inteira e, depois, multiplicava-se o resultado por 100.

Optou-se por trabalhar com o VOT relativo porque este, diferentemente do VOT absoluto, evita vieses que possam estar relacionados, por exemplo, com a taxa de elocução (fala mais lenta ou mais rápida) permitindo uma análise mais sólida dos dados.

Não foi possível, entretanto, calcular a medida da palavra inteira a partir de seu ponto inicial, que deveria coincidir nas gravações com a parte inicial da oclusiva, correspondente ao fechamento. Essa impossibilidade se estabeleceu porque, na inspeção visual dos dados, observou-se que o fechamento e o silêncio da oclusiva surda se confundiam geralmente com um momento de hesitação do participante no teste de nomeação de figuras. Daí, optou-se por considerar a palavra inteira como se ela começasse somente a partir da explosão da oclusiva.

Depois da análise acústica, procedeu-se a análises estatísticas dos valores obtidos dos VOTs relativos das oclusivas surdas.

Seguiram essa esquemática mesma os procedimentos sobre as oclusivas sonoras, [b], [d], [g], extraídas das palavras-alvo *braço*, *dragão*, *grade*. As medidas se realizavam pelo espectrograma de banda larga, aliado à forma de onda. Nesse sentido, media-se a duração da barra de vozeamento e depois se dividia seu valor pela duração da palavra inteira, multiplicando o resultado por 100. Já a ausência da barra de vozeamento foi utilizada como critério para se caracterizar o desvozeamento dessas consoantes.

Depois, procedeu-se à análise da fricativa labiodental surda [f] na palavra-alvo *fraco*. Primeiramente, mediram-se as durações relativas desta consoante e, após isso, foram realizadas análises espectrais das fricativas.

Para se eliciar a fricativa na palavra *fraco*, apresentava-se ao participante uma única figura temática com dois cachorros distintos, os quais se contrastavam, como se fossem dois halterofilistas, sendo um deles aparentemente muito forte, com músculos evidentes, e o outro, muito magro e com aparência frágil. Daí, o experimentador estabelecia comparações entre os dois animais apontando primeiro para o cachorro mais forte e dizendo “este é o cachorro mais forte...” e, logo em seguida, já apontando para o outro cachorro, perguntava “... e este é o mais?” Assim, o participante era instado a responder a pergunta com a palavra *fraco*.

Dessa forma, cada participante produziu três vezes a palavra-alvo *fraco*, já que o teste com todas as figuras era rodado três vezes.

Por fim, tem-se a análise do *tap*.

Resta dizer, ainda, que no *Praat*, especificamente no *Spectrogram settings*, trabalhamos principalmente com as seguintes configurações: *view range* (Hz): 0.0 a 4000.0. (exceto no caso das fricativas, para as quais se consideraram valores maiores no *view range*); *window length*: 0.005 (banda larga) e *dynamic range* (dB): 40.0.

## 6 ANÁLISES E DISCUSSÕES

Pretende-se, neste capítulo, analisar e discutir os dados produzidos a partir dos grupos GDF e GSDF desta pesquisa. Tendo isso em vista, apresentam-se, na sequência, as seguintes análises: do VOT das oclusivas surdas; do VOT das oclusivas sonoras (vozeamento x desvozeamento); da duração das fricativas; dos espectros de Fourier das fricativas; do *tap* em onset complexo. Assim, vejamos.

### 6.1 Do VOT das oclusivas surdas

Os dados dos participantes de ambos os grupos, GDF e GSDF, foram analisados inicialmente no software *Praat*, a partir do qual se mediram os VOTs das oclusivas surdas. Depois, mediu-se também a duração em milissegundo de cada palavra na qual a oclusiva se encontrava. Usou-se a medida do VOT relativo para o estabelecimento de comparações entre as produções de um grupo e as de outro.

As palavras-alvo no teste, já assinaladas anteriormente, eram todas dissílabas paroxítonas, com o onset complexo na sílaba tônica. Assim, tinha-se no onset complexo a seguinte estrutura: oclusiva surda + líquida não lateral (o *tap*)+ [a] como vogal tônica<sup>5</sup>, como se pode observar nas palavras-alvo [p]rato, [t]rave e [k]raque.

Estas e outras palavras-alvo do teste, como as que traziam as oclusivas sonoras e a fricativa labiodental seguidas pelo *tap* no onset complexo, além de outras palavras, que eram distratoras<sup>6</sup>, foram produzidas três vezes por cada participante dos grupos GDF e GSDF.

Aqui, vale destacar o seguinte. Geralmente, o VOT de crianças sem comprometimento fonológico apresenta valores relativamente altos se comparados aos de adultos, devido ao trato vocal menor da criança, que produz maior pressão intraoral (Barbosa e Madureira, 2015). Já no que diz respeito à comparação de VOT entre crianças com e sem comprometimento fonológico, como ainda não se encontram trabalhos nesse sentido no Português Brasileiro, convém destacar o de Barroco et.al. (2007) no Português Europeu. Esses autores (2007) estudaram duas crianças do sexo feminino, uma de 8:7 sem comprometimento

<sup>5</sup> Segundo Silva (1996, p.80): “[...] para [r] em grupos, a variável que age sobre os dados, para que ocorra diferenciação entre eles, é a natureza da vogal tônica.”

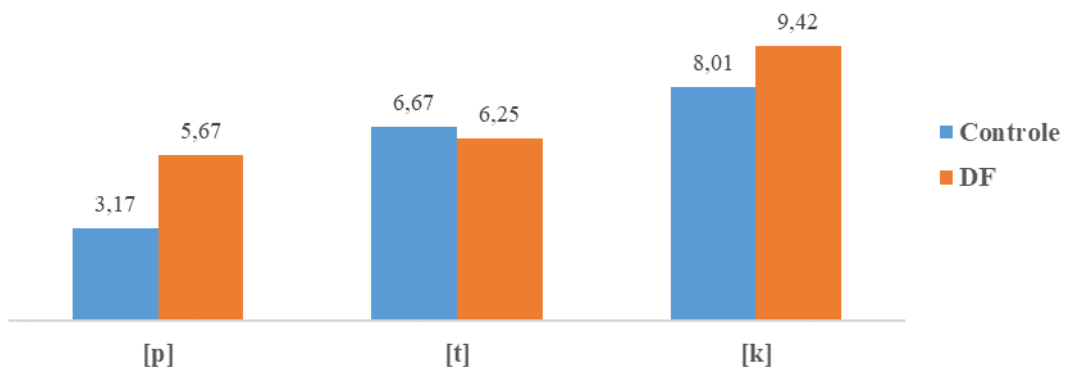
<sup>6</sup> As palavras distratoras eram produzidas a partir das figuras que eram apresentadas às crianças. Por exemplo, numa mesma página se apresentavam três figuras, uma ao lado da outra - uma faca, um prato e um garfo. Nesse caso, a palavra-alvo era apenas o prato e as outras, a faca e o garfo, eram distratoras. Outros tipos de distratoras também se produziam por meio de outras figuras em que se apresentavam as demais palavras-alvo.

fonológico e outra de 7:6 anos com perturbação fonológica (excesso de uso de desvozeamento dos sons para idade), no qual se encontraram valores de VOT mais baixos para a criança com comprometimento fonológico.

Tendo em vista essas considerações, nossas previsões nesse sentido foram as de que o VOT das oclusivas surdas produzidas pelo grupo de crianças com desvio fonológico (GDF) seriam mais baixas que as do grupo de crianças sem o desvio (GSD).

Assim, assumiu-se a hipótese de que haveria diferença significativa entre os grupos nas médias dos VOT das oclusivas surdas. Considerando isso, vejamos o gráfico a seguir, com as médias dos VOTs relativos de cada uma das três oclusivas produzidas pelos dois grupos.

Gráfico 1 – Média de valores do VOT relativo de [p], [t], [k] dos grupos GDF (DF) e GSDF (Controle)



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Inicialmente, como se pode observar no gráfico 1, as médias dos VOTs da bilabial surda [p] foram menores que as da alveolar surda [t], e as desta foram menores que as da velar [k], em ambos os grupos, GDF e GSDF. Cumpre ressaltar que essa ordem de valores, menores para as bilabiais, intermediárias para as alveolares e maiores para as velares, é esperada para línguas como o Português.

Quanto às diferenças entre as produções dos grupos GDF e GSDF em relação ao VOT relativo das oclusivas surdas, pode-se observar, ainda no gráfico 1, que as médias dos VOTs do grupo com desvio fonológico (GDF) foram menores apenas em relação ao VOT da oclusiva alveolar surda [t]. Neste ponto, esse achado converge com o de Barroco *et al.* (2007).

Contudo, em nossas previsões esperávamos médias mais baixas para o grupo com desvio fonológico não apenas em uma oclusiva surda, mas nas três surdas, aspecto este que

não se confirmou, já que as médias dos valores dos VOTs das oclusivas bilabial [p] e velar [k] foram maiores no grupo sem desvio fonológico (GSD).

Tendo isso em vista, buscou-se verificar se tais diferenças seriam estatisticamente significativas. Para tanto, realizou-se um teste estatístico não-paramétrico de Wilcoxon-Mann-Whitney, para variáveis independentes. A escolha desse teste estatístico e não de outros, como o Teste-T, por exemplo, se deu em virtude do tamanho da amostra (tamanho quatro).

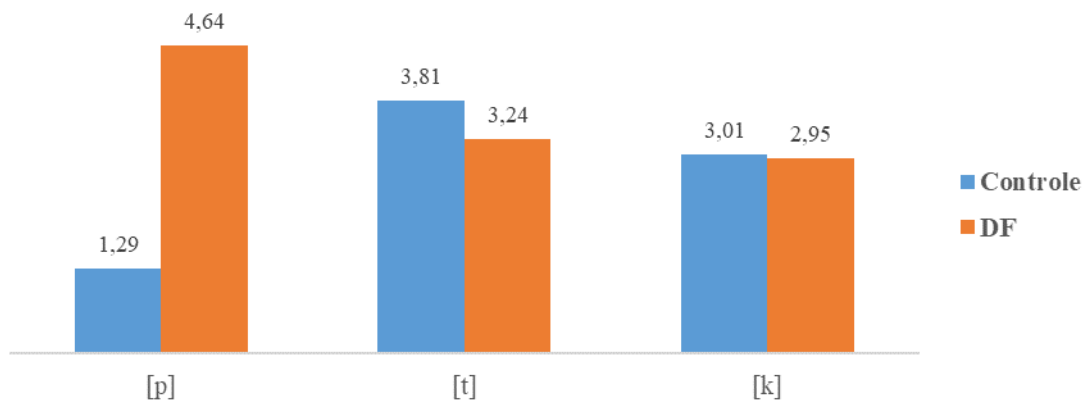
No teste Wilcoxon-Mann-Whitney, a hipótese nula era a de que não haveria diferença significativa entre os grupos e a hipótese alternativa era a de que haveria, sim. Ao se realizarem esses testes, obtiveram-se os seguintes p-valores, para [p] e [k], respectivamente:  $p=0,4857$  e  $p=0,8885$ . Como se pode notar, essas diferenças não foram significativas. Quer dizer, embora as médias do grupo sem desvio fonológico tenham sido descritivamente menores que as do grupo com desvio fonológico, essas diferenças não foram significativas.

Quanto às médias do VOT da alveolar [t], observou-se, como já dito, que as do grupo com desvio fonológico foram menores que as do grupo controle, o que converge com uma de nossas previsões e hipóteses de pesquisa. Cumpre dizer, entretanto, que apesar dessa convergência com nossas previsões, também não se encontrou diferença estatisticamente significativa ( $p \approx 1,0$ ), nas médias dos VOTs da alveolar surda.

Em síntese, constatou-se que os grupos GDF e GSDF, embora apresentem algumas diferenças nas produções dos VOTs relativos das oclusivas surdas, eles não se distinguem significativamente quanto a isso.

A seguir, apresentamos medidas de dispersão desses grupos, como o desvio padrão e o coeficiente de variância, as quais agregam, de certo modo, mais informações sobre os detalhes fonéticos em pauta. Vejamos.

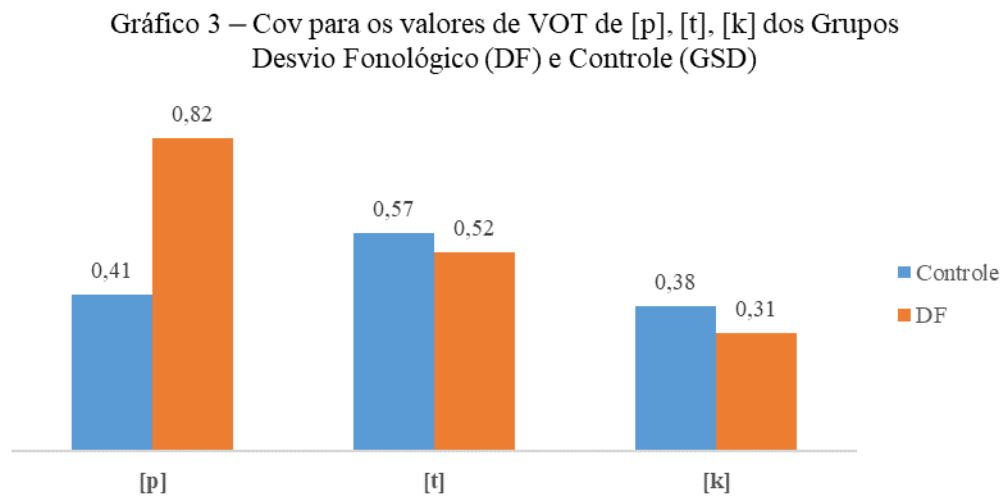
Gráfico 2 – Desvio Padrão dos valores de VOT de [p], [t], [k] dos Grupos Desvio Fonológico (DF) e Controle



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Note-se, no gráfico 2, que o desvio padrão da consoante [p] produzida pelo grupo com desvio fonológico (DesvPad=4,64) é maior que o do grupo GSDF (DesvPad=1,29). Isso sugere a existência de maior instabilidade na produção do VOT da oclusiva bilabial surda pelo grupo com desvio fonológico.

Agora vejamos o gráfico 3, que apresenta o coeficiente de variação (CoV) do VOT das oclusivas surdas produzidos pelos grupos GDF e GSDF.



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Note-se, no gráfico 3, que o CoV<sup>7</sup> de [p] e [t] são maiores que os de [k] em ambos os grupos, GDF e GSDF. Apesar desta elevada variabilidade de [p] e [t] em ambos os grupos, de acordo com Gerosse et.al. (2007), o VOT de /p/ e /t/ diminui com a idade. Já a consoante velar parece ser mais estável em relação à idade.

A seguir, vejamos a Tabela 1, a qual sintetiza os valores das médias, dos desvios padrão e do coeficiente de variação dos VOTs das oclusivas surdas produzidos pelos grupos com desvio fonológico e controle, os quais foram apresentados neste tópico.

<sup>7</sup> O CoV é uma medida relativa de dispersão útil para comparar, em termos relativos, o grau de concentração em torno da média. Ela é calculada através do desvio padrão dividido pela média e multiplicado por 100. Teoricamente, desvios padrões menores que 10% mostram que a série possui uma baixa dispersão. Entre 10% e 20%, indicam média dispersão, e acima de 20%, o CoV tem alta dispersão.

Tabela 1– Média, desvio padrão (DesvP), coeficiente de variância (CoV) do VOT de [p], [t], [k] dos grupos GDF e GSD.

<b>Grupos</b>		
<b>[p]</b>	<b>GSD</b>	<b>GDF</b>
<b>Média</b>	<b>3,17</b>	<b>5,67</b>
<b>DesvP</b>	<b>1,29</b>	<b>4,64</b>
<b>CV</b>	<b>0,41</b>	<b>0,82</b>
<b>[t]</b>	<b>GSD</b>	<b>GDF</b>
<b>Média</b>	<b>6,67</b>	<b>6,25</b>
<b>DesvP</b>	<b>3,81</b>	<b>3,24</b>
<b>CoV</b>	<b>0,57</b>	<b>0,52</b>
<b>[k]</b>	<b>GSD</b>	<b>GDF</b>
<b>Média</b>	<b>8,01</b>	<b>9,42</b>
<b>DesvP</b>	<b>3,01</b>	<b>2,95</b>
<b>CoV</b>	<b>0,38</b>	<b>0,31</b>

Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Como se vê na Tabela 1, consta a síntese das medidas analisadas das variáveis independentes: produções dos grupos GDF e GSDF (controle), e da variável dependente oclusiva surda, cujo parâmetro foi o VOT.

### ***6.1.1 Efeitos do sexo no VOT das oclusivas surdas***

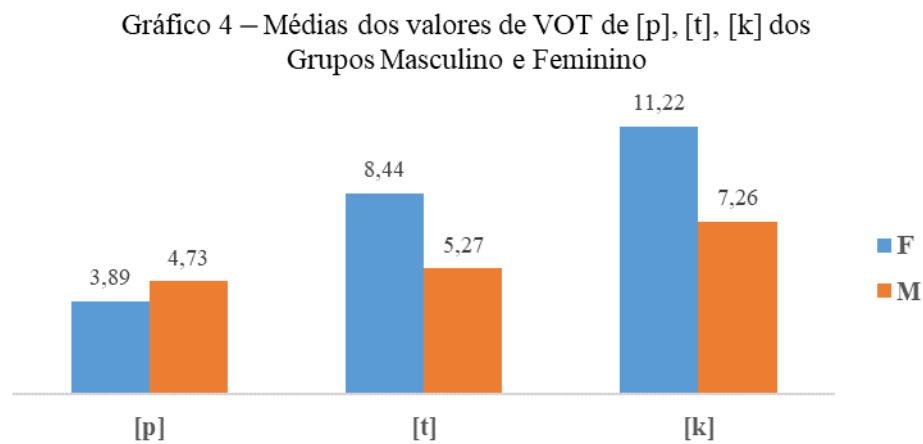
Nesta análise, os participantes de ambos os grupos, GDF e GSDF, foram reagrupados em dois outros grupos distintos, com base na variável independente sexo. Assim, distribuíram-se da seguinte maneira: grupo de meninos (N=5) e grupo de meninas (N=3).

Com base na produção desses grupos de meninos e meninas, realizou-se novamente o teste estatístico não-paramétrico Wilcoxon-Mann-Whitney, para comparar as médias dos valores dos VOTs das oclusivas surdas produzidas por eles.

A hipótese era a de que haveria diferença significativa entre meninos e meninas na produção do VOT, uma vez que as meninas tendem a ter tratos vocais menores que os meninos. Além disso, no sexo feminino o palato tende ser mais baixo. Daí diversos efeitos de

filtragem distintos podem se esperar nesses grupos, alguns dos quais podem incidir no VOT, sobretudo na consoante velar. Sobre esta, aliás, a nossa hipótese é a de que o VOT das meninas será maior. Já em relação à bilabial e a alveolar, também esperamos VOT mais longos para as meninas.

Assim, na sequência, apresentamos a estatística descritiva desses valores, apresentando os resultados do reportado teste estatístico logo em seguida. Vejamos, então, o gráfico 4, que expõe as médias dos valores dos VOTs relativos das oclusivas surdas dos meninos e das meninas.



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Observe-se, no gráfico 4, que a ordem dos valores das oclusivas foi preservada em ambos os grupos (meninos e meninas). Ou seja, tais grupos produziram valores de VOT menores para a bilabial [p], intermediários para alveolar [t] e maiores para a velar [k]. Cumpre lembrar, como já dito antes, que esta ordem de aumento dos valores dos VOTs das oclusivas surdas, menores para bilabial, intermediários para alveolar e maiores para velar, se coaduna com o que se espera em línguas como o Português.

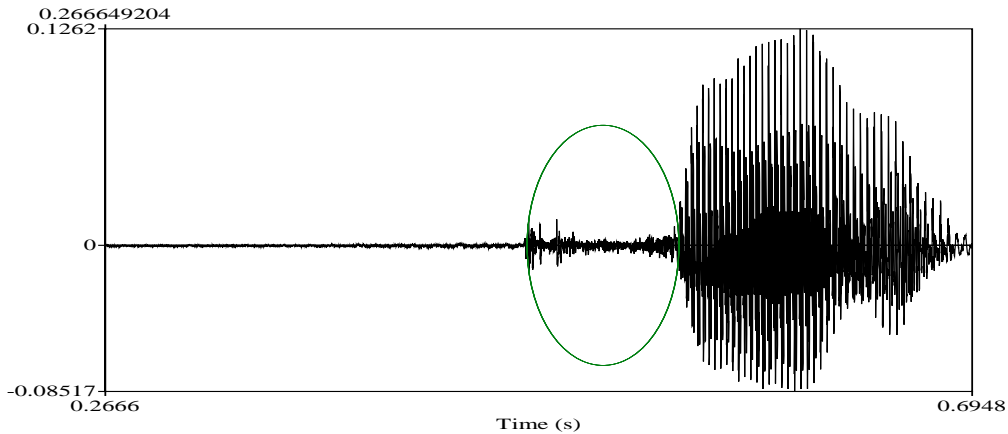
Considerando ainda o gráfico 4, observem-se, agora, as diferenças entre meninos e meninas na produção das oclusivas surdas. Veja-se que a média do VOT da oclusiva bilabial [p] dos meninos, (4,73), é mais alta que a das meninas, (3,89), entretanto, essa diferença não foi estatisticamente significativa ( $p \approx 1,00$ ). A média do VOT da alveolar [t] também não difere significativamente entre meninos e meninas ( $p = 0,2302$ ).

Já no que diz respeito ao VOT da oclusiva velar [k], observa-se que a média das meninas, (11,20), foi mais alta que a dos meninos, (7,26). E mais, verificou-se que essa diferença ( $p = 0,03571$ ) foi estatisticamente significativa, para o nível de significância de 5%.



Vejam-se, a seguir, as figuras 7 e 8 as quais ilustram esse achado com dados de dois participantes, uma menina e um menino.

Figura 7 – Forma de onda da sílaba tônica da palavra-alvo *craque* produzida por uma menina

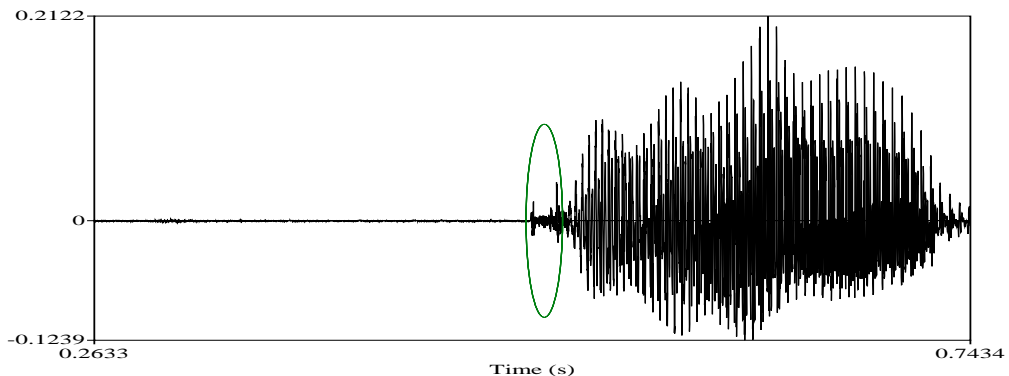


Fonte: elaborado pelo autor.

Legenda: a elipse pontilhada indica o VOT da oclusiva velar surda, cuja duração foi de 75ms, na sílaba tônica da palavra-alvo *craque* produzida por uma menina. Vale ressaltar que, nesta produção, não houve a realização do *tap*. Assim, nesta figura, tem-se a sílaba [ka] em vez de [k r a].

Como se pode ver na figura 3, destaca-se o VOT da oclusiva surda velar [k] produzido por uma menina. Sua duração bruta foi de 75 ms. Já na figura 4, destaca-se o VOT da oclusiva velar surda [k] produzida por um menino. Sua duração foi de 23 ms,

Figura 8 – Forma de onda da sílaba tônica da palavra-alvo *craque* produzida por um menino



Fonte: elaborado pelo autor.

Legenda: a elipse pontilhada indica o VOT da oclusiva velar surda, cuja duração foi de 23ms, na sílaba tônica da palavra-alvo *craque* produzida por um menino. Nesta produção, não houve a realização do *tap*. Assim, na figura, tem-se a sílaba [ka] em vez de [k r a].

Como se pode notar a partir das figuras 3 e 4, em que se comparam dados de um menino e de uma menina, consta que o VOT da oclusiva velar produzida pela menina foi, aproximadamente, três vezes maior que o VOT produzido pelo menino, cujos valores brutos foram, respectivamente, 75 ms e 23 ms.

A razão para a diferença entre as médias dos VOTs da consoante velar [k] de meninos e meninas pode estar associada ao volume dos tratos vocais, como já aludido antes, os quais tendem a ser menores no sexo feminino e, portanto, produzem pressões maiores que atrasam o início de vozeamento, além de solturas mais abruptas e com mais energia. (BARBOSA; MADUREIRA, 2015).

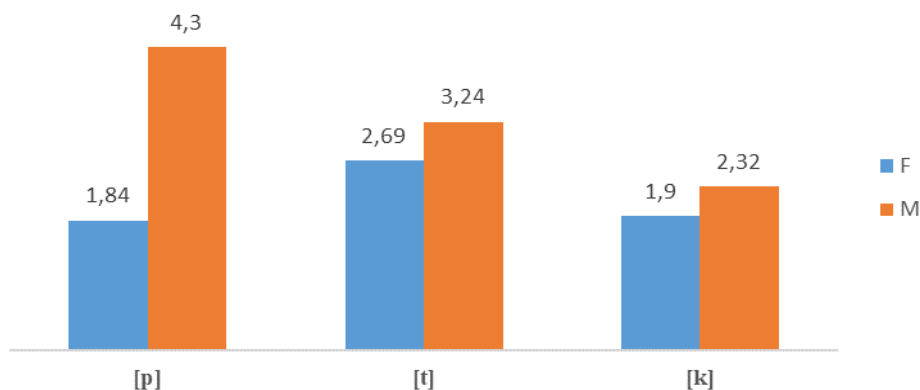
No Português Brasileiro, pelo menos até nossa última revisão, ainda não se encontram trabalhos nesse sentido, isto é, referentes aos efeitos do sexo no VOT produzido por crianças. No geral, como já dito anteriormente, encontram-se apenas trabalhos com produções de falantes adultos, tais como Celeste e Teixeira (2009) e Barbosa e Madureira (2015), os quais foram resenhados em outro capítulo desta tese.

Por outro lado, pode-se dizer, em síntese, que nossos achados sobre os valores do VOT de [k], os quais foram significativamente maiores para as meninas e menores para os meninos, convergem com os de um trabalho já apresentado em outro capítulo desta tese, qual seja: o de Whiteside, Henry e Dobbin (2004), especialmente no que diz respeito ao grupo de crianças na faixa etária entre 5 e 9 anos de idade, a qual, aliás, coincide com a mesma faixa abrangida em nossa pesquisa.

Tendo em vista toda discussão acima, convém dizer que nossos achados, embora não encerrem definitivamente a questão, abrem, pelo menos, caminho para o desenvolvimento de futuras pesquisas a este respeito no Português Brasileiro.

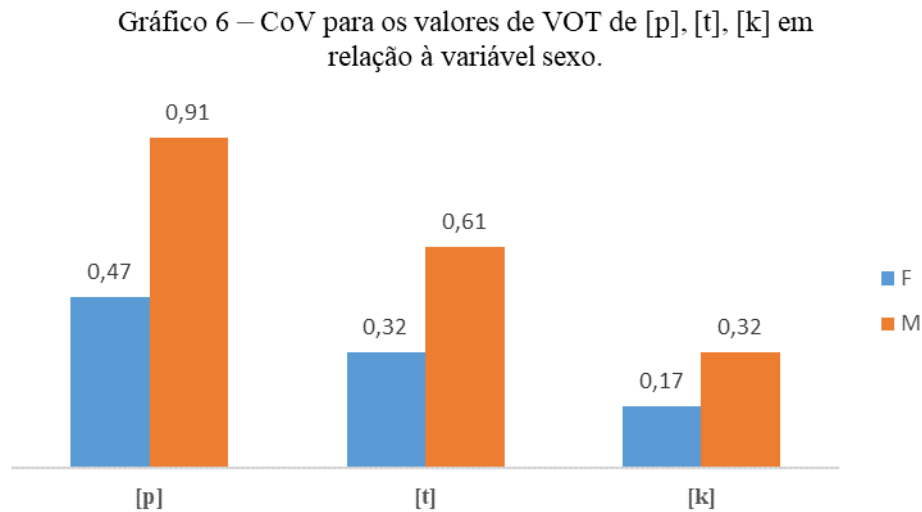
Na sequência, a seguir, destacamos gráficos com medidas de dispersão, desvio padrão e coeficientes de variação, referentes aos grupos de meninos e meninas.

Gráfico 5 – Desvio Padrão (DP) dos valores de VOT para [p], [t], [k] em relação à variável sexo



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

No gráfico acima, apresentam-se os desvios padrão dos valores do VOT produzidos para [p], [t] e [k], pelos meninos (M) e pelas meninas (F).



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

No gráfico acima, mostram-se os coeficientes de variação (CoV) para os valores do VOT de [p], [t] e [k], produzidos pelos meninos (M) e pelas meninas (F).

A seguir, vejamos a tabela 2, que sintetiza todos os valores das médias, dos desvios padrão e dos coeficientes de variância produzidos pelos grupos de meninos e meninas.

Tabela 2 - Média, desvio padrão (DesvP), coeficiente de variação (CoV) dos valores de VOT de [p], [t], [k] produzidos pelos grupos de meninos (M) e meninas (F)

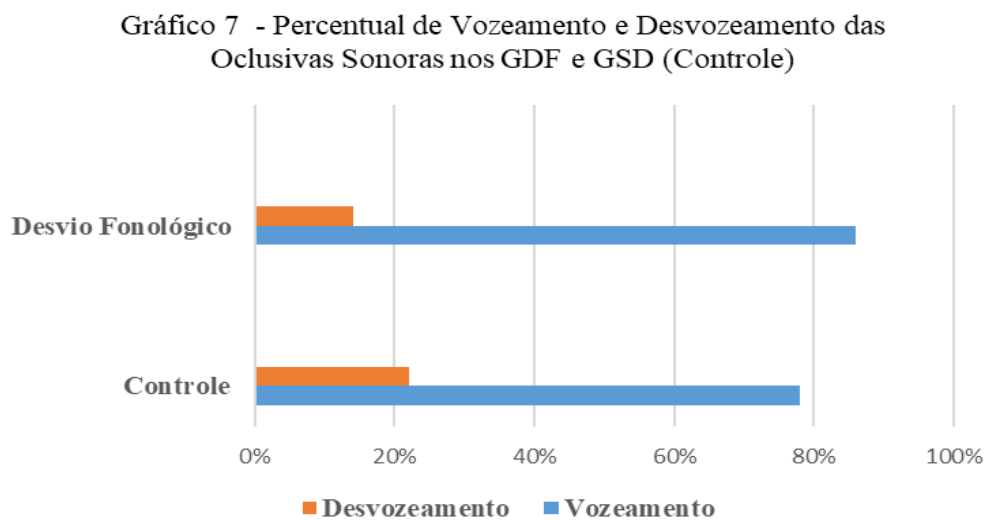
<b>Grupos</b>		
[p]	<b>M</b>	<b>F</b>
<b>Média</b>	<b>4,73</b>	<b>3,89</b>
<b>DesvP</b>	<b>4,3</b>	<b>1,84</b>
<b>CV</b>	<b>0,91</b>	<b>0,47</b>
[t]	<b>M</b>	<b>F</b>
<b>Média</b>	<b>5,27</b>	<b>8,44</b>
<b>DesvP</b>	<b>3,24</b>	<b>2,69</b>
<b>CoV</b>	<b>0,61</b>	<b>0,32</b>
[k]	<b>M</b>	<b>F</b>
<b>Média</b>	<b>7,26</b>	<b>11,22</b>
<b>DesvP</b>	<b>2,32</b>	<b>1,9</b>
<b>CoV</b>	<b>0,32</b>	<b>0,17</b>

Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

### 6.1.2 Análise do VOT das oclusivas sonoras

Quanto às oclusivas sonoras, cada um dos participantes, como já dito anteriormente, produziu três vezes cada uma das três palavras-alvo *braço*, *dragão* e *grade*, no teste de nomeação de figuras. Considerando as produções desses participantes em ambos os grupos, GDF e GSD (controle), obteve-se, ao todo, um corpus de 72 palavras, a partir do qual se analisaram as oclusivas sonoras.

Daí, buscou-se investigar, nas produções dos participantes de ambos os grupos, detalhes fonéticos atinentes ao vozeamento e ao desvozeamento das oclusivas sonoras. Tendo isso em vista, considere-se o gráfico logo a seguir.



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Considerando todas as produções das oclusivas sonoras, tem-se que cerca de 18% delas se realizaram com desvozeamento. O grupo de crianças com desvio fonológico (GDF) realizou desvozeamento em cerca de 14% de suas produções; já o grupo sem o desvio (GSD) atingiu o percentual de 22%.

Apesar dessas diferenças de percentuais de desvozeamento nesses grupos, constatou-se, a partir de um teste Qui-Quadrado, que eles não se distinguem significativamente quanto ao desvozeamento das oclusivas ( $p = 0,358$ ).

Contudo, essas produções podem sugerir certa dificuldade dos participantes de ambos os grupos na produção das oclusivas sonoras. Por outro lado, podem ainda sugerir certa compatibilidade com um aspecto importante das oclusivas sonoras do português europeu. Este aspecto “[...] é o desvozeamento total ou parcial delas, que pode atingir valores superiores a

20% das produções desses segmentos [...]” (LOUSADA, 2006 *apud* BARBOSA; MADUREIRA, 2015, p. 363).

Em contrapartida, este desvozeamento produzido pelos participantes da pesquisa também podem sugerir que os valores dos VOTs parecem não ser tão apropriados para produção das oclusivas sonoras, quer pelas crianças com desvio fonológico, quer pelas crianças sem o desvio.

Nesse sentido, ainda, cumpre ressaltar que Barroco *et al.* (2007) já haviam aludido a tal inadequação dos valores do VOT das sonoras considerando o trabalho de Bortolini *et al.* (1995), os quais, por seu turno, investigaram num estudo longitudinal um grupo controle (N=7), com bebês normais, e um experimental (N=7), com bebês pré-termo, em diferentes níveis de idade. Os resultados mostraram que os sujeitos são mais hábeis na aquisição de valores de VOT para consoantes surdas do que para consoantes sonoras. Segundo Bortolini *et al.* (1995), isso “[...] pode estar associado com o crescimento do controle neuromuscular e com uma atividade muscular mais complexa, necessários para a manutenção do vozeamento durante a oclusão, especialmente nas oclusivas velares<sup>8</sup>.”

### **6.1.3 Análise da fricativa labiodental surda**

Apresenta-se aqui, inicialmente, a análise da duração relativa das fricativas labiodentais surdas produzidas pelos participantes de ambos os grupos. Em seguida, procede-se à análise dos picos espectrais dessas fricativas mesmas.

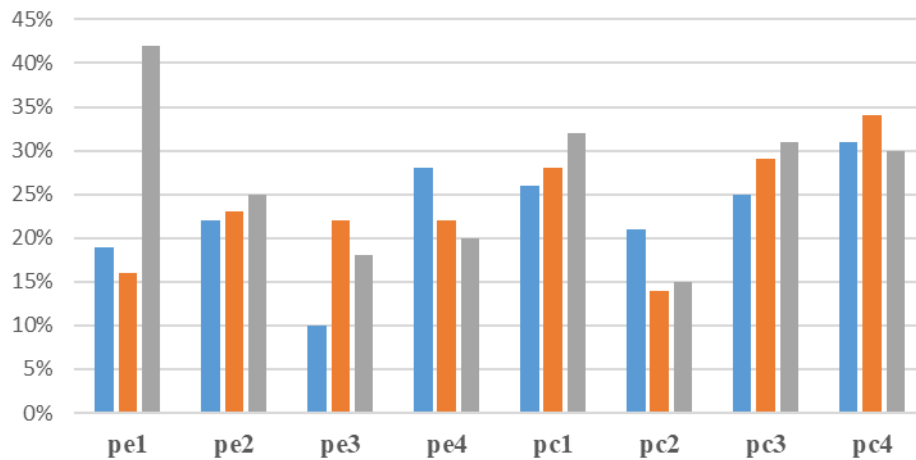
A medida da duração relativa seguiu procedimento semelhante ao cálculo do VOT relativo, descrito anteriormente nas análises anteriores das oclusivas. Assim, mediu-se, inicialmente, o valor bruto da duração da fricativa, em milissegundos; depois, mediu-se o a duração da palavra inteira e, na sequência, dividia-se a duração da fricativa pela duração da palavra inteira, multiplicando esse valor por 100.

Vejamos, no gráfico, a seguir, como se comportaram os grupos GDF e GSDF em relação à duração relativa da fricativa labiodental surda [f].

---

<sup>8</sup> A tradução deste trecho está baseada na de Barroco *et al.* (2007, p. 155). Já a versão original, a qual se encontra no trabalho de Bortolini *et al.* (2006) é esta: “This difficulty may be related to the increased neuromuscular control and complex muscly activity necessary for mantaing voicing during the closure, especially for velar stops.” Aqui interessa destacar, ainda, a frase seguinte do trecho original, que não é mencionada por Barroco *et al.*: “It is importante to recognize the possibility that increased variability in preterm children may be related to some neuromuscular immaturity.”

Gráfico 8 - Durações relativas da fricativa produzidas pelos participantes dos dois grupos



Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Legenda: as letras “pe” correspondem aos participantes do grupo experimental, ou seja, do grupo com desvio fonológico (GDF), e as letras “pc” são códigos referentes aos participantes do grupo controle, isto é, o grupo sem o desvio (GSDF). As cores distintas correspondem a 3 repetições distintas da duração da fricativa labiodental surda produzidas por cada participante.

Como se pode observar no gráfico, as durações relativas da fricativa produzida por quase todos os participantes de ambos os grupos situam-se num intervalo de 14% a 34%, com exceção das produções do participante 1 do grupo com desvio fonológico (pe1), que ultrapassa esse intervalo atingindo 42% em uma de suas produções, e o participante 3 deste mesmo grupo (pe3), que em pelo menos uma de suas fricativas produziu uma duração relativa de 10%, ficando abaixo do reportado intervalo. Ainda em relação a estes participantes, consta que pe1 (19%, 16%, 42%) apresentou duração relativa mínima de 16% e máxima de 42%, com variação de 26% entre a máxima e a mínima; enquanto pe3 (10%, 22%, 18%) produziu mínima de 10% e máxima de 22%, com variação de 12%.

Entretanto, no mesmo grupo destes participantes com desvio fonológico, os outros dois tiveram pequenas variações em suas produções. Observe-se que pe2 (22%, 23%, 25%) apresentou duração relativa mínima de 22% e máxima de 25%, com uma pequena variação de 3%. O participante pe4 (28%, 22%, 20%), por seu turno, teve mínima de 20% e máxima de 28%, com variação de 8%.

Já no grupo sem desvio fonológico, o GSDF, tem-se que o participante 1 do grupo sem desvio fonológico (pc1) (26%, 28%, 32%) apresentou mínima de 26% e máxima de 32%, com variação entre a mínima e a máxima em torno de 6%; o participante 2 (pc2) deste mesmo grupo (21%, 14%, 15%), mínima de 15% e máxima 21, com variação de 6%; o pc3 (25%, 29%, 31%), mínima de 25% e máxima 31%, com variação de 6%; e pc4 (31%, 34%, 30%),

mínima de 30% e máxima de 34%, com variação de 4%. Note-se que tal variação neste grupo oscilou em torno de 4% a 6%.

Diferentemente deste grupo GSDF, consta que, no grupo com desvio fonológico GDF, como se viu antes, ocorreu variação mais alta entre a duração relativa mínima e máxima da fricativa: 3% a 26%. Convém lembrar que os participantes pe1, pe3 e pe4, do GDF, se afastaram muito da faixa de 4% a 6%, a qual foi alcançada pelo outro grupo (GSDF). Reitere-se que as respectivas variações de pe1, pe3 e pe4, entre a mínima e máxima, foram, respectivamente: 26%, 12% e 8%.

Tendo em vista todas essas considerações, pode-se dizer que o grupo com desvio fonológico GDF apresenta maior variabilidade de duração da fricativa na palavra-alvo *fraco*, ao ser comparado com o grupo sem o desvio GSDF. Essa variabilidade pode estar associada a dificuldades específicas na articulação da fricativa labiodental por crianças com desvio fonológico.

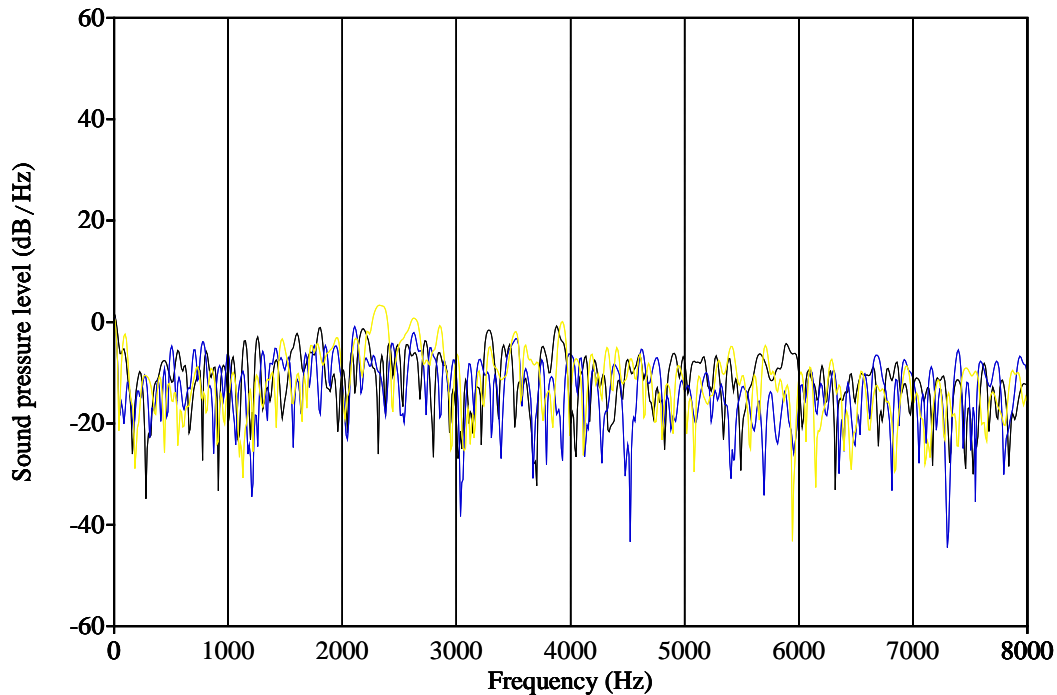
#### **6.1.4 Análise espectral da fricativa labiodental surda**

Nesta análise, extraiu-se o espectro das fricativas atinentes à sílaba tônica da palavra-alvo *fraco*. Com isso buscou-se investigar a configuração espectral da fricativa labiodental surda, por meio da análise FFT, orientada a partir do ponto médio desta consoante. Nesse sentido, ainda, buscou-se também analisar se as produções dos grupos GDF e GSDF se distinguem quanto a tal configuração.

Como cada participante produziu três fricativas, utilizamos três cores distintas para possibilitar a identificação de cada um dos três respectivos espectros dessas fricativas mesmas. Mas, aqui, omitem-se as produções de todos os participantes, e se apresentam apenas as três produções de um único participante, para ilustrar o padrão identificado na configuração espectral da fricativa labiodental surda, que, em linhas gerais, serve para representar o padrão seguido, relativamente, pela maioria dos participantes de ambos os grupos.

Desse modo, a configuração espectral da fricativa labiodental surda derivada da produção da palavra-alvo *fraco* seguiu um padrão de horizontalidade, tal como o que se apresenta no gráfico 9, a seguir.

Gráfico 9 – Espectros, com janela de 25 ms, das três fricativas labiodentais produzidas pelo participante 1 do grupo sem desvio fonológico



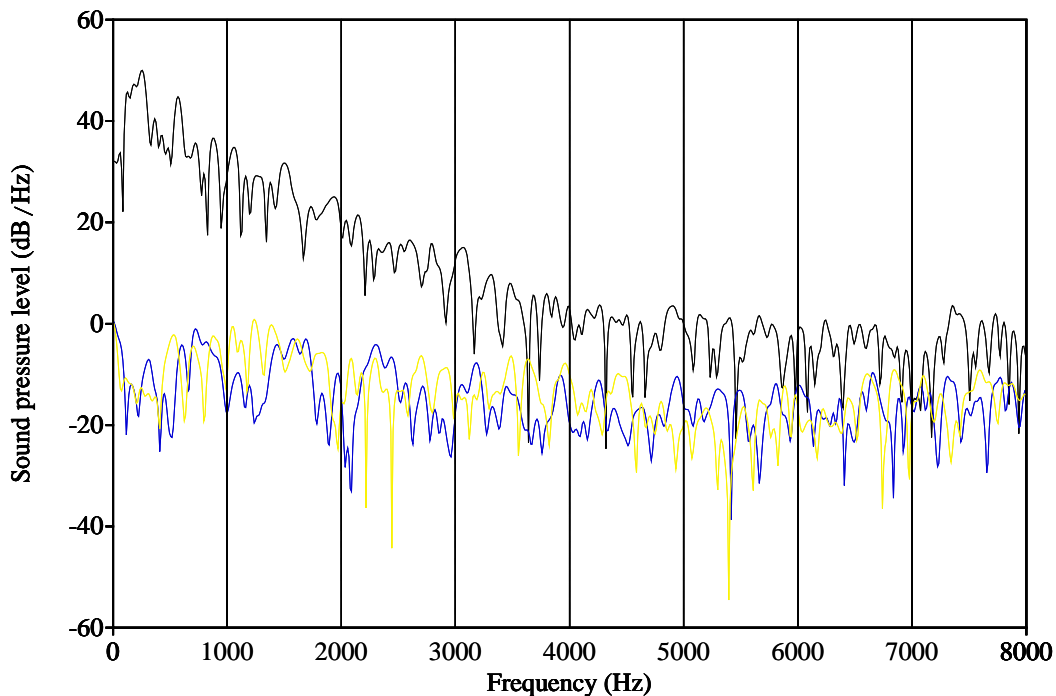
Fonte: Elaborado pelo autor a partir dos dados desta pesquisa.

Veja-se, aí, no gráfico 9, a horizontalidade na composição espectral. Note-se que os formatos dos espectros são similares. Quase todas as produções dos participantes seguiram padrão horizontal semelhante. Apenas uma participante, no caso, a participante 3 do grupo com desvio fonológico, se distanciou desse padrão, visto que em vez de produzir a fricativa labiodental surda na palavra-alvo *fraco*, produziu a fricativa alveolar surda [s], não realizando, portanto, *cluster* com *tap*. Disso resultou, no lugar da palavra *fraco*, outra palavra, que pode ser transcrita, aqui, mais ou menos, da seguinte forma: [ˈsa.kʊ].

Para se ter melhor idéia disso, observem-se, na figura a seguir, os três espectros das fricativas produzidas por esta participante. O preto corresponde à produção da fricativa alveolar surda [s], e os outros dois, o azul e o amarelo, a duas respectivas produções da fricativa labiodental surda, extraídas da palavra-alvo *fraco*. Repare-se que a composição do espectro preto se distingue nitidamente da composição espectral dos outros dois (o azul e o amarelo) em termos de horizontalidade.



Gráfico 10 – Espectros, com janela de 30 ms, das três fricativas produzidas por participante 3 do grupo com desvio fonológico



Fonte: Elaborado pelo autor a partir dos dados da pesquisa.

## 6.2 Análise do *tap* no onset complexo

Nesta análise, se apresentam, inicialmente, os resultados pautados em duas categorias excludentes entre si, a saber: a realização e a não-realização do *tap* em onset complexo. A categoria “realização” do *tap* abrange tanto as realizações do *tap* quanto a de suas variantes em ataque complexo. Os vestígios, ou pistas acústicas, por meio dos quais se consideraram a realização do sinal do *tap* foram, basicamente: i) a elevação de F2; ii) a queda brusca de energia de curta duração cujo clareamento espectrográfico é correlato acústico da batida da ponta da língua. Já a categoria “não-realização” abrange apenas aquelas realizações nas quais não se apresentam vestígios do sinal do *tap* nos *clusters* investigados.

Em seguida, na análise final, se evidenciam e se interpretam, a partir desses *clusters* mesmos, determinadas formas de ondas e espectrogramas atinentes às produções acústicas dos grupos GDF e GSDF.

Tendo em vista essas considerações, começemos, então, pela exposição dos resultados da análise referente às categorias realização e não-realização do *tap* em onset complexo.. No grupo de participantes com desvio fonológico (GDF), consta o seguinte: nos *clusters* constituídos de oclusivas surdas seguidas de *tap*, /pr, tr, kr/, bem como nos de oclusivas sonoras seguidas de *tap*, /br, dr, gr/, e no de fricativa labiodental surda seguida de *tap*, /fr/, a

categoria não-realização apresentou percentual de 100% em todos os contextos. Quer dizer, os participantes com desvio fonológico não realizaram o *tap* em nenhum dos casos analisados.

Já no grupo de participantes sem o desvio fonológico (GSDF) foi diferente. Nos *clusters* constituídos de oclusivas surdas seguidas de *tap*, /pr, tr, kr/, o participante 1 obteve 12% de não-realizações e 88% de realizações; o participante 2 - 33% de não-realizações e 77% de realizações; o participante 3 - 34% de não-realizações e 66% de realizações e o participante 4 - 12% de não-realizações e 88% de realizações. Nos *clusters* constituídos de oclusivas sonoras seguidas de *tap*, /br, dr, gr/, o participante 1 obteve 100% de realizações; o participante 2 - 33% de não-realizações e 77% de realizações; o participante 3 - 45% de não-realizações e 55% de realizações; o participante 4 - 12% de não-realizações e 88% de realizações. No *cluster* constituído de fricativa labiodental surda seguida de *tap*, /fr/ todos os participantes do GSDF obtiveram 100% de realizações.

Esses resultados corroboram nossa hipótese básica, segundo a qual as produções do *tap* em onset complexo dos grupos GDF e GSDF se distinguem significativamente. Frise-se, aí, que o percentual de 100%, atingido pelo GDF, na categoria não-realizações do *tap*, nos mostrou que seria desnecessário proceder a algum teste estatístico de diferenças de proporções entre esses grupos para confirmar a reportada hipótese.

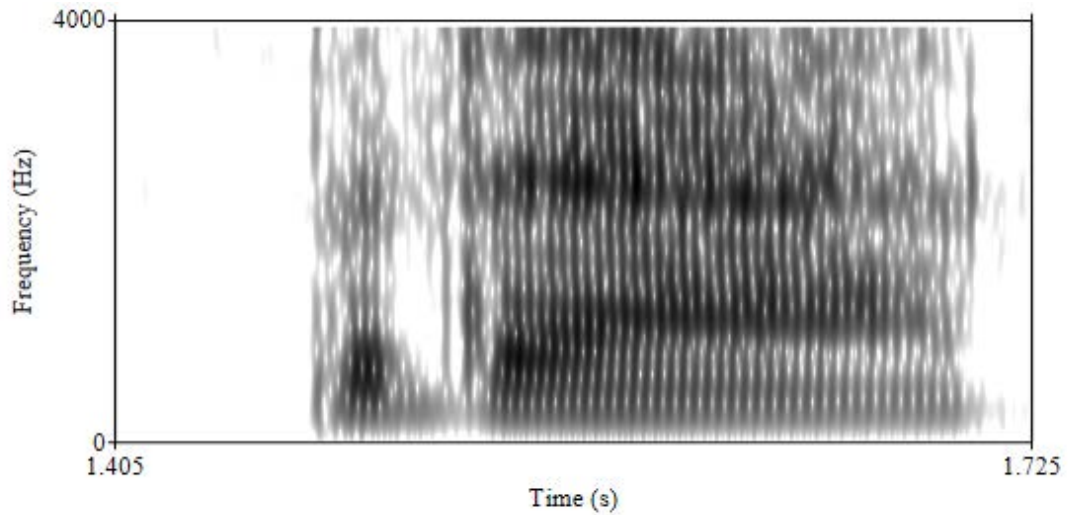
Ainda nesta análise, constatou-se que foram acertadas nossas previsões, segundo as quais os participantes com desvio fonológico tenderiam a não realizar o *tap* em onset complexo, enquanto os participantes sem o desvio tenderiam a realizá-lo.

### **6.2.1 Análise final**

Nesta análise final, como já dito antes, se evidenciam e se interpretam determinadas formas de ondas e espectrogramas atinentes às produções do *tap* em onset complexo dos grupos GDF e GSDF. Aqui, tais produções/evidências são apresentadas de forma alternada: primeiro, apresentam-se duas produções de dois participantes do GSDF; depois, as de dois participantes do GDF e, na sequência, se apresentam novas produções de outros participantes do GSDF seguidas por outras de participantes do GDF e assim por diante.

Começando pelo participante 3 do grupo de crianças sem desvio fonológico (GSDF), observe-se, na figura 9, logo a seguir, o espectrograma da sílaba tônica *pra* que se extraiu da palavra *prato* produzida por tal participante. O *tap*, nesta sílaba, teve duração de 37 ms. Ele é vozeado e está fricativizado.

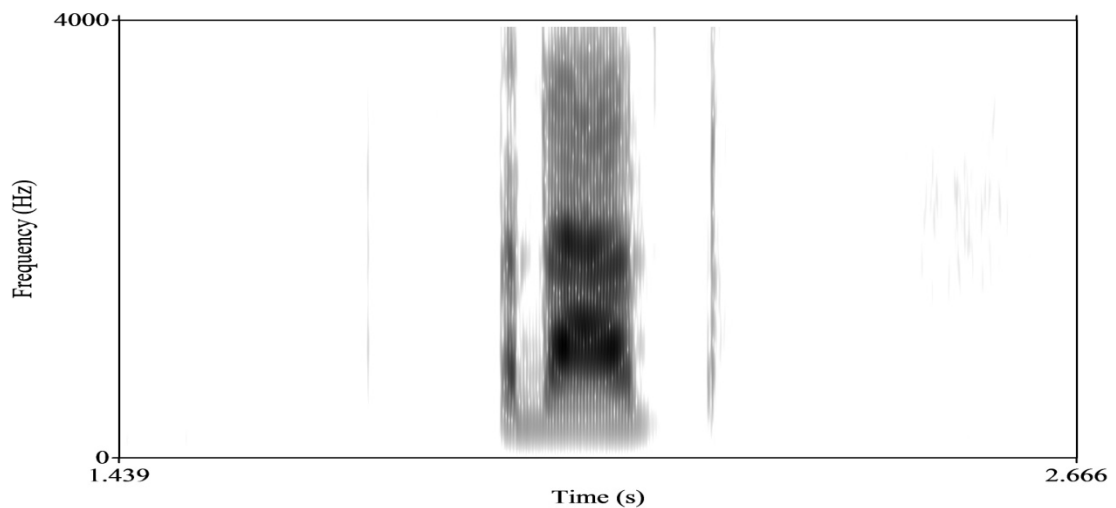
Figura 9 – Espectrograma de banda larga da sílaba *pra* da palavra *prato* produzida pelo participante 3 do GSDF



Fonte: Elaborada pelo autor.

Já na figura a seguir, apresenta-se o espectrograma de banda larga da palavra *prato* produzida pelo participante 2 do GSDF. Observe-se que o *tap* é vozeado e apresenta nítido clareamento nas frequências acima de F1. Quanto à duração, ele tem 45 ms.

Figura 10 - Espectrograma de banda larga da palavra *prato* produzida pelo participante 2 do GSDF

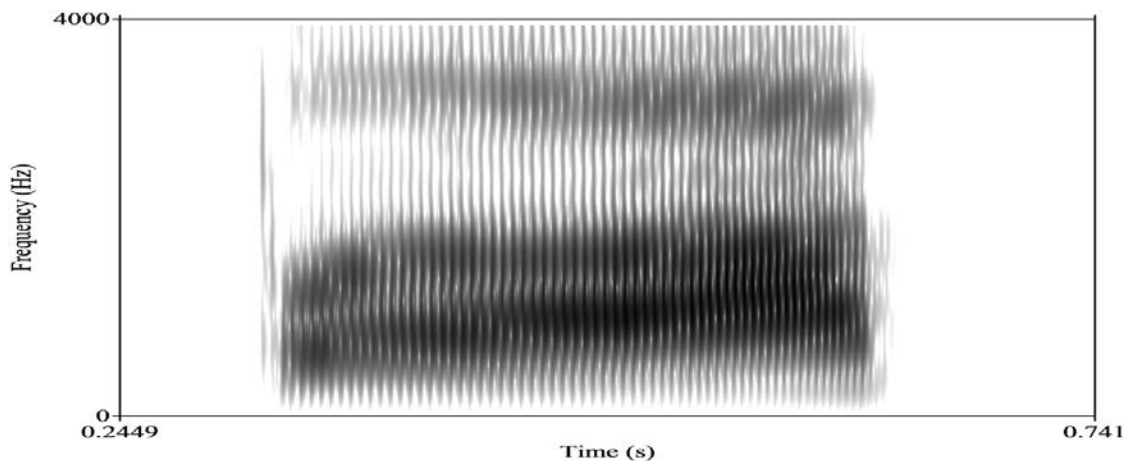


Fonte: Elaborado pelo autor.

Como se viu, nas figuras 9 e 10, apresentam-se evidências de produções do *tap* em onset complexo, de participantes do grupo sem desvio fonológico (GSDF). Agora vejamos produções de dois participantes do grupo com desvio fonológico (GDF).

Na figura 11, tem-se o espectrograma da sílaba tônica da palavra-alvo *prato* produzida pelo participante 1 do grupo com desvio fonológico GDF. Note-se que o *tap* não é realizado por este participante. Não se encontram aí vestígios do sinal do *tap*, tais como o clareamento espectrográfico, que é correlato acústico da batida da ponta da língua, ou a elevação de F2. Observe-se, ainda, que, após a explosão da oclusiva bilabial surda [p], tem-se início uma vogal tônica [a], cuja duração é de 298 ms e que ocorre sem quaisquer interrupções.

Figura 11 – Espectrograma de banda larga da sílaba tônica *pra* extraída da palavra-alvo *prato* produzida pelo participante 1 do GDF

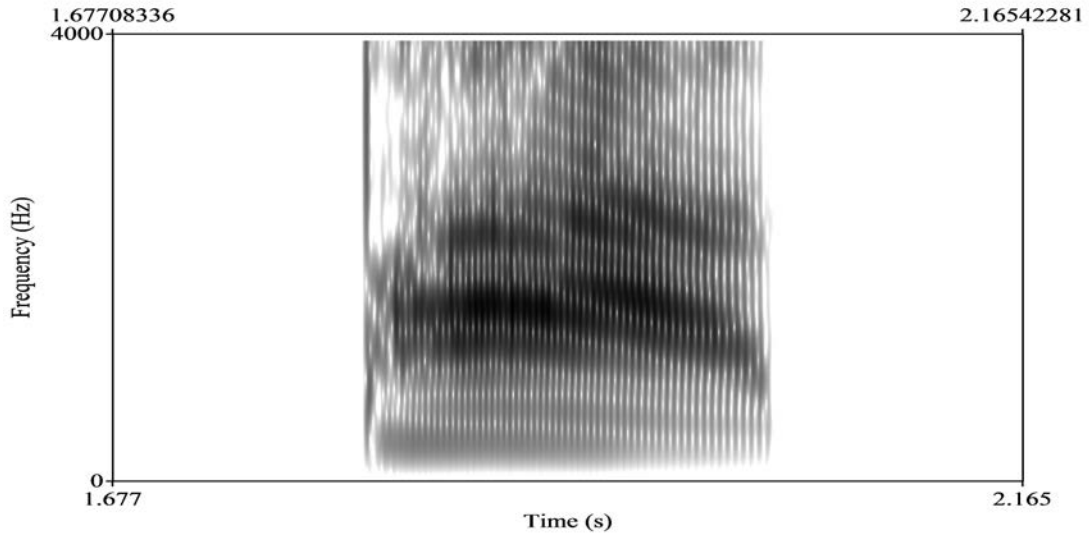


Fonte: Elaborado pelo autor.

Legenda: neste caso, como não se realizou o *tap* em onset complexo, o que se tem aí, no espectrograma, é, na verdade, uma sílaba tônica [pa] em vez de [pra].

Já na figura 12, apresenta-se o espectrograma da sílaba tônica da palavra *prato* produzida pela participante 3, que também é do grupo com desvio fonológico. Note-se que, tal qual o participante 1 de seu mesmo grupo, a participante 3 também não realiza o *tap* em onset complexo. Como se pode observar na figura, não consta o clareamento correspondente ao *tap*, o qual deveria entrecortar a vogal nuclear, nem elevação de F2, tampouco qualquer outra pista acústica que o caracterize. Veja-se que após a soltura da oclusiva surda já se tem, logo em seguida, o início da trajectória dos formantes da vogal tônica da palavra-alvo *prato*, sem vestígios do *tap*. Aliás, observe-se que tanto F2 quanto F3 seguem diretamente suas respectivas trajectórias, sem alterações que possam indicar a realização do *tap*.

Figura 12 – Espectrograma de banda larga da sílaba tônica *pra* da palavra-alvo *prato* produzida pela participante 3 do GDF

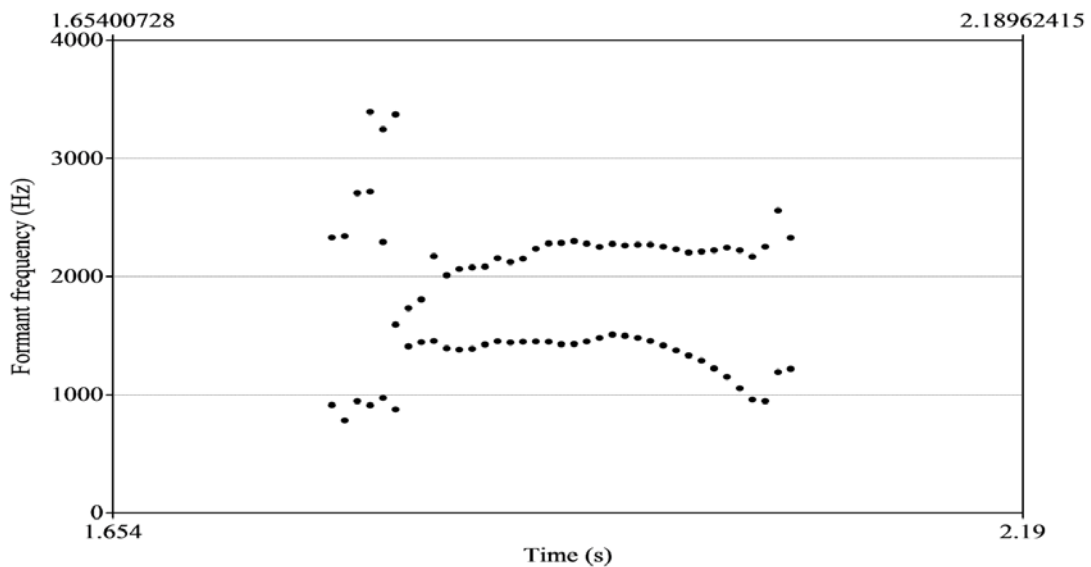


Fonte: Elaborado pelo autor.

Legenda: neste caso, como não se realizou o *tap* em onset complexo, o que se tem aí, no espectrograma, é, na verdade, uma sílaba tônica [pa] em vez de [pra], da mesma forma que se viu na figura anterior.

Ainda em relação a este mesmo dado produzido pelo participante 3 do GDF, mostram-se, na figura 13 as trajetórias de F2 e F3 referentes à vogal tônica, ou nuclear, em pauta.

Figura 13 – Trajetória de F2 e F3 referentes à vogal tônica da sílaba tônica da palavra-alvo *prato* produzida pela participante 3 do GDF



Fonte: Elaborado pelo autor.

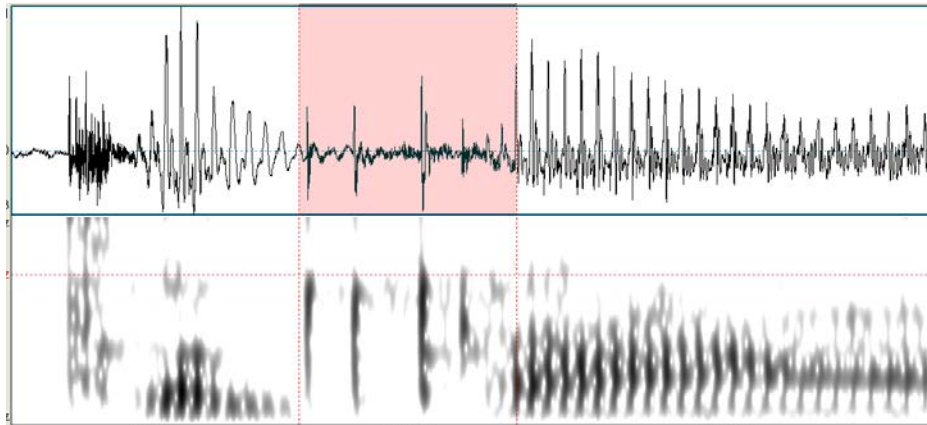
Legenda: neste caso, como não se realizou o *tap* em onset complexo, o que se tem aí, no espectrograma, é, na verdade, uma sílaba tônica [pa] em vez de [pra].

Observe-se na figura 13 que após a soltura da oclusiva surda as trajetórias de F2 e F3 vão se distanciando uma da outra no intervalo de tempo em direção à sílaba seguinte e isso ocorre sem vestígios do *tap* interferindo nessas trajetórias, como se pode notar. Destaque-se aí, sobretudo, que não há elevação de F2, a qual, em certo aspecto, poderia ser tomada como vestígio desta consoante.

Até aqui se evidenciaram as realizações e não-realizações do *tap* em onset complexo referentes à palavra *prato*, produzidas por participantes de ambos os grupos, GDF e GSDF. Como se viu, somente os participantes do grupo sem desvio fonológico (GSDF) realizaram o *tap* em onset complexo. Agora passemos para análise da palavra-alvo *trave*.

Tendo isso em vista, observem-se na figura 14 a forma de onda e o espectrograma de banda larga correspondente à realização do *tap* em trecho da palavra *trave* produzida pelo participante 4 do grupo sem desvio fonológico (GSDF). O segmento não está vozeado e vê-se claramente o elemento vocálico que emerge após a soltura da oclusiva precedente.

Figura 14 – Forma de onda e espectrograma do *tap* no onset complexo produzido na sílaba tônica *tra* da palavra-alvo *trave* pelo participante 4 do GSDF.

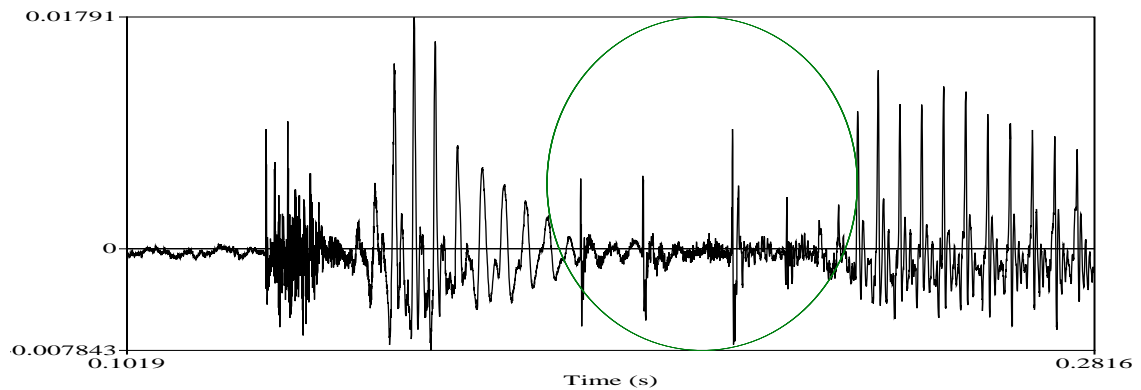


Fonte: Elaborado pelo autor.

Legenda: o trecho selecionado corresponde a, pelo menos, quatro *taps* distintos os quais entrecortam a vogal nuclear (tônica) na palavra-alvo *trave*. Tem-se aí, portanto, um caso especial, visto que se esperaria apenas um único *tap* entrecortando tal vogal.

Já na figura 15, apresentam-se exclusivamente apenas as formas de onda da sílaba tônica *tra* da reportada palavra-alvo e, na sequência, apresenta-se exclusivamente espectrograma disto na figura 16. Vejamos.

Figura 15 – Forma de onda dos “*taps*” em onset complexo produzido na sílaba tônica *tra* da palavra-alvo *trave* pelo participante 4 do GSDF



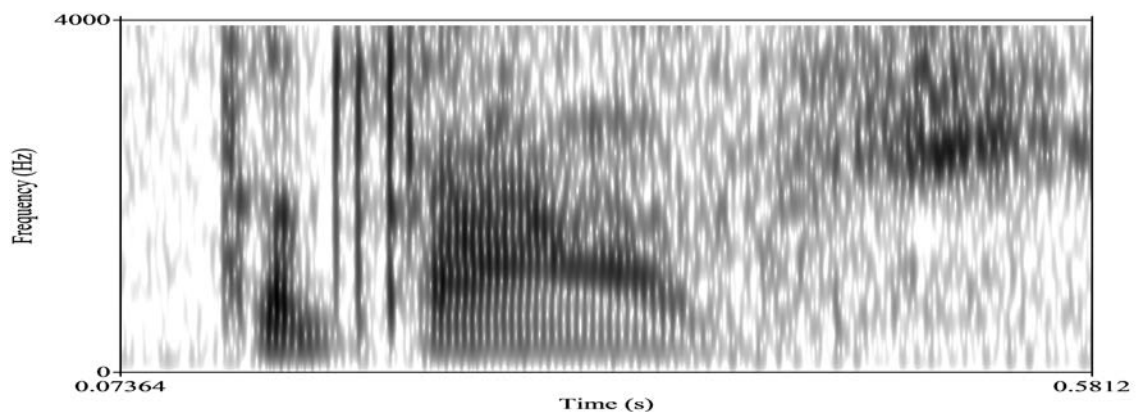
Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada indica os *taps* produzidos em onset complexo na sílaba tônica *tra* da palavra-alvo *trave*.

Como se pode ver na figura 14, no intervalo que vai do fechamento vocálico à abertura vocálica, o qual está destacado pela elipse pontilhada, encontram-se pelo menos quatro *taps* na sílaba tônica *tra* da palavra-alvo *trave*. Esse intervalo dura aproximadamente 51 ms, sendo que do primeiro *tap* para o segundo transcorrem 11 ms; do segundo para o terceiro, 17 ms; do terceiro para o quarto, 11 ms; e do quarto *tap* para a abertura vocálica, 12 ms.

Observe-se agora, no espectrograma da figura 16, que esses *taps* mesmos, além de desvozeados, como dito antes, estão fricativizados.

Figura 16 – Espectrograma do *tap* no onset complexo produzido na palavra-alvo *trave* pelo participante 4 do GSDF



Fonte: Elaborado pelo autor.

Tendo em vista essas considerações, pode-se dizer que o que se tem aí na sílaba tônica *tra* produzida pelo participante 4 do GSDF, se trata, na verdade, de uma vibrante múltipla

alveolar, ou, ainda, como preferimos dizer, um “*tap* múltiplo”, em onset complexo, e não simplesmente um único *tap*, como se esperaria nesse contexto.

A realização de tal segmento – o *tap* múltiplo - no lugar de um único *tap* em onset complexo parece comum em determinados falantes do português brasileiro, crianças ou adultos, independentemente de variações diatópicas, diastráticas, dentre outras.

Chamamos esses falantes de falantes *tépicos*, visto que realizam a vibrante múltipla em onset complexo, ou, por assim dizer, o “*tap* múltiplo” - isto é, vários *taps* - na segunda posição do onset complexo, em vez de realizarem apenas um único *tap* comum, tal qual o ibérico, como geralmente o fazem outros falantes brasileiros nesse contexto.

O uso da vibrante múltipla alveolar - ou *tap* múltiplo - na segunda posição do ataque complexo, tal como o fazem os falantes *tépicos*, parece ser determinado por fatores idiossincráticos. Frise-se, no entanto, que eles não usam apenas essa variante nesse contexto. Também usam outras que, assim como a vibrante múltipla alveolar em onset complexo, refletem um estilo peculiar no que se refere à produção do *tap*, o que não deve ser confundido com um “sintoma de patologia da fala”.

Por conta desse estilo, pode-se ter a impressão de que há nos falantes *tépicos* uma aparente dificuldade, ou ainda, uma aparente falta de controle neuromuscular do articulador ativo – a língua – especialmente quanto à produção do *tap*, já que eles tendem a realizar - com este articulador, mais batidas, ou mais vibrações, do que o necessário para produzir essa consoante, para a qual bastaria uma única batida junto à região alveolar, ou à arcada dentária superior.

De um ponto de vista sociolinguístico, poder-se-ia aventar a hipótese de que os falantes *tépicos* se enquadram num mesmo *idiossocioleto*. Este conceito, por sua vez, não deve ser confundido com o idioleto, o qual é usado para designar o modo de falar característico de um indivíduo, seu modo de pronunciar as palavras etc; e nem com o socioleto, que é usado para designar a variedade linguística de um grupo de falantes que partilham as mesmas características socioculturais (classe socioeconômica, profissão etc.).

Nesse sentido, o *idiossocioleto* se distingue tanto do idioleto quanto do socioleto: do primeiro, porque, embora designe um modo de falar característico de um indivíduo, ou, ainda, um modo característico com que um indivíduo pronuncia determinadas palavras, não se restringe a designar apenas tais aspectos de um único indivíduo, mas se estende a um grupo inteiro de indivíduos que partilham determinadas semelhanças em seus modos de falar e/ou de pronunciar determinadas palavras e sons da língua. E se distingue do socioleto porque os



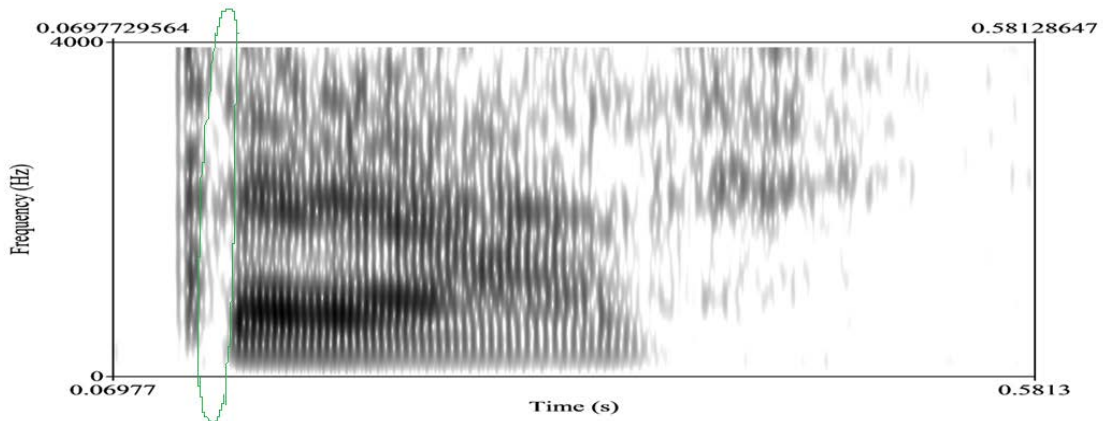
indivíduos do grupo designado por ele - ou seja, pelo *idiossocioleto* - não partilham necessariamente as mesmas características socioculturais.

Assim, o que se entende aqui por *idiossocioleto* se trata de uma variedade linguística que se situa entre o idioleto e o socioleto, mas não se confunde com estes. Designa um grupo de indivíduos que compartilham um modo semelhante de pronunciar determinadas palavras e/ou sons da língua, independentemente de fatores socioculturais e/ou regionais.

Os falantes *tépicos* se enquadram num mesmo *idiossocioleto* porque, como já se disse anteriormente, pronunciam palavras com *tap* de um modo bastante peculiar - bem distinto dos demais falantes brasileiros – quer pelo uso da vibrante múltipla alveolar em onset complexo, quer pelo uso do *tap* mesmo em algumas situações, quer por meio de outras variantes do *tap*. Não obstante, vale frisar que, mediante variação estilística, semelhanças podem emergir entre a pronúncia deles e a de outros falantes.

Agora vejamos a figura 17 na qual se apresenta o espectrograma de banda larga do *tap* na palavra-alvo *trave* produzida pelo participante 1 do GSDF, mesmo grupo do participante abordado anteriormente. Observe-se o clareamento característico do *tap*, o qual está indicado na figura pela elipse verde e tem duração de 18 ms.

Figura 17 – Espectrograma do *tap* em onset complexo produzido na palavra-alvo *trave* pelo participante 1 do GSDF



Fonte: Elaborado pelo autor.

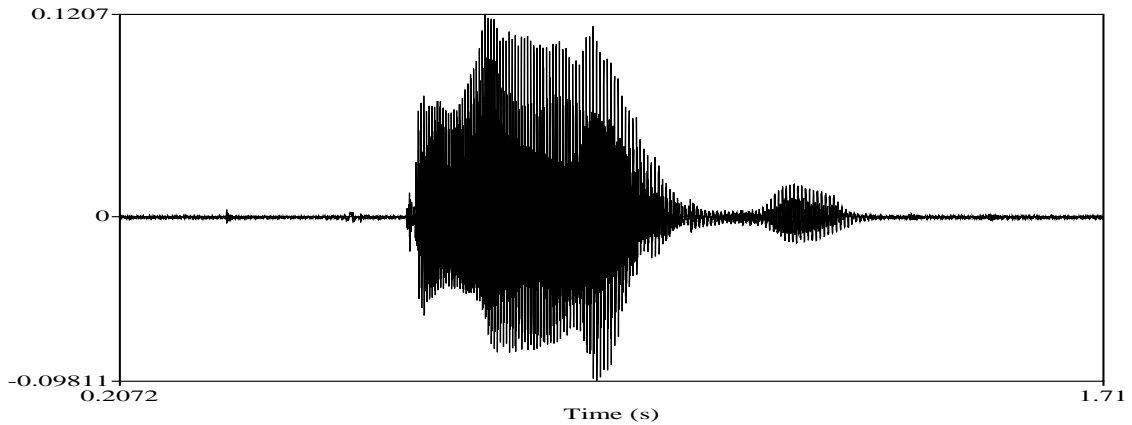
Legenda: a elipse verde indica o *tap*.

Note-se que ambos os participantes deste grupo produziram o *tap*, porém, o que se abordou anteriormente produz um *tap* múltiplo, como já visto, e este último, apenas um único *tap*, como se poderia esperar.

Ainda em relação à palavra-alvo *trave*, apresenta-se, aqui, outra figura, porém referente a uma produção de um participante do grupo GDF. Nesta, entretanto, esse

participante, por sua vez, não produz o *tap* em onset complexo, tampouco uma variante. Como se pode ver em sua produção, na figura 18, logo a seguir, não se encontram vestígios de do *tap* em onset complexo na palavra-alvo *trave*.

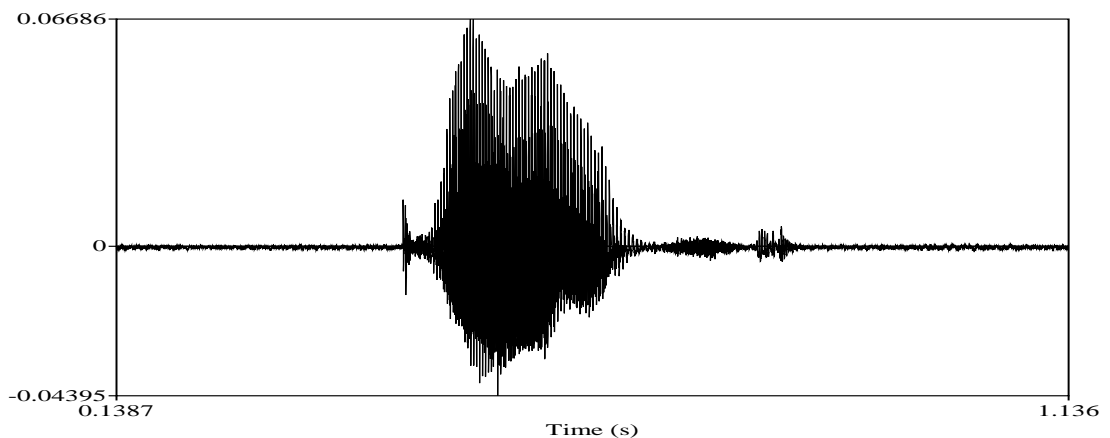
Figura 18 – Forma de onda da palavra-alvo *trave* produzida pelo participante 1 do GDF.



Fonte: Elaborado pelo autor.

Semelhantemente, pode-se constatar, agora na figura 19, a não realização do *tap* na palavra-alvo *trave* na produção de outro participante, no caso, o participante 4 do mesmo grupo do participante anterior, isto é, o grupo com desvio fonológico. Observe-se, nesta figura, que, logo após a soltura da oclusiva, o que se tem é uma vogal tônica seguida pela fricativa labiodental sonora [v] da sílaba seguinte, não havendo, assim, qualquer vestígio do sinal do *tap*.

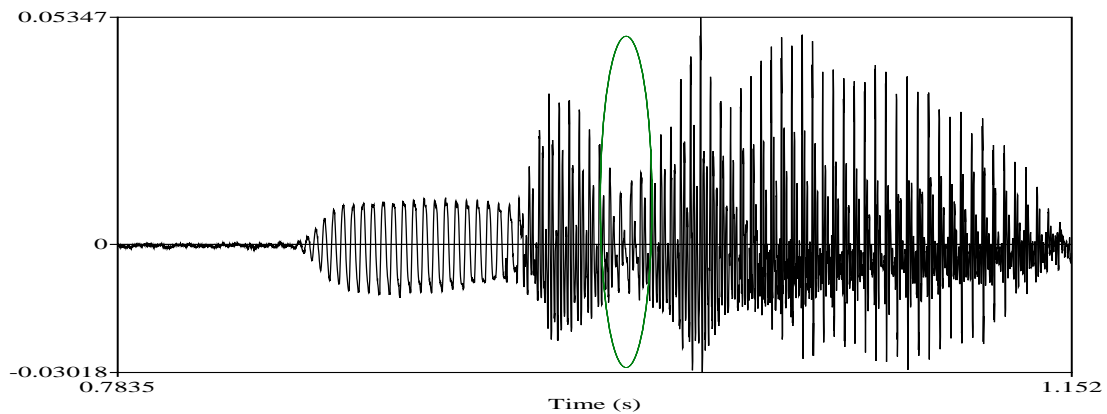
Figura 19 – Forma de onda da palavra-alvo *trave* produzida pela participante 4 do GDF



Fonte: Elaborado pelo autor.

Agora passemos a examinar o *tap* na palavra-alvo *braço*. Veja-se na figura 20, a seguir, a forma de onda da sílaba tônica desta palavra produzida pelo participante 4 do grupo sem desvio fonológico. Note-se que, após soltura da oclusiva sonora [b], emerge uma vogal epentética que é imediatamente seguida pelo *tap*, o qual está destacado pela elipse pontilhada, e é seguido por outra vogal.

Figura 20 – Forma de onda da sílaba tônica *bra* da palavra-alvo *braço* produzida pelo participante 4 do GSDF

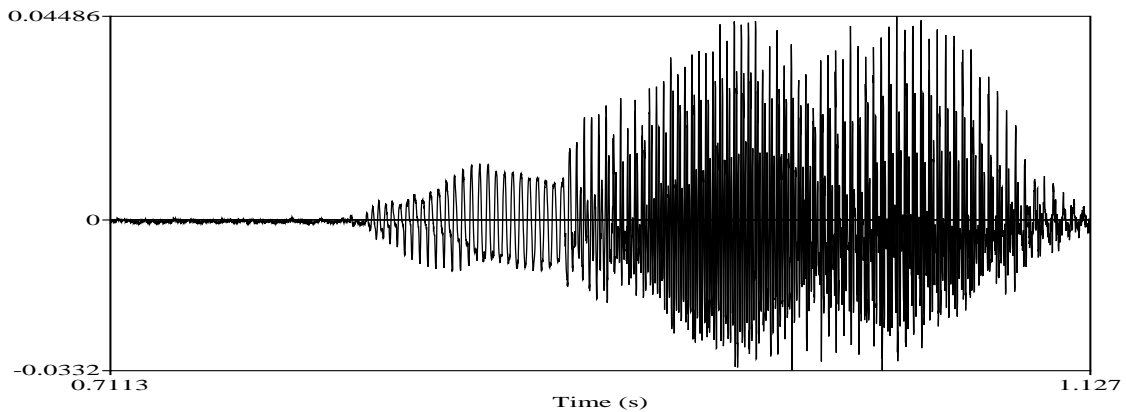


Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada destaca o *tap* na sílaba tônica *bra* da palavra-alvo *braço*.

Agora vejamos na figura 21 a forma de onda da palavra-alvo *braço* produzida por um participante do grupo com desvio fonológico. Observe-se, aí, que não há vestígios do *tap*.

Figura 21 – Forma de onda da sílaba tônica da palavra-alvo *braço* produzida pelo participante 4 do GDF



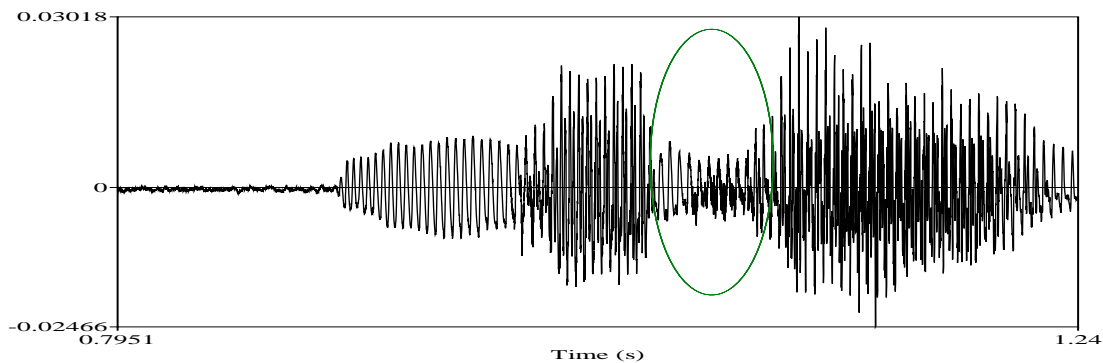
Fonte: Elaborado pelo autor.

Legenda: neste caso, como o participante não realizou o *tap*, o que se tem, nesta figura, é uma sílaba tônica */ba/* em vez de */bra/*.

Como se pode notar, diferentemente do que se viu na figura 20, na figura 21, logo após a soltura da oclusiva bilabial sonora, emerge a vogal tônica [a] da palavra-alvo *braço*, não sendo precedida nem entrecortada pelo *tap*.

Agora vejamos duas produções atinentes à palavra-alvo *dragão*. Observe-se, na figura 22, logo a seguir, a forma de onda da sílaba *dra* da palavra-alvo *dragão* produzida pelo participante 4 do grupo sem desvio fonológico (GSDF). Veja-se que, no início da palavra, e logo após a soltura da oclusiva alveolar sonora emerge uma vogal que é entrecortada pelo *tap*, cuja duração é de 49 ms.

Figura 22 – Forma de onda da sílaba *dra* da palavra-alvo *dragão* produzida pelo participante 4 do GSDF

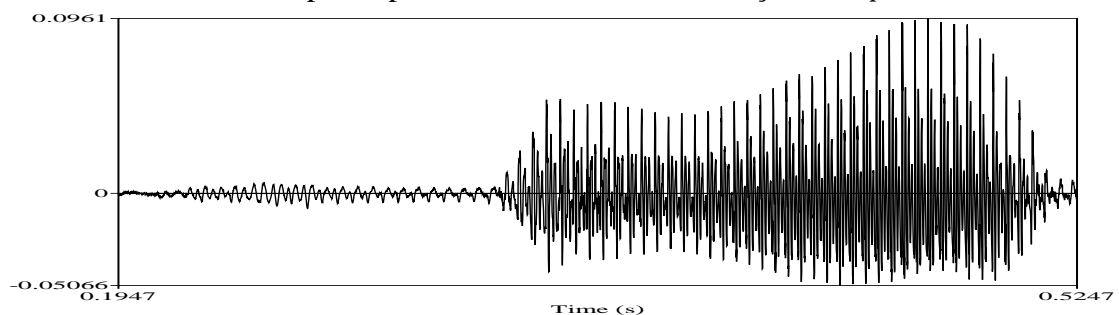


Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada destaca o *tap* na sílaba *dra* da palavra-alvo *dragão*.

Já na figura 23, tem-se a sílaba *dra* da palavra-alvo *dragão* produzida pelo participante 1 do grupo com desvio fonológico (GDF). Note-se que, neste caso, o *tap* não é realizado. No início da palavra *dragão*, logo após a alveolar sonora [d], segue-se a vogal [a] que, por sua vez, não é entrecortada pelo *tap*.

Figura 23 – Forma de onda da sílaba pretônica da palavra-alvo *dragão* produzida pelo participante 1 do GDF, sem realização do *tap*



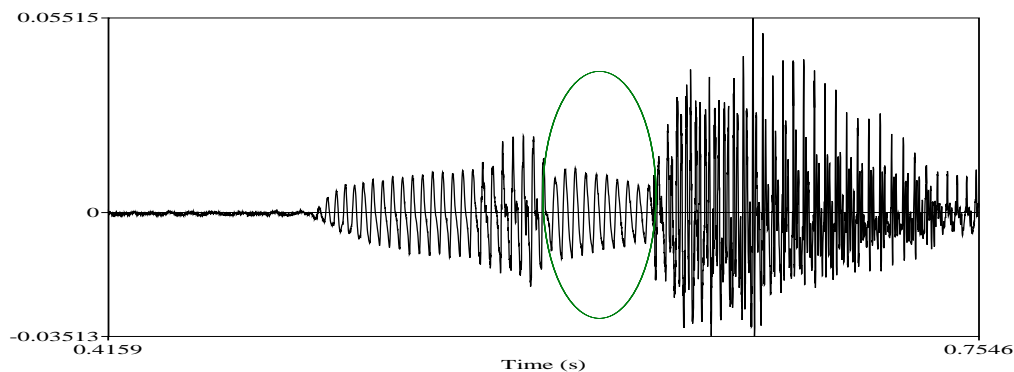
Fonte: Elaborado pelo autor.

Legenda: neste caso, como o participante não realizou o *tap*, o que se tem, nesta figura, é a sílaba *da* em vez de *dra*.

Estabelecendo-se uma comparação entre as figuras 23 e 22, ambas referentes à pretônica da palavra-alvo *dragão*, percebem-se, nelas, com nitidez, diferenças que assinalam a realização do *tap* na produção do participante sem desvio fonológico e a não-realização do *tap* na produção do participante com desvio fonológico.

Vejam agora evidências de realização do *tap* em onset complexo na palavra-alvo *grade*. Observe-se na figura 24 a forma de onda da sílaba tônica da palavra-alvo *grade* produzida pelo participante 4 do grupo sem desvio fonológico.

Figura 24 – Forma de onda da sílaba tônica *gra* da palavra-alvo *grade* produzida pelo participante 4 do GSDF



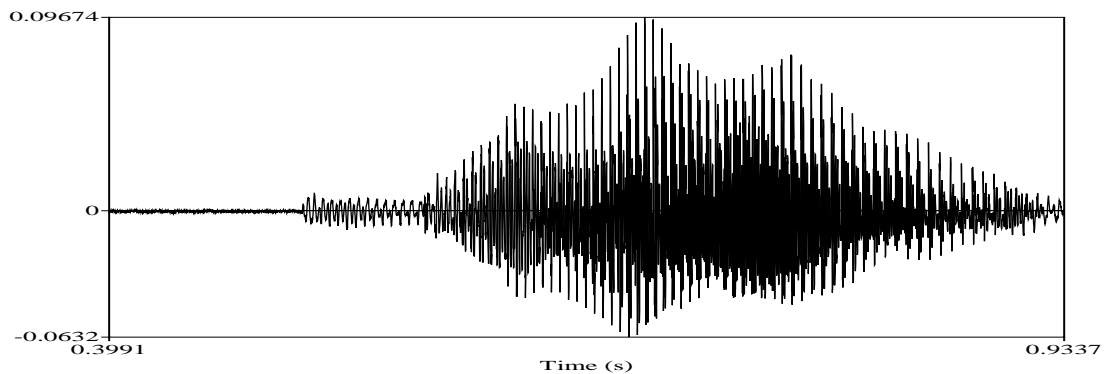
Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada indica o *tap* na sílaba tônica *gra* da palavra-alvo *grade*.

Note-se que, aí, na figura 24, tem-se uma consoante alveolar sonora seguida por uma vogal epentética, e esta, por sua vez, é seguida pelo *tap* (destacado pela elipse pontilhada) o qual tem duração de 44 ms.

Agora observemos a figura 25, na qual também consta a sílaba tônica da palavra-alvo *grade*, porém produzida pelo participante 1 do grupo com desvio fonológico. Observe-se que, neste caso, o *tap* não é realizado.

Figura 25 – Forma de onda da sílaba tônica da palavra-alvo *grade* produzida pelo participante 1 do GDF.

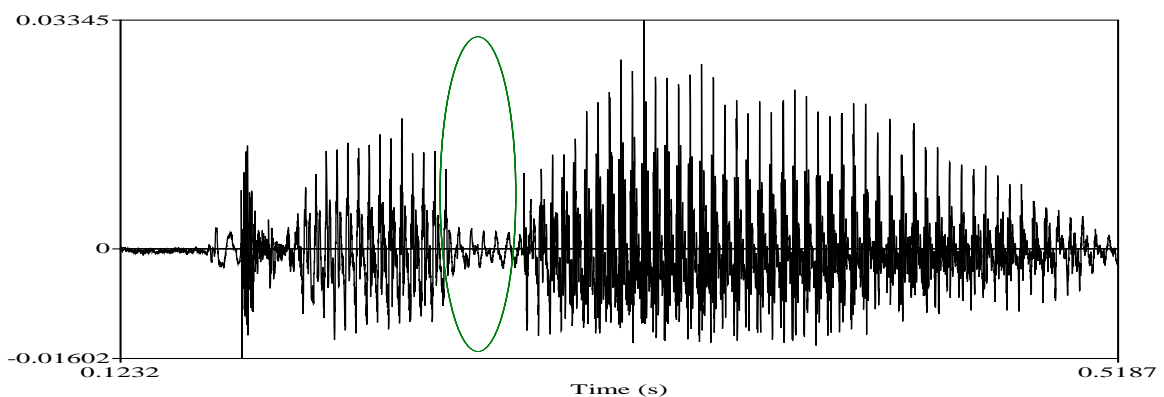


Fonte: Elaborado pelo autor.

Legenda: neste caso, como o participante não realizou o *tap*, o que se tem, nesta figura, é a sílaba *ga* em vez de *gra*.

Ainda em relação à palavra-alvo *grade*, interessa apresentar uma produção de outro participante do grupo sem desvio fonológico. Tendo isso em vista, observemos, a seguir, na figura 26, a forma de onda da sílaba tônica da palavra-alvo *grade* produzida pela participante 3 do grupo sem desvio fonológico.

Figura 26 – Forma de onda da sílaba tônica *gra* da palavra-alvo *grade* produzida pelo participante 3 do GSDF



Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada indica o *tap*, o qual é precedido por vogal epentética e seguida pela vogal tônica da palavra-alvo *grade*.

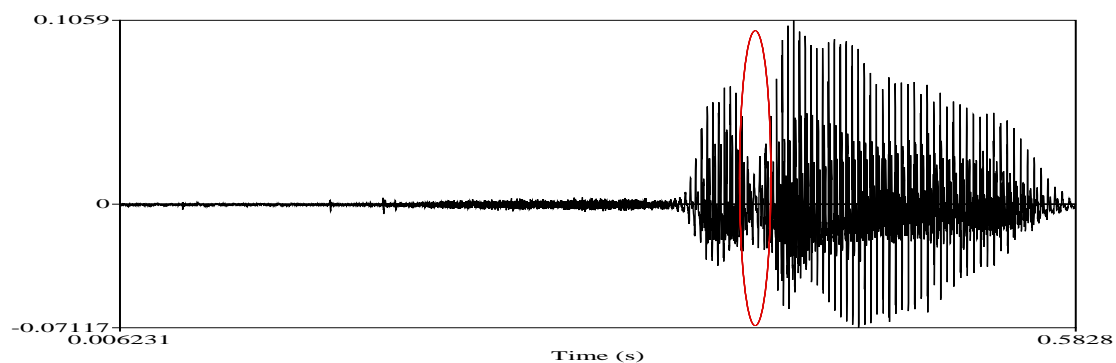
Note-se que, assim como o outro participante do GSDF (ver figura 24), este, do mesmo grupo, também realiza o *tap* em onset complexo, o qual é sonoro e tem duração de 31 ms. Observe-se que o *tap*, aí, é precedido por uma vogal epentética e seguido pela vogal tônica [a] da palavra-alvo *grade*.

Cumpra dizer, ainda, que tal participante dessonoriza a oclusiva velar sonora [g], como se pode notar. Aliás, vale relembrar que a dessonorização da velar [g] mostrou-se, com

certa recorrência, não só em outras produções desta participante mesma, mas também em algumas das produções de outros participantes de ambos os grupos GSDF e GDF, como se viu na análise específica das oclusivas sonoras.

Passemos agora para a análise da palavra-alvo *fraco*. Vejamos na figura 27 a produção de um participante do grupo sem desvio fonológico.

Figura 27 – Forma de onda sílaba tônica *fra* da palavra-alvo *fraco* produzida pelo participante 4 do GSDF



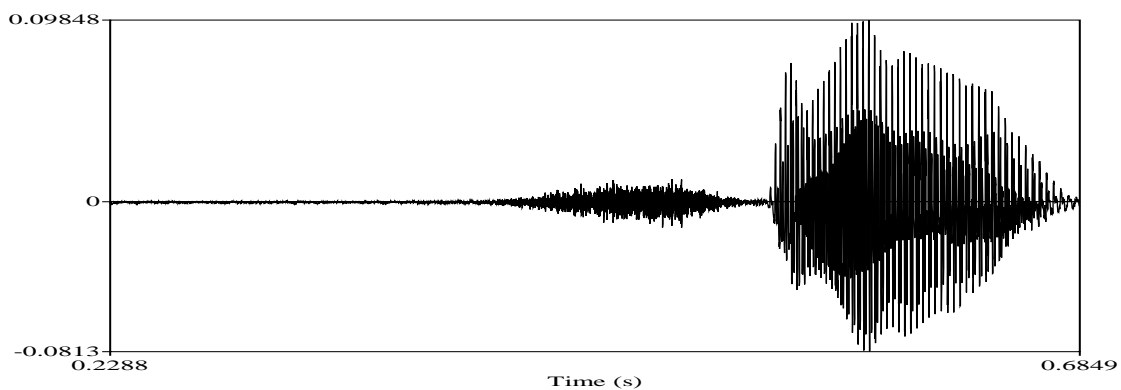
Fonte: Elaborado pelo autor.

Legenda: a elipse pontilhada indica o *tap*.

Observe-se na figura 27 que, após a fricativa labiodental surda, tem-se uma vogal epentética que é seguida pelo *tap*, o qual, por seu turno, tem duração de 17 ms. A duração da sílaba tônica, em pauta, é de 425 ms. Logo, o *tap*, neste caso, equivale a 4% da duração desta sílaba.

Agora vejamos a figura 28 na qual também consta a sílaba tônica da palavra-alvo *fraco*, porém referente à produção de um participante do grupo com desvio fonológico. Observe-se que este participante não realiza o *tap*.

Figura 28 – Forma de onda da sílaba tônica da palavra-alvo *fraco* produzida pelo participante 4 do GDF.



Fonte: Elaborado pelo autor.

Legenda: neste caso, como o participante não realizou o *tap*, o que se tem, nesta figura, é a sílaba *fa* em vez de *fra*.

Veja-se, aí, ainda na figura 28, que a fricativa labiodental surda é seguida imediatamente pela vogal tônica sem quaisquer vestígios do sinal do *tap*.

Como se viu, apresentaram-se, nesta análise final, a partir de dados acústicos, evidências de que os participantes do grupo sem desvio fonológico (GSDF) realizaram o *tap* em onset complexo, e de que os do grupo com desvio fonológico (GDF) não o realizaram.

Cumprе ressaltar, ainda, que estes não realizaram o *tap* em nenhum dos *clusters* investigados (/pɾ/ , /tɾ/ , /kɾ/ , /bɾ/ , /dɾ/ , /gɾ/, /fɾ/) - como já dito na análise anterior - e que todos eles – os participantes do GDF – tinham graus diferentes de desvio fonológico, os quais, de acordo com seus diagnósticos, variavam entre leve, moderado e severo.

Assim, independentemente da gravidade do desvio fonológico, nenhum dos participantes do GDF realizou o *tap* em onset complexo. Estas evidências se afinam, em certo aspecto, com um estudo de Ribas (2009), já reportado anteriormente, segundo o qual que 92% dos sujeitos com desvio fonológico, com idades entre 5 e 10 anos, não têm a estrutura silábica com onset complexo adquirida.



## 7 CONCLUSÃO

O objetivo principal desta tese consistiu em explicar padrões acústicos do *tap* em onset complexo na fala de crianças com e sem desvio fonológico à luz da Fonética Acústica. Nesse sentido, foram investigadas as produções acústicas do *tap* em onset complexo de dois grupos distintos, o GDF, constituído de crianças com desvio fonológico, e o GSDF, constituído de crianças sem o desvio fonológico.

A análise do *tap* em onset complexo corresponde à última análise apresentada na tese, a qual foi dividida em duas partes. Na primeira, consideraram-se duas categorias distintas e excludentes entre si, a saber: a realização e a não-realização do *tap* em onset complexo. Ambas foram analisadas a partir da assinatura acústica desta consoante nas produções orais de todos os participantes do GDF e do GSDF.

A categoria “realização” se guiou a partir da presença do sinal do *tap* e de suas variantes nos dados produzidos pelos mencionados participantes. Já a categoria “não-realização” se pautou, sobretudo, na ausência de vestígios do sinal do *tap* nesses dados mesmos.

Na segunda parte, evidenciaram-se as produções orais de ambos os grupos de crianças, por meio da apresentação de formas de ondas e espectrogramas derivadas dessas produções mesmas, interpretando-se o modo pelo qual tais crianças realizam, ou não realizam, o *tap* em onset complexo. As evidências daí apresentadas são de extrema relevância, visto que este é o primeiro trabalho do país, centrado na região Nordeste, que as expõe, quer pela espectrografia, quer pela forma de onda.

Também se objetivou nesta tese explicar e comparar outros detalhes fonéticos, relativos às consoantes que formavam *clusters* com *tap*, ou seja, as consoantes que o precediam *tap* nas palavras-alvo, *prato*, *braço*, *trave*, *craque*, *braço*, *dragão*, *grade* e *fraco*. Assim, o enfoque dos objetivos específicos se deu sobre as consoantes que o precediam o *tap* nos seguintes *clusters*: /pɾ/ , /tɾ/ , /kɾ/ , /bɾ/ , /dɾ/ , /gɾ/ e /fɾ/.

Frise-se que todos os objetivos desta tese foram alcançados, do principal aos específicos. Para dar conta destes, realizou-se análise acústica das reportadas consoantes nos *clusters* com *tap*, produzidas por participantes dos grupos GDF e GSDF.

A análise das consoantes que precediam o *tap* nesses *clusters* se dividiu em três partes. Na primeira, apresentou-se a análise das oclusivas surdas [p], [t] e [k], considerando-se o parâmetro VOT relativo; na segunda, a das oclusivas sonoras [b], [d] e [g], considerando-se também o VOT relativo, o vozeamento e o desvozeamento; e, na terceira, a da fricativa

labiodental surda [f], foram considerados os parâmetros duração relativa e a configuração espectral (FFT).

A pergunta norteadora desta pesquisa foi a seguinte: em que medida as produções acústicas do *tap* em onset complexo de crianças com desvio fonológico se distinguem das de crianças sem o desvio?

A hipótese básica foi a de que as produções acústicas do *tap* em onset complexo de crianças com desvio fonológico (GDF) se distinguem significativamente das de crianças sem o desvio (GSDF). Daí as previsões eram as de que o grupo com desvio fonológico (GDF) tenderia a não realizar o *tap*, não apresentando vestígios do sinal desta consoante, enquanto o grupo sem o desvio (GSDF) tenderia a realizá-lo, quer por meio de sua forma padrão (semelhante ao de um adulto), quer por meio de variantes livres, já conhecidas ou ainda desconhecidas.

Tal hipótese e tais previsões são, em certo aspecto, velhas e novas. São velhas por terem se pautado, em certa medida, na literatura dos estudos em aquisição fonológica do português brasileiro, os quais, antes mesmo de nós, já apontavam na direção daquilo que se aventou em nossa hipótese básica e em nossas previsões. Não obstante, os caminhos e a perspectiva trilhados nesta tese são outros totalmente distintos de todos os estudos que nos antecederam. Eis aí um dos motivos por que nossa hipótese e previsões são novas.

Outra razão pela qual se pode dizer que elas são novas se funda na visão que defendemos a respeito do saber, do conhecimento que, de algum modo, já está estabelecido ou preestabelecido. Tal visão pode ser sintetizada com as seguintes palavras de Coracini (2007, p.187), com as quais concordamos: “[...] todo saber pode ser questionado, ou pelo menos, deve ser visto na sua relatividade à situação: momento histórico-social, local geográfico [...].”

Tendo isso em vista, vale lembrar que os mencionados estudos em aquisição fonológica, como, por exemplo, os de Ribas (2007, 2011), se restringem apenas ao Sul e ao Sudeste do país, além de partirem de análises de oitiva sobre bancos de dados, como o DESFONO, e não de análise acústica, como se fez aqui. Diante disso, também cumpre lembrar que, no Nordeste, nosso estudo é pioneiro, seja em termos específicos, como no que se refere às evidências apresentadas mediante análise acústica de *clusters* com *tap* produzidos por crianças com e sem desvio fonológico, seja em termos gerais, como no que diz respeito a cada uma das análises desenvolvidas nesta tese.

Posto isso, aclare-se que, nas análises, os resultados corroboraram nossa hipótese básica. Ou seja, as produções acústicas do *tap* em onset complexo do grupo de crianças com desvio fonológico (GDF) se distinguem significativamente das do grupo de crianças sem o

desvio (GSDF). Constatou-se um elevado percentual de não-realização do *tap* pelos participantes do GDF, o qual atingiu a marca de 100% em todos os *clusters* investigados. Diante disso, não foi necessário proceder a um teste estatístico para constatar a existência de diferença significativa entre GDF e GSDF. Por conseguinte, as previsões foram acertadas. Isto é, as crianças com desvio fonológico tendem a não realizar o *tap* em onset complexo, enquanto as crianças sem o desvio não só tendem a realizá-lo, como também o realizam, seja por meio de sua forma padrão, seja através de suas variantes.

Assim, nossos resultados convergem, em certo aspecto, com determinados estudos em aquisição fonológica pautados em análise de oitiva, como o de Ribas (2009), segundo o qual 92% dos sujeitos com desvio fonológico não realizam o *tap* em onset complexo. Isso, de um lado, reitera o valor das contribuições trazidas por esses estudos de oitiva e, de outro, instiga novas reflexões sobre o estatuto da análise acústica, embora não coloque em xeque seu inestimável valor, tampouco, como argumenta Silva (2010), sua precedência em relação à análise de oitiva.

Quanto às hipóteses específicas, cumpre lembrar que, na primeira delas, afirmava-se que os grupos GDF e GSDF se distinguiriam em relação à produção das oclusivas surdas que precediam o *tap* em onset complexo. O parâmetro utilizado para isso, como se disse anteriormente, foi o VOT relativo. Nesse sentido, constataram-se diferenças nas produções de GDF e GSDF, mas não foram significativas, uma vez que, para as médias dos valores dos VOTs das oclusivas surdas [p], [t] e [k], obtiveram-se, respectivamente, os seguintes p-valores, com base no teste não-paramétrico de Wilcoxon-Mann-Whitney:  $p=0,4857$ ;  $p \cong 1,0$  e  $p=0,8885$ .

Entretanto, considerando-se os efeitos da variável sexo no VOT, encontrou-se diferença significativa ( $p= 0,03571$ ) para o nível de significância de 5%, em relação à média dos valores do VOT de [k], sendo a média das meninas mais alta que a dos meninos. Esse achado - frise-se - é de extrema relevância, visto que, até então, não existiam, no português brasileiro, estudos sobre efeitos do sexo no VOT produzido por crianças.

Já em outra hipótese específica, afirmava-se que haveria diferença entre os grupos GDF e GSDF quanto às oclusivas sonoras, sobretudo, no que se refere ao vozeamento e ao desvozeamento dessas consoantes, tendo sido o VOT relativo o parâmetro acústico pelo qual se testou esta hipótese. Nas análises, viu-se que 18% do total das consoantes oclusivas sonoras se realizaram com desvozeamento. O grupo de crianças com desvio fonológico (GDF) realizou desvozeamento em cerca de 14% de suas produções; enquanto o grupo sem o desvio (GSDF) atingiu o percentual de 22%. Isto, de certo modo, contrariou nossas previsões

segundo as quais haveria mais desvozeamento nas produções do grupo com desvio fonológico. Todavia, constatou-se, a partir de um teste estatístico Qui-Quadrado, que tais diferenças não foram estatisticamente significativas ( $p = 0,358$ ).

Aqui, interessa dizer que a análise acústica, por meio do software *Praat*, pode ser bastante útil para a clínica fonoaudiológica, seja na construção de diagnósticos, seja no acompanhamento de processos terapêuticos. Ela viabiliza a obtenção de dados precisos, nítidos e seguros, sobretudo, no que se refere ao vozeamento e ao desvozeamento das consoantes oclusivas. Com isso, a identificação de consoantes que se realizam, com vozeamento ou desvozeamento, se torna mais precisa e confiável do que aquela que se faz de oitiva na clínica. Aliás, frise-se que, para nós, a análise acústica do vozeamento/desvozeamento se revelou muito prática. Nesse ponto, em específico, acreditamos que, a partir da introdução da análise acústica na prática clínica/terapêutica, tratamentos mais eficientes e eficazes podem emergir.

Em outra hipótese específica, afirmava-se que os grupos GDF e GSDF se distinguiriam quanto à fricativa labiodental surda. Na testagem desta, o parâmetro duração relativa dessa fricativa foi considerado. Viu-se, porém, que os grupos não se distinguem quanto isso. Também se propôs outra hipótese específica sobre a fricativa labiodental surda, de acordo com a qual os grupos se distinguiriam quanto à sua composição espectral. Mas esta hipótese também não foi confirmada.

No entanto, descobriu-se que a composição espectral da fricativa labiodental surda se constitui como um padrão relativamente horizontal, o qual se verificou nas produções das crianças de ambos os grupos.

Ficam para futuros trabalhos alguns problemas pendentes: a análise da natureza das vogais na estrutura CCV com o *tap* em onset complexo; a análise do VOT das oclusivas surdas em “onset simples” na fala de meninos e meninas, que, parece seguir o mesmo padrão previsto para falantes adultos, homens e mulheres, do português brasileiro, como se viu nesta tese; a análise do *tap* em onset complexo seguido por outras vogais distintas de [a]; análise do *tap* em onset complexo em sintagmas como: “outro dragão”, “pedra trinta”, “quadra três”, “quatro *drones*”. Para nós, o que é de suma relevância, no momento, consiste em sugerir três coisas: a primeira, relativa à criação de laboratórios de Fonética em mais universidades do país, sobretudo, no norte e no Nordeste, para que se desenvolvam mais pesquisas nessas regiões, a fim de se obter uma visão mais consistente sobre a realidade da aquisição fonético-fonológica do português brasileiro; a segunda, referente à formação de professores de língua portuguesa, de modo que nesta formação não se explore apenas estudos sobre aquisição típica,

mas também sobre aquisição atípica, incluindo-se aí questões atinentes ao desvio fonológico; a terceira, relativa à construção de uma nova linha de pensamento por meio da qual a criança (ou o adulto), com “desvio fonológico”, ou com quaisquer outros “distúrbios lingüísticos”, aprenda a se impor como falante, de maneira que expresse suas palavras sem baixar a cabeça àqueles que tentam humilhá-la ou inferiorizá-la em detrimento de sua fala.

## REFERÊNCIAS

- ALBANO, E. **Codificação estatística das categorias fonéticas**: vestígio da dinâmica da fala na fonotaxe lexical. Belo Horizonte, 2007b. Versão revista da Conferência de abertura do IX Encontro Nacional de Fonética e Fonologia, intitulada “Modelos Fônicos Dinâmicos: para uma Contribuição Brasileira”, proferida em 26 de novembro de 2006 no auditório da reitoria da UFMG.
- ALBANO, E. Fonologia gestual e aquisição do sistema fônico hoje. *In*: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM DE PAULA, M. (Org.). **Estudos em aquisição fonológica**. Pelotas: Editora e Gráfica Universitária Pelotas, 2009. v. 2, p. 225-240
- ALBANO, E. **O gesto e suas bordas**: esboço de fonologia acústico-articulatória do português brasileiro. Campinas: Mercado de Letras: 2001.
- ALBANO, E. Representações dinâmicas e distribuídas: indícios do português brasileiro adulto e infantil. **Letras de Hoje**, Porto Alegre, v. 42, n. 1, p. 131-150, 2007a.
- ALBANO, E. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. Uma Introdução à Dinâmica em Fonologia. **Revista da ABRALIN**, Curitiba, v. 11, n. 1, p. 1-30. Jul. 2012.
- ALENCAR, M. **Aspectos socio-dialetais da língua falada em Fortaleza**: as realizações dos fonemas /r/ e /r/. 2007. Tese (Doutorado) - Universidade Federal do Ceará, Fortaleza, 2007.
- AMERICAN PSYCHIATRIC ASSOCIATION. **Diagnostic and statistical manual of mental disorders, fifth edition (DSM-V)**. Arlington, VA: American Psychiatric Association, 2013.
- AMERICAN PSYCHIATRIC ASSOCIATION. **Manual diagnóstico e estatístico de transtornos mentais**. Tradução Cláudia Dornelles. 4. ed. rev. Porto Alegre: Artmed, 2002.
- ARAGÃO, S. A neutralização dos fonemas /v, z, ʒ / no falar de Fortaleza. *In*: RIBEIRO, S; SÔNIA, B; CARDOSO, S. (Org.). **Dos sons às palavras**: nas trilhas da língua portuguesa. Salvador: EDUFBA, 2009. p. 189-200.
- BARBOSA, P. A.; MADUREIRA, S.; **Manual de fonética acústica experimental**: aplicações e dados do português. São Paulo: Cortez, 2015.
- BARROCO, M. A. L. *et al.* 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, jun. 2007.
- BERTI, L. **Aquisição incompleta do contraste entre /s/ e /ʃ/ em crianças falantes do português brasileiro**. 2006. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 2006
- BERTI, L. Contrastes e contrastes encobertos na produção da fala de crianças. **Pró-fono**, Barueri, v. 22, n. 4, p. 531-536, 2010.

BERTI, L.; FREITAS, M. O gesto fônico na aquisição “desviante”: movimento entre produção e percepção. O gesto fônico na aquisição desviante. **Revista da ABRALIN**, Curitiba, v. 11, n. 1, p. 197-220, jul. 2012.

BOERSMA, P.; WEENINCK, D. **Praat**: doing phonetics by computer (Versão 6.0.21) [Programa Computacional]. [S. l.], 2016. Disponível em: <<http://www.fon.hum.uva.nl/praat>>. Acesso em: 20 dez. 2016.

BRASIL, B.; MELO, R.; MOTA, H.; DIAS, R.; MEZZOMO, C.; GIACCHINI, V. O uso da estratégia de alongamento compensatório em diferentes gravidades do desvio fonológico. **Revista da Sociedade Brasileira de Fonoaudiologia**, São Paulo, v. 15, n. 2, p. 231-237, 2010.

CÂMARA JR., M. **Princípios da linguística geral**: como introdução aos estudos superiores da língua portuguesa. 2. ed. Livraria Acadêmica: Rio de Janeiro, 1954

CELESTE, M.; TEIXEIRA, É. Efeito do sexo e idade na produção do VOT. **Revista de Letras da Universidade Católica de Brasília**, Brasília, DF, v. 2, n. 1, p. 28-39, jul. 2009.

CHO, T.; LADEFOGED, P. Variation and universals in VOT: evidence from 18 languages. **Journal of Phonetics**, Amsterdam, v. 27, p. 207-229, 1999.

CORACINI, M. **Um fazer persuasivo**: o discurso subjetivo da ciência. Campinas: Pontes, 2007.

CRISTÓFARO-SILVA, T.; MIRANDA, I. Aquisição de econtros consonantais tautossilábicos: uma abordagem multirrepresentacional. **Revista Linguística**, Rio de Janeiro, v. 7, n. 1, p. 14-30, jun. 2011.

FRY, D. **Acoustic phonetics**: a course of basic readings. Cambridge: Cambridge University Press, 1976.

GIACCHINI, V.; MOTA, H.; MEZZOMO, C. Diferentes modelos de terapia fonoaudiológica nos casos de simplificação do onset complexo com alongamento compensatório. **Rev. CEFAC**, Cmpinas, v. 13, n. 1, p. 57-64, Jan./Fev. 2011.

GIANNETTI, E. **A ilusão da alma**: biografia de uma ideia fixa. São Paulo: Companhia das Letras, 2009.

GIERUT, J. Treatment efficacy: functional phonological disorders in children. **Journal Speech Language Hearing R.**, Hannover, v. 41, n. 1, p. S85-S100, Feb. 1998.

GRUNWELL, P. Os desvios fonológicos evolutivos numa perspectiva lingüística. *In*: YAVAS, M. (Org.). **Desvios fonológicos em crianças**. Porto Alegre: Mercado Aberto, 1990.

HERNANDORENA, C. **A aquisição da fonologia do português**: estabelecimento de padrões com base em traços distintivos. 1990. Tese (Doutorado) – Faculdade de Letras, Pontifícia Universidade Católica do rio Grande do Sul, Porto Alegre, 1990.

JESUS, L.; SHADDLE, C. H. Acoustic analysis of European Portuguese uvular and voiceless tapped alveolar fricatives. **Journal of the International Phonetic Association**, Cambridge, v. 35, n. 1, p. 27-44, 2005.

LADEFODEG, P. **Phonetic data analysis**. Oxford: Blackwell Publishing, 2003.

LADEFODEG, P.; MADDIESON, I. **The sounds of the world's languages**. Oxford: Blackwell Publishers, 1996.

LAMPRECHT, R. **Perfil da aquisição normal da fonologia do Português**: descrição longitudinal de 12 crianças: 2:9 a 5:5. 1990. Tese (Doutorado) – Faculdade de Letras da PUCRS, Porto Alegre, 1990.

LISKER, L.; ABRAMSON, A. A cross-language study of voicing in initial stops: acoustical measurements. **Word**, [s. l.], v. 20, n. 3, p. 384-422, 1964.

MACAMBIRA, J. **Fonologia do português**. Fortaleza: Secretaria de Cultura e Desporto, 1985.

MARUSSO, A. S. Princípios básicos da teoria acústica de produção da fala. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 13, p. 19-43, 2005.

MEZZOMO, C. **Aquisição de fonemas na posição de coda medial do português brasileiro, em crianças com desenvolvimento fonológico normal**. 1999. Dissertação (Mestrado) – Faculdade de Letras, Pontifícia Universidade Católica do rio Grande do Sul, Porto Alegre, 1999.

MEZZOMO, C.; MOTA, H.; DIAS, R.; GIACCHINI, V. O uso da estratégia de alongamento compensatório em crianças com desenvolvimento fonológico normal e desviante. **Letras de Hoje**, Porto Alegre. v.3. n.3 p. 35-41, jul./set. 2008.

MEZZOMO, C.; MOTA, H.; DIAS, R. A estratégia de alongamento compensatório e sua relação com a consciência fonológica. **Distúrb. Comum.**, São Paulo, v. 23, n. 2, p. 125-131, ago. 2011.

MIRANDA, A. **A aquisição do “r”**: uma contribuição à discussão sobre seu status fonológico. 1996. Dissertação (Mestrado) – Faculdade de Letras, Pontifícia Universidade Católica do rio Grande do Sul, Porto Alegre, 1996.

MIRANDA, I. **Aquisição e variação estruturada de encontros consonantais tautossilábicos**. 2007. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

NÚCLEO DE ATENÇÃO MÉDICA INTEGRADA NAMI. **Núcleo de Atenção Médica Integrada Nami – Unifor**. Fortaleza, 2016. Disponível em: <[www.unifor.br/nami](http://www.unifor.br/nami)>. Acesso em: 26 out. 2016.

NIETZSCHE, F. **Além do bem e do mal**: prelúdio a uma filosofia do futuro. Tradução: Paulo César de Souza. São Paulo: Companhia das Letras, 2005.



ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Classificação de transtornos mentais e de comportamento da CID-10**: descrições clínicas e diretrizes diagnósticas. Tradução Dorgival Caetano. Porto Alegre: Artmed, 1993.

PAHOCA-LEVY, I. **Uma nova face da nau dos insensatos**: a dificuldade de vozear obstruintes em crianças de idade escolar, 1993. Tese (Doutorado em Linguística) – Universidade de Campinas, Campinas, 1993.

RAMOS, A. **Processos de estrutura silábica em crianças com desvios fonológicos**: uma abordagem não-linear. 1996. Tese (Doutorado em Linguística Aplicada) – Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1996.

RIBAS, P. **Aquisição do onset complexo no português brasileiro**. 2002. Dissertação (Mestrado) – Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

RIBAS, P. **Onset complexo nos desvios fonológicos**: descrição, implicações para a teoria, contribuições para terapia. 2006. Tese. (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2006.

RIBAS, P. Os dados de aquisição fonológica atípica na constituição silábica. *In*: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM DE PAULA, M. (Org.). **Estudos em aquisição fonológica**. Pelotas: Editora e Gráfica Universitária Pelotas, 2009. v. 2, p. 63-76.

RIBAS, P. Aquisição fonológica atípica: características dos dados de crianças com desvio fonológico evolutivo. *In*: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM DE PAULA, M. (Org.). **Estudos em aquisição fonológica**. Pelotas: Editora e Gráfica Universitária Pelotas, 2011. v. 3, p. 223-250.

RIZZOTO, A. **Os processos fonológicos de estrutura silábica no desenvolvimento fonológico normal e nos desvios fonológicos evolutivos**. 1997. Dissertação (Mestrado) - Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.

ROSSI-BARBOSA, L; CALDEIRA, A.; HONORATO-MARQUES, R.; SILVA, R.; Prevalência de transtornos fonológicos em crianças do primeiro ano do ensino fundamental. **Revista da Sociedade Brasileira de Fonoaudiologia**, São Paulo, v. 16, n. 3, p. 330-336, July/Sept. 2011. ISSN 1982-0232.

SERAINÉ, F. Introdução ao atlas linguístico e folclórico do Cariri. **Separata da Revista do Instituto do Ceará**, Fortaleza, Tomo 86, v. 91, 1972.

SILVA, P. O estatuto da análise acústica nos estudos fônicos. **Cadernos de Letras da UFF – Dossiê**: Letras e cognição, Niterói, n. 41, p. 213-229, 2010.

STEVENS, K. **Acoustic phonetics**. Cambridge: MIT Press, 1998.

TEIXEIRA, E. Reflexões sobre a relação entre processos fonológicos aquisicionais e processos marcadores de estigmatização sociolinguística. *In: SIMPÓSIO SOBRE A DIVERSIDADE LINGUISTICA NO BRASIL*, 1., 1986, Salvador. **Atas...** Salvador: UFBA, 1986. p. 101-107.

TEIXEIRA, E. Perfil de desenvolvimento fonológico em português (P.D.F.P). **Estudos Linguísticos e Literários**, Salvador, n. 12. p. 225-237, 1991.

TEIXEIRA, E. Um estudo sobre Processos de Simplificação Fonológica na aquisição do português. *In: RIBEIRO, S.; SÔNIA, B.; CARDOSO, S. Dos sons às palavras: nas trilhas da língua portuguesa*. Salvador: EDUFBA, 2009. p. 173-185.

TRUBETZKOY, N. **Principles de phonologie**. Paris: Klincksieck, 1949.

VASSOLER, A. M. O. **Coordenação gestual na produção de encontros consonantais em crianças com desenvolvimento de linguagem típico e atípico**. 2016. Tese (Doutorado) - Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista, São José do Rio Preto, 2016.

VIDOR, D. **Aquisição das líquidas não-laterais por crianças com desvios fonológicos evolutivos**: descrição, análise e comparação com o desenvolvimento normal. 2000. 159 f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.

VOGELEY, A. Desvios fonológicos. *In: AZEVEDO, N.; LIMA DA FONTE, R. Aquisição da linguagem, seus distúrbios e especificidades: diferentes perspectivas*. Curitiba: CRV, 2011. p. 151-164.

WERTZNER, H. Transtorno fonológico. *In: FERREIRA-GONÇALVES, G.; KESKE-SOARES, M.; BRUM DE PAULA, M. (Org.). Estudos em aquisição fonológica*. Pelotas: Editora e Gráfica Universitária PREC-UFPel, 2011. v. 3 p. 133-151.

WHITESIDE, Sandra; HENRY, Luiza; DOBBIN, Rachel. Sex differences in voice onset time: a developmental study of phonetic context effects in British English. **The Journal of the Acoustic Society of America**, [S. l.], v. 116, n. 2, p. 1179-1183, 2004.