

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS

FABIANA ANDRADE PENIDO

**PONDERAÇÃO DE PISTAS AUDITIVAS POR CRIANÇAS E
ADULTOS FALANTES NATIVOS DO PORTUGUÊS BRASILEIRO**

BELO HORIZONTE

2017

FABIANA ANDRADE PENIDO

**PONDERAÇÃO DE PISTAS AUDITIVAS POR CRIANÇAS E
ADULTOS FALANTES NATIVOS DO PORTUGUÊS BRASILEIRO**

Tese apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Linguística Teórica e Descritiva.

Área de Concentração: Linguística Teórica e Descritiva

Linha de pesquisa: 1B – Fonética e Aspectos Sonoros da Gramática

Orientador: Prof. Dr. Rui Rothe-Neves

Belo Horizonte
Faculdade de Letras
2017

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

P411p Penido, Fabiana Andrade.
Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro [manuscrito] / Fabiana Andrade Penido. – 2017.
192 f., enc. : il., grafs (color), tabs (p&b)
Orientador: Rui Rothe-Neves.
Área de concentração: Linguística Teórica e Descritiva.
Linha de Pesquisa: Fonética e Aspectos Sonoros da Gramática.
Tese (doutorado) – Universidade Federal de Minas Gerais, Faculdade de Letras.
Bibliografia: f. 160-169.
Anexos: f. 170-192.

1. Percepção da fala nas crianças – Teses. 2. Língua portuguesa – Fonologia – Teses. 3. Língua portuguesa – Fonética – Teses. 4. Percepção da fala – Teses. 5. Audição – Teses. 6. Fonética acústica – Teses. I. Rothe-Neves, Rui, 1970-. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 414



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS



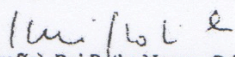
FOLHA DE APROVAÇÃO

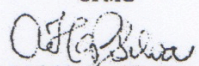
PONDERAÇÃO DE PISTAS AUDITIVAS POR CRIANÇAS E ADULTOS FALANTES NATIVOS DO PORTUGUÊS BRASILEIRO

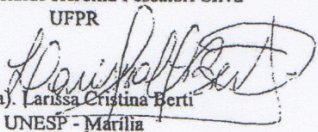
FABIANA ANDRADE PENIDO

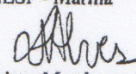
Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Doutor em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Organização Sonora da Comunicação Humana.

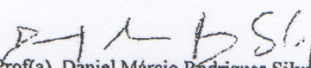
Aprovada em 26 de junho de 2017, pela banca constituída pelos membros:


Prof(a). Rui Rothe-Neves - Orientador
UFMG


Prof(a). Adelaide Hercília Pescatori Silva
UFPR


Prof(a). Larissa Cristina Berti
UNESP - Marília


Prof(a). Luciana Mendonça Alves
Imih


Prof(a). Daniel Márcio Rodrigues Silva
UFMG

Belo Horizonte, 26 de junho de 2017.

AGRADECIMENTOS

A Deus, por ser o meu companheiro de todas as horas e por sempre abrir todos os caminhos para o progresso dos meus estudos.

Ao Prof. Dr. Rui Rothe-Neves, meu querido orientador, pela tranquilidade de sempre, pelas conversas repletas de conhecimento científico e por todos esses seis anos de convivência harmoniosa.

À minha coorientadora da Universidade de Edimburgo, Alice Turk, pelo carinho, atenção e disponibilidade durante o tempo que permaneci em Edimburgo até a conclusão deste trabalho.

Aos meus pais amados, Zeca e Sandra, pelo carinho e incentivo à realização deste projeto. Vocês são bacanas demais!

Ao meu amado marido, Leo, por ser um companheiro maravilhoso e por fazer parte da minha vida, inclusive me acompanhando nas aventuras escocesas. Você é o melhor!

Aos queridos Carol, Didi e Alice, pelo carinho de sempre.

A todos os meus amigos, tios, tias, primos, primas, sogro, sogra e sobrinhas, por me incentivarem durante mais esta etapa. Em especial, à minha linda “Vovó Consuêlo”, por ser um exemplo de vida, fé e sabedoria.

Ao meu grande amigo Lucas, “Luquito”, por ter se esforçado na instalação do *software* de síntese de fala, o que permitiu o prosseguimento desta pesquisa.

À “turma da perícia”, que sempre ficou na torcida para que tudo desse certo. Em especial, a Lê, Lílian, Cleu, Si, Raquel, Pri, Wal, Tati, Gizelle, Rosana, Deise, Maria Cecília e Stephanie, pela disponibilidade e pelo empréstimo de seus ouvidos.

Ao Governo do Estado de Minas Gerais e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo incentivo e investimento nos meus estudos.

À Universidade de Edimburgo, pela receptividade e atenção durante a realização do doutorado sanduíche.

Aos alunos do Colégio Salesiano e da Escola Infantil “O Pequeno Construtor”, pelo envolvimento com o experimento e pelo carinho.

Ao pequeno João, meu companheirinho preferido, pelo amor que supera qualquer cansaço.
Te amo além do infinito!

RESUMO

Objetivo: Este estudo investigou a mudança no desenvolvimento da ponderação de pistas auditivas para a percepção das fricativas coronais não vozeadas ([f] e [s]) e das plosivas bilabial e alveolar ([b] e [d]) do português brasileiro. A principal hipótese testada foi a de que as crianças atribuem maior peso às pistas dinâmicas do sinal acústico de fala enquanto os adultos utilizam mais a pista estática, conforme propõem Nittrouer, Manning e Meyer (1993). **Metodologia:** A amostra foi composta por crianças de quatro anos de idade, crianças de sete anos e adultos. Em cada grupo de idade foram testados doze sujeitos. As crianças e os adultos foram recrutados em duas instituições de ensino e em uma instituição pública de Belo Horizonte/MG. Para cada par de palavras investigado, dentre eles /sapa/-/sapa/, /faba/-/saba/ e /bata/-/data/, foram produzidos dois contínuos sintetizados. Duas pistas auditivas foram manipuladas: uma pista foi manipulada de forma idêntica em ambos os contínuos e a outra pista foi manipulada de forma diferente nos dois contínuos. Foi realizada uma tarefa de classificação, em que os participantes ouviram os estímulos por meio de um fone de ouvido e tiveram que decidir qual categoria era mais coerente com o que foi apresentado dentre as possibilidades existentes. **Resultados:** Os dados coletados foram analisados estatisticamente por meio do modelo de análise probito, da análise de variância (ANOVA) e da análise multivariada da variância (MANOVA) para as fricativas e do teste “t pareado” para as plosivas. Os resultados revelaram que para os pares de palavras e pseudopalavras compostos pelas fricativas /sapa/-/sapa/ e /faba/-/saba/, os adultos e as crianças de sete anos de idade ponderaram mais fortemente a pista estática de altura da frequência do ruído fricativo quando comparados com as crianças mais jovens. No entanto, nenhum dos três grupos de idade foi mais influenciado pela pista dinâmica de transição dos formantes. Além disso, verificou-se uma mudança no padrão de respostas dos grupos de idade para a classificação das palavras e das pseudopalavras, indicando um efeito do estatuto lexical. Já para o par de palavras /bata/-/data/, composto pelas plosivas, as crianças de quatro anos atribuíram menor peso à pista dinâmica de transição dos formantes do que as crianças de sete anos e os adultos. Ainda neste par, a pista estática de amplitudes do *burst* não foi ponderada de forma significativa por nenhuma faixa etária. Conforme esperado, em todos os pares de palavras testados, as crianças mais jovens apresentaram um padrão de respostas mais inconsistente, menos categórico, com curvas de classificação menos íngremes, na qual foi possível observar uma modificação no padrão de respostas ao longo do desenvolvimento da criança. **Conclusões:** A partir dos resultados pode-se concluir que nem sempre as crianças mais jovens ponderaram mais fortemente a pista dinâmica quando comparadas com os adultos. Concluiu-se também que as estratégias de ponderação de pistas parecem modificar de acordo com: a) a experiência da criança com sua língua nativa; b) o contraste fonológico testado e c) a informatividade que as pistas auditivas fornecem aos ouvintes para a classificação dos contrastes.

Palavras chave: Percepção da fala; Ponderação; Desenvolvimento; Criança.

ABSTRACT

Objective: This study investigated the developmental changes in cue weighting for the distinction of voiceless coronal fricatives ([ʃ] and [s]) and of bilabial and alveolar stops ([b] and [d]) in Brazilian Portuguese. The main hypothesis tested was that children weight the dynamic cues more than adults, whereas adults used more the static cues (Nittrouer, Manning & Meyer, 1993). **Methodology:** The sampling was constituted of children of ages four and seven years old, as well as adults. Twelve subjects were tested in each of the three age groups. Both children and adults were recruited from two education centers and from a government institution in Belo Horizonte/MG. For each words pair investigated, among them /ʃapa-/sapa/, /ʃaba-/saba/ and /bata-/data/, two synthesised continua were produced. Two auditory cues were manipulated: one cue was manipulated identically in both continua and the other cue was manipulated differently in the two continuums. The participants performed a classification task in which they heard the stimuli on headphones and decided which category was the most coherent to what had been presented among the existing possibilities. **Results:** The data gathered was statistically analyzed using the probit model, the analysis of variance (ANOVA) and the multivariate analysis of variance (MANOVA) for the fricatives and the paired t-test for the plosives. The results revealed that for the minimal pair /ʃapa-/sapa/ and for the nonword pair /ʃaba-/saba/, adults and seven-year-old children weighted the static cue of centre frequency of the frication noise more heavily than the younger children. However, none of the three age groups was more influenced by the dynamic cue of formants transition. In addition, there was a change in the responses pattern of the three age groups for the classification of words and nonwords, indicating an effect of the lexical status. For the words pair /bata-/data/, composed by stops, the four-year-old children weighted more heavily the dynamic cue of formants transition than seven-year-old children and the adults. Still in this pair, the static cue of burst amplitudes wasn't significantly weighted by any age group. As expected, in all word pairs, younger children showed more inconsistent and categorical less response pattern with shallower classification curves, in which was observed a change in responses pattern throughout the development of the child. **Conclusions:** From the results it can be concluded that younger children have not always weight more heavily the dynamic cue than the adults. It was also concluded that the strategies of cue weighting seem to modify according to: a) the child's experience with his native language; b) the phonological contrast tested; and c) the informativeness that the auditory cues provide to the listeners for the classification of the contrasts.

Key words: Speech Perception; Weighting, Development; Child.

LISTA DE ILUSTRAÇÕES

- FIGURA 1** - Resultados da tarefa de classificação das crianças de três, cinco, sete anos de idade e adultos para os contrastes /ja/-/sa/ e /ju/-/su/.....42
- FIGURAS 2 e 3** - Oscilograma e espectrograma da fricativa coronal não vozeada [s] (à esquerda) e da fricativa coronal vozeada [z] (à direita).....60
- FIGURA 4** - Barra de vozeamento na plosiva [b] da sílaba [ba], visualizada pela região mais escurecida na base do espectrograma.....64
- FIGURA 5** - Três primeiros formantes (F1, F2 e F3) da vogal [a] seguida da fricativa [s], na palavra [sapa].....66
- FIGURAS 6 e 7** - F1 e F2 da vogal [a] seguinte à fricativa [ʃ], na palavra /ʃapa/ (à esquerda) e F1 e F2 da vogal [a] seguinte à fricativa [s], na palavra /sapa/ (à direita) – Informante do sexo masculino de 33 anos de idade.....67
- FIGURAS 8 e 9** - F1 e F2 da vogal [a] seguinte à plosiva [b], na palavra /bata/ (à esquerda) e F1 e F2 da vogal [a] seguinte à plosiva [d], na palavra /data/ (à direita) – Informante do sexo masculino de 33 anos de idade.....68
- FIGURA 10** - Exemplo da anotação realizada nos dados de fala dos informantes; informante masculino, 34 anos.....70
- FIGURA 11** - Região de mudança abrupta da amplitude visualizada no oscilograma entre a consoante [s] e a vogal [a], na palavra /sapa/.....71
- FIGURA 12** - Ponto inicial da vogal [a] na palavra /sapa/. No oscilograma observa-se o primeiro pulso glótico e no espectrograma nota-se o ponto inicial do segundo formante vocálico (seta preta).....72
- FIGURA 13** - Porção de 0.03s de ruído e 70ms da vogal [a] na palavra /sapa/.....73

FIGURA 14 - Tela do Praat que disponibiliza o <i>Formant listing</i> , da região selecionada no espectrograma da figura 13.....	73
FIGURA 15 - Formantes da primeira vogal [a] da palavra /sapa/, com indicação dos pulsos glóticos regulares no oscilograma e da porção estável de 20ms da vogal.....	74
FIGURA 16 - Formantes da primeira vogal [a] da palavra /bata/, com indicação dos pulsos glóticos regulares no oscilograma e da porção estável de 20ms da vogal.....	74
FIGURA 17 - Seleção da vogal tônica [a] da palavra /sapa/.....	75
FIGURA 18 - Porção do segmento fricativo [s] na palavra /sapa/ que foi selecionada e extraída para análise.....	76
FIGURA 19 - Fatia do espectro do ruído fricativo [s] com a indicação dos picos de amplitude do quarto e do quinto formante (F4 e F5).....	77
FIGURA 20 - Janela de objetos do Praat para a obtenção do valor referente ao centro de gravidade (seta preta).....	78
FIGURA 21 - Barra de explosão selecionada na plosiva [b], da palavra /bata/.....	79
FIGURA 22 - Forma espectral do <i>burst</i> da plosiva [b], na palavra /bata/ com a indicação dos picos de amplitude do quarto formante (F4).....	80
FIGURA 23 - Duração da fase de pré-sonorização do segmento plosivo [b], na palavra /bata/. A seta indica a duração do <i>burst</i> , sinalizada pelas duas barras azuis.....	81
FIGURA 24 - Espectro acústico das fricativas do árabe egípcio. O espectro foi produzido pela análise LPC selecionando a porção média do ruído fricativo. As setas mostram os picos espectrais da fricativas.....	85
FIGURAS 25 e 26 - Forma espectral do <i>burst</i> da plosiva [b] (esquerda) e da plosiva [d] (direita).....	90

FIGURA 27 - Processo de produção da fala utilizado na síntese de formantes.....	100
FIGURA 28 - Configuração dos ressoadores em cascata (R5 à R1).....	101
FIGURA 29 - Configuração dos ressoadores em paralelo com as amplitudes dos picos espectrais (formantes).....	102
FIGURA 30 - Frequência fundamental da vogal [a] que seguiu a fricativa [ʃ], na palavra /ʃapa/, sinalizada pela linha azul. No oscilograma observa-se o comportamento da amplitude de vozeamento (seta preta).....	105
FIGURAS 31 e 32 - Espectrograma da vogal [a] seguinte a fricativa [ʃ]. À esquerda observa-se o espectrograma do som original (vogal natural) e à direita o espectrograma da vogal sintetizada.....	106
FIGURAS 33 e 34 - Fatia do espectro (do inglês, <i>Spectral slice</i>) da fricativa [s]. À esquerda observa-se a fatia do espectro do som original (fricativa natural) e à direita a fatia do espectro da fricativa sintetizada.....	108
FIGURAS 35 e 36 - Fatia do espectro (do inglês, <i>Spectral slice</i>) da fricativa [ʃ]. À esquerda observa-se a fatia do espectro do som original (fricativa natural) e à direita a fatia do espectro da fricativa sintetizada.....	110
FIGURAS 37 e 38 - Espectrograma da fricativa [s]. À esquerda observa-se o espectrograma do som original (fricativa natural) e à direita o espectrograma da fricativa sintetizada.....	110
FIGURAS 39 e 40 - Espectrograma da fricativa [ʃ]. À esquerda observa-se o espectrograma do som original (fricativa natural) e à direita o espectrograma da fricativa sintetizada.....	111
FIGURAS 41 e 42 - À esquerda observa-se o espectrograma da plosiva [b] sintetizada e à direita observa-se o espectrograma da plosiva [d] sintetizada.....	114

FIGURA 43 - Envoltórios espectrais obtidos em 50 ms em torno do ponto central do primeiro (em preto, linha contínua) ao décimo (em vermelho) ruídos fricativos.....	117
FIGURA 44 - Espectrograma do melhor exemplar da palavra /fapa/ sintetizada.....	121
FIGURA 45 - Espectrograma do melhor exemplar da palavra /sapa/ sintetizada.....	122
FIGURA 46 - Espectrograma do melhor exemplar da palavra /bata/ sintetizada.....	123
FIGURA 47 - Espectrograma do melhor exemplar da palavra /data/ sintetizada.....	124
FIGURA 48 - Tela de exibição dos estímulos do par mínimo /fapa/-/sapa/.....	125
FIGURA 49 - Tela de exibição dos estímulos do par mínimo /bata/-/data/.....	125
FIGURA 50 - Tela de exibição dos estímulos do par de pseudopalavras /faba/-/saba/.....	126
FIGURA 51 - Resultados da tarefa de classificação dos adultos para o par mínimo /bata/-/data/.....	128
FIGURAS 52 e 53 - Resultados da tarefa de classificação dos adultos para o par mínimo /fapa/-/sapa/.....	129
FIGURAS 54 e 55 - Resultados da tarefa de classificação dos adultos para o par de pseudopalavras /faba/-/saba/.....	130
FIGURA 56 - Procedimento experimental utilizado com as crianças de quatro e de sete anos e com os adultos.....	134
FIGURA 57 - Resultados da tarefa de classificação para o par mínimo /fapa/-/sapa/ e para o par de pseudopalavras /faba/-/saba/.....	140
FIGURA 58 - Resultados da tarefa de classificação para o par mínimo /bata/-/data/.....	141

QUADRO 1 - Parâmetros manipulados ao longo do tempo no <i>software</i> Sensyn. Vogal [a] que seguiu a fricativa [ʃ].....	104
QUADRO 2 - Parâmetros manipulados ao longo do tempo no <i>software</i> Sensyn. Fricativa [s], das palavras /sapa/ e /saba/.....	107
QUADRO 3 - Parâmetros manipulados ao longo do tempo no <i>software</i> Sensyn. Fricativa [ʃ], das palavras /ʃapa/ e /ʃaba/.....	109
QUADRO 4 - Parâmetros manipulados ao longo do tempo no <i>software</i> Sensyn. Plosiva [b], da palavra /bata/.....	112
QUADRO 5 - Parâmetros manipulados ao longo do tempo no <i>software</i> Sensyn. Plosiva [d], da palavra /data/.....	113
GRÁFICO 1 - Formantes da porção inicial e estável da vogal [a] seguinte às fricativas [s] e [ʃ].....	93
GRÁFICO 2 - Formantes da porção inicial e estável da vogal [a] seguinte às plosivas [b] e [d].....	96
GRÁFICO 3 - Q-Q plot - Efeito da transformação dos dados em arcosenos.....	137
GRÁFICO 4 - Inclinação (<i>slope</i>) das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Fricativas.....	146
GRÁFICO 5 - Fronteira fonêmica das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Fricativas.....	147
GRÁFICO 6 - Fronteira fonêmica e inclinação (<i>slope</i>) das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Plosivas.....	148

LISTA DE TABELAS

TABELA 1 - Duração das fricativas [s] e [ʃ] em (ms).....	82
TABELA 2 - Frequência dos formantes das fricativas [s] e [ʃ] em (Hz).....	83
TABELA 3 - Amplitude dos formantes das fricativas [s] e [ʃ] em (dB).....	84
TABELA 4 - Centro de gravidade das fricativas [s] e [ʃ] em (Hz).....	86
TABELA 5 - Duração da fase de pré-sonorização [b] e [d] em (ms).....	89
TABELA 6 - Duração da fase de soltura da oclusão [b] e [d] em (ms).....	89
TABELA 7 - Duração da vogal [a] seguinte às fricativas (ms).....	91
TABELA 8 - Frequências dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ].....	92
TABELA 9 - Duração da vogal [a] seguinte às plosivas em (ms).....	94
TABELA 10 - Frequência dos formantes da vogal [a] seguinte às plosivas [b] e [d].....	95
TABELA 11 - Valores das frequências dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] em (Hz), informante 1.....	118
TABELA 12 - Valores das frequências dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] em (Hz), informante 2.....	119
TABELA 13 - Valores das frequências dos formantes do continuum vocálico [a] seguinte às plosivas [b] e [d] em (Hz).....	120
TABELA 14 - Valores médios e desvio padrão da fronteira fonêmica e da inclinação (<i>slope</i>) das curvas de classificação.....	143

SUMÁRIO

1 INTRODUÇÃO	18
1.1 Contextualização do tema.....	18
1.2 Justificativa e hipóteses.....	22
1.3 Objetivo geral.....	24
1.4 Objetivos específicos.....	24
1.5 Estrutura da tese.....	24
2 FUNDAMENTAÇÃO TEÓRICA	27
2.1 A aquisição da linguagem.....	27
2.1.1 Teorias linguísticas de aquisição da linguagem.....	28
2.2 O desenvolvimento fonológico da criança: a percepção.....	29
2.3 A categorização dos sons da fala.....	31
2.3.1 A percepção categórica.....	33
2.3.2 O desenvolvimento das categorias fônicas.....	37
2.4 A ponderação perceptiva de pistas auditivas.....	40
2.4.1 A ponderação perceptiva de pistas auditivas nos estudos desenvolvidos com os sons fricativos e plosivos.....	45
2.4.2 As hipóteses que buscam explicar as diferenças na ponderação de pistas auditivas entre crianças e adultos.....	53
3 CAPÍTULO 1 - ESTUDO PILOTO DAS PROPRIEDADES ACÚSTICAS DAS FRICATIVAS CORONAIIS NÃO VOZEADAS E DAS PLOSIVAS VOZEADAS [b] E [d] DO PORTUGUÊS BRASILEIRO EM POSIÇÃO INICIAL DE PALAVRA	59
3.1 Descrição das propriedades acústicas das fricativas, com foco nas fricativas [s] e [ʃ].....	59
3.2 Descrição das propriedades acústicas das plosivas, com foco nas plosivas [b] e [d].....	63
3.3 Descrição das propriedades acústicas da vogal [a], com foco na transição dos formantes (F1, F2 e F3).....	65
3.4 Metodologia do estudo piloto.....	68
3.4.1 Participantes do estudo.....	68
3.4.2 Procedimento.....	69
3.4.3 Análise das propriedades acústicas dos estímulos.....	69
3.4.3.1 Medidas dos formantes e da duração vocálica.....	71
3.4.3.2 Medidas dos segmentos fricativos e plosivos.....	75
3.4.3.2.1 Propriedades acústicas dos segmentos fricativos.....	76

3.4.3.2.2 Propriedades acústicas dos segmentos plosivos.....	79
3.5 Resultados e considerações do estudo piloto.....	81
3.5.1 Propriedades acústicas das fricativas [s] e [ʃ].....	81
3.5.1.1 Duração das fricativas.....	81
3.5.1.2 Formantes e picos de amplitude do ruído fricativo.....	82
3.5.1.3 Centro de gravidade do ruído fricativo.....	86
3.5.2 Propriedades acústicas das plosivas [b] e [d].....	87
3.5.2.1 Duração absoluta dos segmentos.....	87
3.5.2.2 Picos de amplitude e formantes da soltura da oclusão.....	88
3.5.3 Medidas da duração e dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] e às plosivas [b] e [d].....	90
3.5.3.1 Medidas da duração e dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ]....	90
3.5.3.2 Medidas da duração e dos formantes da vogal [a] seguinte às plosivas [b] e [d]....	94
3.5.4 Resumo das principais conclusões do estudo piloto.....	96
4 CAPÍTULO 2 – SÍNTESE DE FALA DOS ESTÍMULOS.....	98
4.1 Síntese do sinal de fala.....	98
4.2 Principais tipos de sintetizadores.....	99
4.2.1 Sintetizador articulatório.....	99
4.2.2 Sintetizador de formantes.....	99
4.2.2.1 Ressonadores em cascata.....	101
4.2.2.2 Ressonadores em paralelo.....	101
4.3 Parâmetros de controle para síntese dos sons da fala.....	102
4.3.1 Síntese da vogal [a] seguinte às fricativas [s] e [ʃ] e às plosivas [b] e [d].....	103
4.3.2 Síntese das fricativas [ʃ] e [s].....	106
4.3.3 Síntese das plosivas [b] e [d]	111
5 CAPÍTULO 3 – ESTUDO PILOTO COMPOSTO PELA TAREFA DE CLASSIFICAÇÃO DOS ESTÍMULOS SINTETIZADOS /ʃapa/-/sapa/, /ʃaba/-/saba/ E /bata/-/data/.....	115
5.1 Metodologia do estudo piloto composto pela tarefa de classificação dos estímulos sintetizados.....	115
5.1.1 Participantes do estudo.....	115
5.1.2 Estímulos.....	116
5.1.2.1 Continua compostos pelas palavras /ʃapa/-/sapa/ e /ʃaba/-/saba/.....	117

5.1.2.1.1 Ruído fricativo.....	117
5.1.2.1.2 Porção vocálica seguinte às fricativas.....	118
5.1.2.2 Continua compostos pelas palavras /bata/-/data/.....	119
5.1.2.2.1 Plosivas [b] e [d].....	119
5.1.2.2.2 Porção vocálica seguinte às plosivas.....	120
5.1.3 Continua.....	121
5.1.4 Procedimento experimental.....	123
5.1.5 Resultados e considerações do estudo piloto.....	127
5.1.5.1 Resultados da tarefa de classificação com o par mínimo /bata/-/data/.....	127
5.1.5.2 Resultados da tarefa de classificação com o par mínimo /fapa/-/sapa/, dos informantes 1 e 2.....	128
5.1.5.3 Resultados da tarefa de classificação com o par de pseudopalavras /faba/-/saba/, dos informantes 1 e 2.....	129
6 CAPÍTULO 4 – A PERCEPÇÃO DAS FRICATIVAS CORONAIIS NÃO VOZEADAS [ʃ]-[s] E DAS PLOSIVAS VOZEADAS [b]-[d] POR CRIANÇAS E ADULTOS FALANTES NATIVOS DO PB.....	131
6.1 Metodologia do estudo de ponderação de pistas auditivas.....	131
6.1.1 Sujeitos da pesquisa.....	131
6.1.2 Estímulos, continua e procedimento experimental.....	133
6.1.3 Coleta de dados com as crianças de quatro e de sete anos de idade e com os adultos.....	133
6.2 Plano de análise.....	135
6.3 Resultados e discussão.....	139
7 CONSIDERAÇÕES FINAIS.....	159
8 REFERÊNCIAS BIBLIOGRÁFICAS.....	162
ANEXOS.....	172

1 INTRODUÇÃO

1.1 Contextualização do tema

O tema desta tese é a percepção da fala e, mais especificamente, o desenvolvimento da ponderação de pistas auditivas no mapeamento das ondas sonoras às categorias linguísticas¹.

A percepção da fala é o processo pelo qual somos capazes de interpretar e compreender os sons utilizados nas línguas humanas. Perceber e compreender a fala pode ser considerada uma tarefa complexa, uma vez que o ouvinte deve mapear o sinal acústico de fala variável e contínuo em alvos linguísticos. Além disso, cada som da fala é constituído de diversas propriedades acústicas, ou pistas, que poderão ser utilizadas pelo ouvinte quando ele necessita realizar a distinção entre estes sons.

De acordo com Wright (2004, p. 36) o termo “pista auditiva refere-se à informação do sinal acústico que permite ao ouvinte perceber a existência de um contraste fonológico”. É bem estabelecido que os ouvintes não prestam atenção a todas as informações acústicas disponíveis no sinal de fala. Ao invés disso, os ouvintes prestam mais atenção em algumas pistas auditivas do que em outras, ou seja, eles aprendem que alguns aspectos no sinal acústico de fala são mais importantes em sinalizar um determinado som distintivo do que outros. Desta forma, os ouvintes atribuem diferentes pesos para as diferentes pistas auditivas em uma dada situação. Isto é, cada pista auditiva adquire um grau diferente de importância para a percepção de certo som da língua num contexto específico. A esse fenômeno dá-se o nome de “ponderação de pistas”². Estes padrões parecem mudar com o aumento da experiência linguística e as crianças que estão em fase de aprendizado dos contrastes sonoros de sua língua materna necessitam aprender a detectar quais são as pistas auditivas mais relevantes no sinal acústico de fala para contrastar os sons.

¹ Como o tema desta tese é a percepção da fala, por vezes utiliza-se “percepção” como uma forma mais rápida de “percepção da fala”. Não há nenhuma intenção de tratar do processo de percepção a não ser no âmbito da percepção da fala.

² Tanto quanto pudemos recuperar, o “peso perceptivo” (*perceptual weight*) atribuído por falantes a uma pista foi investigado experimentalmente pela primeira vez por Dorman, Studdert-Kennedy e Raphael (1977) e, no contexto do desenvolvimento, por Morrongiello, Robson, Best e Clifton (1984), ainda sem utilizar o termo “ponderação de pistas”.

Os contrastes compostos pelas fricativas coronais não vozeadas (/f/-vogal e /s/-vogal/) e pelas plosivas vozeadas (/b/-vogal/ e /d/-vogal/), sons investigados neste estudo, acompanhados por uma vogal possuem dois tipos principais de pistas auditivas que são consideradas relevantes diante da necessidade de classificar estes contrastes: uma pista dinâmica e uma pista estática. As pistas dinâmicas referem-se às pistas dispostas no sinal acústico de fala que modificam ao longo do tempo, enquanto as pistas estáticas referem-se às pistas que não mudam significativamente ao longo do tempo. Tanto no caso das fricativas quanto das plosivas, a pista dinâmica relaciona-se a mudança espectral dinâmica que ocorre durante a transição dos formantes entre a consoante e a vogal subsequente, especialmente a transição do segundo formante (F2). Já a pista estática relaciona-se às diferenças na altura da frequência do ruído fricativo quando falamos das fricativas e às diferenças na amplitude da fase de soltura (do inglês, *burst*) quando direcionamos às plosivas vozeadas.

Estudos desenvolvidos com falantes nativos do inglês americano mostram que, para certos contrastes, crianças e adultos pesam as pistas auditivas diferentemente. Nittrouer e colaboradores verificaram consistentemente que, na classificação dos contrastes [s]-vogal e [ʃ]-vogal, as crianças parecem dar mais peso ao início da transição dos formantes das vogais quando comparadas com os adultos e relativamente menos peso que os adultos para as características espectrais do ruído fricativo. Tal estratégia seria utilizada para facilitar a análise do sinal linguístico e para indicar os gestos articulatórios necessários para produzir aquele contraste em sua língua nativa (NITTROUER, 1992; 1996; NITTROUER e MILLER, 1997a; NITTROUER e STUDDERT-KENNEDY, 1987).

Nittrouer, Manning e Meyer (1993) propõem um modelo para explicar essa mudança no desenvolvimento da ponderação de pistas auditivas, denominado “Mudança evolutiva da ponderação de pistas” (*Developmental Weighting Shift* ou DWS). As pesquisadoras afirmam que os pesos atribuídos a vários parâmetros acústicos mudam quando a criança adquire experiência em sua língua nativa, caracterizada pelo uso da língua, e que tal mudança no desenvolvimento da ponderação está relacionada ao aumento da sensibilidade à estrutura fonética que ela adquire ao longo do desenvolvimento.

O modelo DWS baseia-se na premissa de que as crianças processam a fala de maneira mais global, em unidades maiores, tal como as sílabas e as palavras monossilábicas. Já os adultos processam a fala de forma mais analítica, em unidades menores. De acordo com Nittrouer *et*

al. (2000), essa diferença de processamento tem impacto na percepção da fala em termos da atenção que os ouvintes dão às pistas auditivas. Desse modo, crianças jovens prestam mais atenção que os adultos às pistas que delimitam as porções do sinal acústico correspondentes às sílabas; ou seja, à transição dos formantes vocálicos. Outras explicações alternativas para as diferenças na ponderação de pistas auditivas entre crianças e adultos foram propostas, porém como se verá mais adiante elas mostraram-se inválidas (SUSSMAN, 2001; EISENBERG, SHANNON e SCHAEFER *et al.*, 2000).

Ao longo de anos de pesquisa com os sons fricativos, os estudos que investigaram o desenvolvimento da ponderação de pistas auditivas em crianças e adultos indicaram que as crianças mais jovens pesam mais fortemente a pista de transição dos formantes do que as crianças mais velhas e os adultos (NITTROUER, 1992; 1996; NITTROUER e MILLER, 1997a; NITTROUER e STUDDERT-KENNEDY, 1987). Porém, estudos mais recentes desenvolvidos com falantes nativos do holandês, do inglês britânico e do português brasileiro evidenciaram que nem sempre esta tendência está presente diante de contrastes compostos pelos sons fricativos, plosivos e nasais (GERRITS, 2001; MAYO e TURK, 2004; PENIDO, 2012).

O único estudo direcionado à investigação do desenvolvimento da ponderação de pistas auditivas por falantes nativos do português brasileiro (PB) foi realizado por Penido (2012). Neste estudo, investigamos as estratégias de ponderação de pistas das crianças de quatro e sete anos de idade e dos adultos utilizando o par mínimo /aʃaR/-/asaR/ e o contraste /ʃu/-/su/, compostos pelas fricativas coronais não vozeadas [ʃ-s]. As pistas auditivas manipuladas foram a transição do segundo formante vocálico (F2) e a altura da frequência do ruído fricativo, ou seja, uma pista dinâmica e uma pista estática.

Os resultados do estudo de Penido (2012) corroboraram o que é proposto pelo modelo DWS somente com o par mínimo /aʃaR/-/asaR/. Neste caso, as crianças de quatro e de sete anos de idade foram mais influenciadas pela pista dinâmica *transição do F2* quando comparadas com os adultos. Além disso, apenas no grupo das crianças de quatro anos verificou-se que a interação das pistas auditivas foi significativa. Isso quer dizer que, parece que nesta idade, as crianças não utilizam as pistas independentemente. O contraste /ʃu/-/su/ revelou outro padrão de respostas quando comparado com o par mínimo /aʃaR/-/asaR/, onde a pista de altura da frequência do ruído fricativo foi considerada uma pista relevante somente para as

crianças de quatro anos de idade. Desta forma, a frequência do ruído fricativo não foi aparentemente uma pista tão importante para diferenciar as fricativas [ʃ-s] no PB quanto no inglês americano.

Algumas considerações se fizeram importantes no estudo de Penido (2012). No modelo DWS, a preferência pela pista dinâmica ao invés da pista estática pelas crianças mais jovens relaciona-se às diferenças nas estratégias de processamento do sinal de fala, que mudam com o desenvolvimento da linguagem, não com os detalhes fonéticos das pistas auditivas que estão disponíveis nos diferentes contextos vocálicos. Assim, era esperado que os participantes do estudo de Penido (2012) mantivessem o mesmo padrão de respostas tanto diante do par mínimo /aʃaR/-/asaR/ quanto do contraste /ʃu/-/su/ e que este padrão não fosse afetado pelo contexto vocálico. Tal fato não foi observado, uma vez que os adultos ponderaram a transição do F2 mais fortemente do que a altura da frequência do ruído fricativo no contexto vocálico composto pela vogal [u], enquanto as crianças de quatro anos ponderaram a transição do F2 mais fortemente no contexto vocálico composto pela vogal [a]. Acusticamente, a pista de transição do F2 é menos intensa no contraste /ʃu/-/su/ em comparação com o par mínimo /aʃaR/-/asaR/ devido à labialização no contexto vocálico [u] e, então, menos disponível quando comparada com o contexto vocálico composto pela vogal [a]. Desta forma, Penido (2012) verificou que o que emerge neste cenário não é um desenho de desenvolvimento conforme propõe o modelo DWS, da pista dinâmica para a pista estática, mas um padrão onde as crianças de quatro anos de idade são menos estáveis e utilizam qualquer pista auditiva que esteja mais disponível no contexto, de forma isolada ou não. Já os adultos mostram um uso independente e mais consistente das pistas auditivas disponíveis no sinal de fala.

Outra observação importante verificada no estudo de Penido (2012) relaciona-se à categorização das respostas dos participantes, principalmente dos adultos, diante do par mínimo /aʃaR/-/asaR/ e do contraste não lexicalizado /ʃu/-/su/. As respostas dos adultos foram claramente categóricas no par mínimo, diferentemente do contraste não lexicalizado. Este resultado reproduziu o padrão descrito no segundo experimento de Nittrouer e Miller (1997, p.2259) com o par mínimo do inglês "sue-shoe" e com o contraste não lexicalizado "sa-cha". Assim, Penido (2012) relata que é mais provável que a natureza categórica destas respostas venha do estatuto lexical ao invés do contexto vocálico.

Diante do exposto no estudo de Penido (2012) algumas questões necessitam de esclarecimento, tais como (a) as crianças mais jovens modificam as suas estratégias de ponderação de pistas auditivas diante de palavras que constituem par mínimo e de palavras que não constituem par mínimo no PB (pseudopalavras), mantendo o mesmo contexto vocálico (por exemplo, par mínimo /fapa/-/sapa/ e par de pseudopalavras /faba/-/saba/)?; (b) devido ao fato de que as crianças mais jovens possuem pouca experiência com a língua nativa e estão construindo as categorias de sua língua materna, ou seja, estão construindo a fonologia, elas apresentam maior dificuldade em classificar o par de pseudopalavras que não constitui par mínimo no PB?

Este estudo se propôs a estender a investigação das estratégias de ponderação de pistas auditivas de crianças e adultos falantes nativos do PB e examinar as variáveis que podem apresentar influências sobre estas estratégias.

1.2 Justificativa e hipóteses

A investigação da ponderação de pistas auditivas por crianças e adultos não foi objeto de estudos em diversas línguas. A grande maioria dos estudos, como se viu, foi realizada com contrastes do inglês (por exemplo, NITTROUER, 1992; NITTROUER e MILLER, 1997; NITTROUER, 2002; JONES, 2002; MAYO e TURK, 2004, 2005a). Apenas Gerrits (2001) investigou outra língua, no caso o holandês, com resultados parcialmente divergentes das previsões realizadas a partir do inglês. Conforme relatado, no PB, apenas um estudo foi desenvolvido com o intuito de pesquisar este fenômeno (Penido, 2012), onde as pistas auditivas de altura da frequência do ruído fricativo e de transição do F2 foram manipuladas no par mínimo /aʃaR/-/asaR/ e no contraste /ʃu/-/su/.

Apesar de grande parte dos estudos que investigaram as estratégias de ponderação de pistas auditivas utilizarem a mesma metodologia para criação dos estímulos nenhum destes estudos pesquisou o tema utilizando pares mínimos e pseudopalavras com o mesmo contexto vocálico. Direcionando aos estudos realizados com as fricativas coronais não vozeadas observa-se que Nittrouer e colaboradores investigaram o tema em questão utilizando o par mínimo "sue"- "shoe" e os contrastes "sa"- "cha" do inglês americano. Gerrits (2001) utilizou apenas o par mínimo do holandês "sjok"- "sok" em sua investigação. Já Mayo e Turk (2004) realizaram o experimento apenas com contrastes do inglês britânico "si"- "shy", "bay"- "day",

“ta”-“da”, “tea”-“dea”, enquanto Penido (2012) utilizou um par mínimo e um contraste do português brasileiro “achar”-“assar” e “chu”-“su”, porém com contextos vocálicos divergentes e com as fricativas inseridas em posição intervocálica. É por este motivo que neste estudo foram selecionados o par mínimo “chapa”-“sapa” e o par de pseudopalavras “chaba”-“saba”, em posição inicial de palavra seguido do mesmo contexto vocálico, isto é, a vogal [a]. Além das fricativas [ʃ] e [s], o par mínimo composto pelas plosivas vozeadas “bata”-“data” também foi utilizado. Selecionou-se as plosivas bilabial [b] e a alveolar [d] vozeadas, uma vez que estes sons possuem uma variação maior nas amplitudes da fase de soltura, fato que facilita a síntese desses sons. Estes estímulos permitiram uma investigação mais controlada sobre as estratégias de ponderação de pistas auditivas de crianças e adultos e contribuíram para um melhor esclarecimento sobre como as crianças constroem e organizam as categorias fonológicas de sua língua nativa durante o desenvolvimento linguístico.

De modo geral, investigou-se o desenvolvimento da ponderação de pistas auditivas, para a percepção das fricativas coronais não vozeadas e das plosivas vozeadas [b] e [d] por crianças e adultos, falantes nativos do português brasileiro. Antes disso, tornou-se necessário realizar uma investigação e uma descrição detalhada sobre as características acústicas dos estímulos produzidos por meio da síntese de fala.

O presente trabalho pretendeu colaborar para a pesquisa no âmbito das estratégias de ponderação de pistas auditivas de crianças e adultos falantes nativos do PB. As hipóteses deste estudo são:

- As crianças mais jovens, que possuem pouca experiência linguística, pesarão as pistas auditivas disponíveis no sinal acústico de fala de forma diferente dos adultos;
- As crianças da faixa etária de quatro anos de idade serão menos estáveis e utilizarão qualquer pista auditiva que esteja mais disponível no sinal acústico de fala quando elas necessitam classificar as fricativas coronais não vozeadas e as plosivas vozeadas [b] e [d] do PB. Presume-se que apesar das crianças mais jovens utilizarem adequadamente a língua para se comunicar parece que nesta faixa etária dos quatro anos de idade elas ainda não possuem as categorias fonológicas bem definidas;

- Os três grupos de idade apresentarão respostas mais consistentes diante do par mínimo /ʃapa/-/sapa/ do que do par de pseudopalavras /ʃaba/-/saba/, devido à frequência de uso desses pares na língua.

1.3 Objetivo geral

Investigar o desenvolvimento da ponderação de pistas auditivas para a percepção das fricativas coronais não vozeadas ([ʃ] e [s]) e das plosivas bilabial e alveolar ([b] e [d]) por crianças e adultos falantes nativos do português brasileiro.

1.4 Objetivos específicos

- Verificar se as crianças e os adultos apresentam estratégias de ponderação de pistas auditivas semelhantes para a classificação das fricativas coronais não vozeadas e das plosivas vozeadas [b] e [d] diante do par mínimo e do par de pseudopalavras compostos pelo mesmo contexto vocálico;
- Verificar se as crianças e os adultos modificam o padrão de respostas diante do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/ devido à frequência de uso desses pares no PB;
- Investigar se as crianças mais jovens, que possuem pouca experiência linguística, pesam as pistas auditivas disponíveis no sinal acústico de fala de forma diferente dos adultos diante das palavras compostas pelas fricativas e plosivas;
- Investigar se as crianças da faixa etária de quatro anos de idade são menos estáveis e utilizam qualquer pista auditiva que esteja mais disponível no sinal acústico de fala quando elas necessitam classificar as fricativas coronais não vozeadas e as plosivas vozeadas [b] e [d] do PB.

1.5 Estrutura da tese

Esta tese compõe-se de sete seções. Antes de buscar um esclarecimento em relação ao desenvolvimento da ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro tornou-se necessário elucidar questões prévias que surgiram com a realização do estudo de Penido (2012), conforme comentado anteriormente.

Na primeira seção elaborou-se a Introdução, em que se realizou uma breve contextualização do tema central deste estudo com os principais pontos da revisão de literatura e se apresentou o problema formulado pela pesquisa, a justificativa, as hipóteses e os objetivos.

Na segunda seção, desenvolveu-se uma revisão da literatura, constituída de temas e pesquisas relevantes ao objeto de estudo em questão, ou seja, a percepção da fala e a ponderação de pistas auditivas pelas crianças e adultos. Realizou-se um levantamento dos estudos disponíveis na literatura referente à percepção da fala numa perspectiva do desenvolvimento da linguagem. Além disso, abordou-se o desenvolvimento fonológico da criança com foco no desenvolvimento das categorias fônicas, bem como estudos relacionados à categorização dos sons da fala, incluindo a percepção categórica e os experimentos comportamentais para investigação do tema. Por fim, apresentou-se uma explanação das pesquisas direcionadas a resolução da seguinte questão: as crianças utilizam as mesmas informações do sinal acústico de fala que os adultos para classificar os sons da fala? Cabe ressaltar que parte da revisão de literatura foi publicada na dissertação de mestrado desenvolvida em 2012 e foi mantida nesta tese por ser considerada um conteúdo composto por definições básicas para o entendimento do tema de percepção da fala. No entanto, nesta tese explorou-se de forma mais profunda as publicações direcionadas à investigação da ponderação de pistas auditivas, bem como as principais hipóteses para explicar as diferenças nos resultados entre crianças e adultos.

Na terceira seção, descreveu-se o capítulo 1, na qual realizou-se uma investigação minuciosa sobre as propriedades acústicas dos estímulos utilizados no experimento de ponderação de pistas auditivas (/fapa-/sapa/, /faba-/saba/ e /bata-/data/).

Na quarta seção, desenvolveu-se o capítulo 2, que envolveu uma descrição sobre a técnica de síntese de fala utilizada para construir os estímulos.

Após a descrição das propriedades acústicas dos estímulos e da técnica de síntese de fala utilizada para construí-los, elaborou-se o terceiro capítulo onde se desenvolveu um estudo piloto com o objetivo de avaliar a qualidade dos estímulos sintetizados.

Na sexta seção, analisou-se as estratégias de ponderação de pistas de crianças e adultos diante dos pares de palavras compostos pelas fricativas [ʃ] e [s] e pelas plosivas [b] e [d] do PB.

Por fim, na sétima seção, formulou-se as conclusões e as considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 A aquisição da linguagem

As pesquisas sobre aquisição da linguagem constituem uma área multidisciplinar, na confluência entre as abordagens linguística, cognitiva e sociointeracionista ou pragmática, uma vez que o início e o desenvolvimento da aquisição da linguagem dependem de diversos fatores (ROTHER-NEVES e VITOR, 2008).

Os pesquisadores que investigam esta área têm por objetivo central descobrir como ocorre o processo de aquisição da linguagem. Ou seja, “de que modo o ser humano parte de um estado no qual não possui qualquer forma de expressão verbal e, naturalmente, incorpora a língua de sua comunidade nos primeiros anos de vida” (CORREA, 1999; p. 339).

Correa (1999) explica que o material empírico de que estas pesquisas dispõem são dados da produção, percepção e compreensão de enunciados linguísticos por crianças, obtidos em condições naturais ou experimentais.

Os estudos direcionados à aquisição da linguagem seguem uma trajetória que percorre diferentes modelos teóricos para a fundamentação da análise dos dados. De modo geral, pode-se dizer que a Fonologia, ao longo dos anos, preocupou-se muito mais com os dados da produção da fala do que com os da percepção. Este fato pode ser evidenciado nos vários modelos fonológicos fornecidos pelas teorias ao longo do tempo.

Na Fonologia, as investigações basearam-se na Fonologia Gerativa, passando pela Fonologia Natural, e desta para a Fonologia Autossegmental. Todos esses modelos teóricos, considerados mais tradicionais, têm como pressuposto que o processo de aquisição fonológica se dá a partir da aplicação de regras pelo aprendiz e que tal processo ocorre de forma derivacional. Dessa forma, o falante teria uma representação subjacente (*input*) e, a partir da aplicação de uma série de regras, chegaria à realização fonética (*output*). Uma das principais vertentes de investigação teórica do desenvolvimento fonológico é a Teoria da Otimidade. Além desta, existem as abordagens mais recentes tais como a Fonologia de Uso, a Teoria de Exemplos e a Fonologia Probabilística.

2.1.1 Teorias linguísticas de aquisição da linguagem

A partir do final da década de 1950, os estudos de Noam Chomsky estimularam os trabalhos sobre aquisição da linguagem. A teoria gerativa, desenvolvida por ele, defende que o ser humano é provido de uma gramática inata, denominada “Gramática universal” (GU), que contém todas as regras de todas as línguas. A criança toma como base para seu desenvolvimento a fala dos adultos, que serve de estrutura para o desenvolvimento de suas próprias regras (CHOMSKY, 1957).

Chomsky propôs que a criança possui um dispositivo de aprendizagem da linguagem (DAL) inato, que é ativado e se processa a partir de sentenças (*input*), o qual produz como resultado a língua a que a criança está exposta. Este dispositivo é formado por uma série de regras, e a criança, em contato com a língua a que é exposta, seleciona as regras que funcionam para aquela língua, desativando as que não funcionam.

Em 1981, em desenvolvimento à teoria gerativa, surge o modelo de princípios e parâmetros, em que se nota um distanciamento em relação à tradição da gramática gerativa dos primeiros momentos, devido à rejeição do conceito de regra. Esta teoria, também apresentada por Chomsky, modifica a concepção que se tinha de gramática universal. De acordo com esta versão da teoria, a gramática universal é formada por princípios, que são utilizados igualmente por todas as línguas, e parâmetros, que possuem representações na língua em que a criança encontra-se inserida. Então, cabe à criança escolher, com base no *input*, o valor que o parâmetro deve apresentar.

Em 1993, Prince e Smolensky e McCarthy e Prince propõem uma nova teoria linguística, a “Teoria da otimidade” (do inglês, *Optimality Theory* – OT). Este novo modelo teórico postula a existência de uma gramática universal (GU), isto é, o conhecimento inato da linguagem por todos os seres humanos e, ainda, um conjunto de restrições (CON, do inglês *constraints*), compartilhado pelas línguas do mundo. O que diferencia uma língua da outra é o ranqueamento das restrições, que mostra por que determinados padrões são permitidos em uma língua e proibidos em outras (BONILHA, 2003).

Na Teoria da Otimidade há o pressuposto de que as crianças nascem com restrições universais, as quais serão ranqueadas gradativamente, de acordo com a hierarquia de sua

língua materna. Esta teoria não se baseia em regras, mas em restrições universais. Por fim, as abordagens mais recentes, a Fonologia de Uso e a Fonologia Probabilística, sugerem que o conhecimento linguístico é baseado no uso, ou seja, na experiência com a língua à qual o indivíduo é exposto, e é gerenciado probabilisticamente em conexões organizadas em redes. Apesar destas teorias serem consideradas versáteis e poderem lidar com os dados tanto da percepção da fala quanto da produção, não é o objetivo desse estudo apresentar uma revisão aprofundada desse tema, que pode ser encontrada em Escudero e Boersma (2001), Bonilha (2003), Boersma (2009), Bybee (2001) e Pierrehumbert, (2003)³.

2.2 O desenvolvimento fonológico da criança: a percepção

O interesse pelo estudo da percepção de fala tem sido cada vez maior, devido ao reconhecimento de sua importância no processo de desenvolvimento da linguagem da criança. O processo gradual de aquisição dos padrões de fala de um adulto pela criança é denominado “desenvolvimento fonológico”. De acordo com a colocação de Santos (2008, p. 465):

Sempre que refletimos em relação a este desenvolvimento, vêm-nos à mente os sons que a criança vai utilizar para veicular o que ela sabe, pensa ou quer. No entanto, a fonologia de uma língua não deve ser reduzida aos sons e nem à sua aquisição, portanto; deve também estar relacionada à percepção do conjunto de fonemas de uma língua. Então, o que a criança deve “aprender”? Ou o que deve fazer parte de seu conhecimento fonológico (SANTOS, 2008)?

Para adquirir a competência fonológica adulta, as crianças em desenvolvimento devem aprender a discriminar os padrões sonoros da língua e organizá-los em categorias fônicas apropriadas (HANZAN e BARRET, 2000). Além disso, a criança descobre os padrões de combinação dos segmentos em unidades, estruturas maiores, as regras fonotáticas e fonológicas que dão forma às palavras de sua língua. Para além dos segmentos, as crianças também precisam descobrir as propriedades prosódicas de sua língua e como elas se organizam. No entanto, para conseguir organizar todo este conhecimento, as crianças devem descobrir pistas naquilo que ouvem (SANTOS, 2008).

³ Apesar desta tese tratar do tema de percepção da fala, o presente estudo não possui como objetivo testar os pressupostos das teorias de percepção da fala, tais como a Teoria Motora, a Teoria Realista Direta e a Abordagem Auditiva Geral. Uma revisão mais detalhada destas Teorias pode ser encontrada em Klatt (1989), Miller e Eimas (1995), Diehl, Lotto e Holt (2004), Holt e Lotto (2008) e Samuel (2011).

De acordo com Pujol, Lavigne-Rebillard e Uziel (1991), o processo de aquisição fonológica tem início muito antes de a criança emitir os primeiros sons. O sistema auditivo humano é funcional antes dos seis meses de gestação (22^a a 24^a semana gestacional) e evolui durante o último trimestre de gestação, em ambientes com a presença de sons para os quais o feto responde comportamental, elétrica e neuroquimicamente (LECANUET, 1998). Querleu, Renard, Versyp *et al.* (1988) desenvolveram experimentos em que microfones foram implantados no lado externo da parede uterina de mulheres grávidas, tendo apurado que seu ambiente acústico permite aos bebês que escutem sons ambientais. Em pesquisas análogas realizadas em ovelhas, outros pesquisadores observaram que, dentro do útero, ouve-se apenas a estrutura prosódica (rítmica e entoacional), mas não os segmentos (HUANG, GERHARDT, ABRAMS *et al.*, 1997).

Segundo Flavell, Miller e Miller (1999), as perguntas que interessam quanto à percepção da fala em bebês são de dois tipos: “O que eles gostam de ouvir?” e “O que eles escutam quando ouvem?”. Nas três últimas décadas, muitas pesquisas foram realizadas com o intuito de esclarecer estes questionamentos (para detalhes metodológicos, ver a bibliografia citada nas referências). Desde o nascimento, o bebê humano parece estar biologicamente pré-sintonizado e predisposto a processar os sons da fala. Além disso, parece preferir o som da fala a outros sons, mas principalmente pela fala materna (“maternês”), caracterizada pelo seu tempo lento, tom alto e entonação muito exagerada (FERNALD e SIMON, 1984; GRIESER e KUHL, 1988).

A partir da investigação inicial realizada por Eimas, Siqueland, Jusczyk *et al.* (1971), um grande número de estudos deixou claro que bebês de aproximadamente um mês de vida conseguem discriminar muitos dos contrastes fonéticos utilizados nas línguas do mundo. No entanto, essa habilidade de discriminação não constitui evidência de que os bebês possuem um reconhecimento segmental (a habilidade de identificar os elementos fonéticos da língua). No mínimo, os bebês conseguem discriminar os contrastes acústicos fundamentais ao reconhecimento segmental.

Jusczyk, Pisoni e Mullenix (1992) verificaram que bebês entre um e quatro meses de idade conseguem normalizar diferentes falantes e diferentes ritmos de fala. Ou seja, eles conseguem detectar a mesma vogal quando falada por diferentes falantes, que produzem a vogal com diferentes propriedades acústicas, e podem reconhecer o mesmo elemento

quando falado em diferentes ritmos de fala. Os bebês em torno desta idade também conseguem detectar variações nos padrões entoacionais e reconhecer a mesma sílaba em diferentes enunciados.

Jusczyk, Houston e Newsome (1999) mostraram, ao investigar bebês americanos de sete meses, que muito antes de produzirem os primeiros sons as crianças já identificam padrões acentuais em sua língua. Além disso, utilizam traços prosódicos para distinguir palavras estrangeiras de palavras pertencentes à língua materna. Entretanto, dos sete aos dez meses de idade ocorre uma diminuição desta habilidade de detectar determinados contrastes estrangeiros, o que indica que o bebê desenvolveu alguns aspectos de um reconhecimento fonético específico da língua (WERKER e TESS, 1984; KUHL, 2004). Segundo estes pesquisadores, dos dez aos doze meses de idade o bebê pode ter reorganizado as categorias perceptivas para se adequar à estrutura fonética de sua língua materna.

Em resumo, pode-se dizer que, do ponto de vista da percepção da fala, uma criança não aprende uma língua, ela esquece todas as outras. Conforme relatado nesta seção, durante a fase inicial do desenvolvimento da linguagem, as crianças deverão não somente *aprender* a discriminar os padrões sonoros de sua língua materna, mas também a *organizá-los* consistentemente em suas categorias fonêmicas. Esta habilidade é denominada categorização fonêmica e será descrita a seguir.

2.3 A categorização dos sons da fala

A categorização dos sons da fala é considerada um processo cognitivo fundamental (GOUDBEEK, 2006). De acordo com a colocação de Goudbeek (2006):

Sem categorias, as pessoas seriam incapazes de reconhecer as diferentes cores do arco-íris, apreciar um tom musical, falar sobre tipos de animais no zoológico e reconhecer amigos ou ler a escrita de alguém. Em outras palavras, a categorização está presente em todas as situações onde experiências prévias orientam suas interpretações presentes. Compreender a linguagem falada é um excelente exemplo de tal situação, uma vez que envolve a categorização de múltiplos níveis, que vão do reconhecimento de consoantes e vogais até a interpretação de estruturas gramaticais e do contexto. (GOUDBEEK, 2006, p. 15, tradução nossa).

Goudbeek (2006) relata que a aquisição das categorias dos sons da língua pode ocorrer em duas situações: 1^a) quando o bebê adquire sua primeira língua, sem qualquer categoria presente; e 2^a) diante do aprendizado de uma língua estrangeira, enquanto já existe pelo menos uma língua.

Os bebês têm que aprender a categorizar os sons de entrada em categorias sonoras de sua língua nativa. Já os aprendizes de uma nova língua lidam com o problema de que as categorias sonoras de sua nova língua podem ser diferentes de sua língua nativa. Assim, eles têm que tentar integrar as novas categorias sonoras em um sistema existente. Apesar das diferenças possíveis entre a situação do bebê e do aprendiz de segunda língua, Goudbeek (2006) relata em seus estudos que alguns dos processos subjacentes são os mesmos. A tarefa que eles têm que enfrentar é a mesma. Ambos têm que reconhecer padrões estatísticos no estímulo, o que varia em muitas dimensões relevantes.

O conhecimento acumulado sobre a percepção da fala ressalta o desenvolvimento da criança como um importante aspecto na competência perceptiva (MONTGOMERY e CLARKSON, 1997; OHDE, HALEY e McMAHON, 1996). Já ao nascer, o bebê possui a capacidade de discriminar contrastes consonantais.

Durante o primeiro ano de vida, a acuidade dessa habilidade declina para os sons que não são distintivos no contexto linguístico da criança. Ela passa a discriminar seletivamente os sons distintivos da língua a que é exposta, fazendo uma reorganização perceptual, em um claro processo de adaptação à solicitação do meio em que vive (JENKINS, 1992; KUHL, 1983, 1987, 2004). Nesta adaptação está incluído o processo de agrupamento de fones inicialmente diferentes, que passam a ser percebidos como sendo equivalentes. Este fenômeno é conhecido como “categorização”.

Goudbeek (2006) define duas tarefas que há que se distinguir, uma vez que serão abordadas posteriormente: 1^a) identificação ou classificação – consiste no mapeamento de um a um dos estímulos, que são gravados e apresentados, um de cada vez, aos ouvintes, e estes são solicitados a classificar cada estímulo apresentado com base em um rótulo escrito ou pictórico; e 2^a) discriminação – consiste na interpretação de dois ou mais pares de estímulos como iguais ou diferentes. Se o ouvinte for incapaz de discriminar dois estímulos, não poderá atribuí-los a duas categorias diferentes. Este processo é denominado “percepção

categórica” e relaciona-se a falta de discriminabilidade entre os sons que pertencem a mesma categoria.

2.3.1 A percepção categórica

Nesta seção, apresenta-se um excuro sobre o tema da percepção categórica, o qual não segue linearmente o que até agora se expôs. Servirá, entretanto, para esclarecer conceitos e, principalmente, procedimentos metodológicos importantes para a discussão que se segue.

Grande parte do trabalho sobre categorização da fala está preocupada com este fenômeno de percepção categórica, especialmente com a capacidade ou incapacidade dos ouvintes de perceber diferenças entre os sons da fala que pertencem à mesma categoria fonêmica (HARNAD, 1987; LIBERMAN, HARRIS, HOFFMAN, *et al.*, 1957; REPP, 1984). É característico dos sistemas perceptivos humanos agrupar estímulos em categorias cognitivamente eficientes, de modo a facilitar o armazenamento e a evocação de informação. O mecanismo complexo de percepção categórica capacita o indivíduo a reconhecer sons distintivos consistentemente, a despeito de uma grande variação em parâmetros acústicos cruciais.

Como resultado de décadas de pesquisa, identificou-se a percepção categórica com um paradigma experimental particular. No “paradigma de percepção categórica”, um estímulo que faz parte de um continuum entre os sons da fala é apresentado aos ouvintes em duas tarefas psicoacústicas: uma de discriminação e uma de classificação. Acreditava-se que os resultados de percepção categórica demonstravam forte relação entre o desempenho dos ouvintes nestas duas tarefas: os estímulos de fala classificados como pertencentes à mesma categoria são mais difíceis de discriminar; e os estímulos da fala classificados como pertencentes a categorias diferentes são mais fáceis de discriminar. Assim, o desempenho na discriminação é previsível a partir do desempenho na classificação.

O primeiro estudo de percepção categórica foi realizado por Liberman, Harris, Hoffman e Griffith (1957), nos Laboratórios Haskins (LIBERMAN, HARRIS, HOFFMAN *et al.*, 1957; LIBERMAN, HARRIS, KINNEY *et al.*, 1961). Estes pesquisadores investigaram a classificação e a discriminação dos ouvintes de um estímulo *continuum* entre consoantes oclusivas /bV - dV - gV/. Liberman, Harris, Hoffman *et al.* (1957) não observaram

concordância entre os resultados obtidos e a hipótese testada: os resultados da discriminação foram melhores que o previsto pela classificação. Este achado aparentemente representou a capacidade do ouvinte de distinguir os sons da fala com base não apenas na categorização fonêmica, mas também nas diferenças acústicas entre um estímulo e outro. Em que pesem as conclusões dos pesquisadores, o primeiro estudo é frequentemente citado como um exemplo paradigmático de percepção categórica (REPP, 1984) e a correlação significativa entre a discriminação prevista (classificação) e a discriminação observada foi tomada como uma importante evidência deste fenômeno.

Comparar os resultados das tarefas de classificação e de discriminação fonêmica – “percepção categórica” – de um estímulo contínuo foi por muito tempo considerado um método útil para investigar o armazenamento e a recuperação das categorias fonêmicas da memória de longo prazo. No entanto, Schouten, Gerrits e Van Hessen (2003) observaram que a proximidade da relação entre as duas tarefas – ou seja, o grau de percepção categórica – depende de uma série de fatores, alguns dos quais são desconhecidos ou aleatórios.

Gerrits (2001) realizou diversos experimentos com o intuito de verificar se existe uma forte relação entre a discriminação e a classificação dos ouvintes ao longo de um estímulo contínuo entre dois sons da fala (“Hipótese de percepção categórica”). A pesquisadora esperava que os ouvintes fossem relativamente piores para detectar diferenças entre estímulos pertencentes à mesma categoria fonêmica, uma vez que estes continham diferenças acústicas que não transmitiam um significado e, portanto, eram irrelevantes para a compreensão da fala. De outro lado, esperava que ouvintes fossem relativamente melhores para detectar diferenças entre estímulos pertencentes a categorias fonêmicas diferentes.

Gerrits (2001) concluiu que a hipótese de percepção categórica deve ser revisada. Ela propõe que não ocorre uma relação estreita entre a classificação e a discriminação dos sons da fala, e sim uma grande variabilidade no grau em que a tarefa de classificação prediz a discriminação. Parecem existir efeitos marcados de fatores processuais diversos, tais como: natureza e qualidade dos estímulos, intervalo entre estímulos, sujeitos e, principalmente, tarefa de discriminação.

Conforme revisado por Silva e Rothe-Neves (2009), uma vez que as tarefas de discriminação e classificação são baseadas em medidas comportamentais, torna-se

necessário realizar uma análise específica dos experimentos, com o intuito de separar o comportamento perceptivo dos processos de atenção e decisão do sujeito. Estas tarefas, consideradas experimentos comportamentais, serão descritas brevemente a seguir.

A tarefa de discriminação padrão entre os estudos de percepção categórica é a tarefa ABX, na qual o ouvinte é solicitado a indicar se o estímulo X é idêntico ao estímulo A ou ao estímulo B. Ao fazê-lo, ele pode recorrer tanto aos traços auditivos quanto às etiquetas fonêmicas atribuídas para os sons A e B. No entanto, Massaro e Cohen (1983) ressaltam que esta tarefa apresenta um viés em relação aos resultados obtidos a partir da percepção categórica. Em decorrência da duração relativamente pequena dos traços acústicos na memória auditiva do ouvinte no momento em que o estímulo X é apresentado, o sujeito tenta comparar os traços auditivos deste com os estímulos de A e B, e estes traços já podem ter decaído, de forma que o ouvinte é obrigado a recorrer apenas às classificações atribuídas a A e a B. Tal fato produziria resultados voltados à percepção categórica.

Massaro e Cohen (1983) relatam que a tarefa AX tem por objetivo evitar estratégias que dependam exclusivamente da etiquetagem fonêmica, uma vez que reduz a carga na memória auditiva. Nesta tarefa, uma combinação de estímulos é apresentada: AA, AB, BB, BA. O ouvinte é solicitado a indicar se os estímulos são iguais ou diferentes. A desvantagem desta tarefa é que quando questionados se os dois estímulos são diferentes ou iguais os sujeitos podem decidir responder “diferente” somente se estiverem certos de suas decisões. Isso indica que a tarefa AX não está livre de viés, uma vez que a resposta dos sujeitos é determinada por um critério subjetivo.

Outra tarefa utilizada em experimentos psicofísicos é a de realizar uma escolha forçada entre duas alternativas, ou 2AFC (do inglês, *two alternatives forced choice*), em que dois estímulos diferentes são apresentados (AB ou BA), e o sujeito deve determinar a ordem na qual eles são apresentados. Esta tarefa é frequentemente utilizada em experimentos direcionados à percepção da intensidade. O sujeito ouve dois estímulos que variam em intensidade e responde indicando a ordem em que eles foram apresentados: intenso-fraco ou fraco-intenso. Gerrits (2001) resalta que no caso de estímulos de fala torna-se necessário explicar aos sujeitos o que termo *ordem* significa e mencionar as categorias fonêmicas nas instruções. Tal fato é analisado pela pesquisadora como um risco para encorajar o comportamento classificatório pelo ouvinte.

A tarefa AX de quatro intervalos, ou 4IAX (do inglês, *four intervals AX*), é considerada mais sensível às diferenças acústicas entre os estímulos de fala. O julgamento consiste em quatro combinações possíveis: ABAA, BAAA, AAAB e BABB. O intervalo temporal entre o segundo e o terceiro estímulo é mais longo. O sujeito deve indicar qual dos dois pares apresenta estímulos idênticos: o primeiro ou o segundo par. Supõe-se que os ouvintes determinem as diferenças entre o estímulo de cada par e, em seguida, qual das duas diferenças é a menor.

A tarefa 4I2AFC (do inglês, *four intervals two alternatives forced choice*) é semelhante à 4IAX, porém é considerada menos trabalhosa. Neste caso, somente quatro sequências são possíveis: AABA, ABAA, BBAB e BABB. O ouvinte deve indicar se o estímulo díspar está na segunda ou na terceira posição. Assim, o sujeito poderia ignorar o primeiro e o quarto estímulo, desempenhando uma tarefa 2AFC. Estes estímulos ignorados poderiam ser utilizados como referência, de modo a tornar possível uma decisão baseada apenas nos traços sensoriais. Nesta tarefa, a decisão do sujeito é a menos sensível a critérios subjetivos.

Gerrits (2001) ressalta que o desempenho na discriminação não mostrará relação com aquele da classificação, quando a tarefa de discriminação incita os ouvintes a comparar os traços auditivos do estímulo. De outro lado, se a tarefa de discriminação utilizada leva os ouvintes a realizar a classificação fonêmica para discriminar o estímulo, o desempenho da discriminação mostrará uma relação estreita com o desempenho da classificação.

A partir da literatura discutida, conclui-se que a relação entre os resultados da discriminação e os da classificação simplesmente refletem o grau pelo qual os ouvintes empregam os mesmos processos perceptivos durante estas tarefas. Neste estudo, será realizada apenas a tarefa de classificação, em que os sujeitos escutarão os estímulos, por meio dos fones de ouvido, e decidirão qual categoria é mais coerente com o que foi apresentado dentre as possibilidades existentes (Ex.: /ʃapa/-/sapa/). Esta tarefa induz a classificação fonêmica e, conseqüentemente, produz a categorização, que é um fenômeno que pode ser provocado e, portanto, pode ser utilizado para esclarecer sobre os estímulos e as pistas auditivas que o provocam. Conforme colocado por Silva e Rothe-Neves (2009), a utilização de técnicas que induzam o comportamento classificatório é útil, uma vez que “permite identificar por meio de quais pistas auditivas é possível provocar nos falantes de uma dada língua a ativação da – imagem mental – ou a representação de uma dada categoria sonora” (p. 326).

2.3.2 O desenvolvimento das categorias fônicas

De volta ao tema central do desenvolvimento linguístico, uma das questões mais importantes da percepção da fala diz respeito ao modo como os ouvintes começam a perceber os sons de uma forma particular para sua língua nativa. Para se comunicar eficientemente, um ouvinte tem que discriminar a variação acústica do sinal de fala linguisticamente relevante e generalizar toda variação que é irrelevante. O que conta como relevante e irrelevante depende do inventário fonético da língua específica.

Estudos realizados nos últimos quarenta anos mostram que bebês são aparentemente capazes de distinguir todas (ou quase todas) as diferenças fonéticas encontradas nas línguas e, então, fazê-las por processos que rendem representações categóricas organizadas (EIMAS, SIQUELAND, JUSCZYK *et al.*, 1971; EIMAS e MILLER, 1992).

Para as crianças, assim como para os adultos, o mapeamento entre sinais acústicos e representações categóricas não é invariável. O limite entre as categorias infantis pode ser alterado pela variação de fatores contextuais. E, da mesma forma, as representações categóricas das crianças são especificadas por várias propriedades do sinal acústico (EIMAS, 1985; MILLER e EIMAS, 1983). Além disso, aos seis meses de idade as crianças podem formar classes de equivalência, apesar da variação acústica de diferenças nos falantes, da entonação e do contexto (KUHL, 1983). Esses fenômenos mostram que as crianças, assim como os adultos, escutam com base na variação natural presente na fala. Esta habilidade é necessária para que os bebês desenvolvam seu repertório de constituintes básicos da linguagem. Aos doze meses de idade, os processos de percepção fornecem representações que são adequadas para a descoberta e a construção das palavras e dos componentes sintáticos. De acordo com Werker (1993), Best (1994) e Kuhl (2004), a experiência com a língua nativa modifica a percepção fonética durante os primeiros anos de vida.

Werker e Lalonde (1988) investigaram esta questão em estudos realizados com bebês e adultos ingleses e indianos e como eles percebem variações no som de [d]. Para os falantes do hindi, a diferença entre o som [d] na palavra inglesa *doll* versus o som [d] em *your doll* – um contraste fonético entre uma dental [d] versus uma retroflexa [d] – sinalizaria duas categorias distintas. No inglês, ambos os sons [d] sinalizam somente uma forma possível –

foneticamente classificada como a alveolar [d]. Os bebês que tinham o inglês como língua materna discriminaram facilmente os dois sons [d] da língua hindi, enquanto os adultos falantes do inglês apresentaram maiores dificuldades nesta tarefa. Aos dez meses de idade, os aprendizes do inglês começaram a perceber a fala de acordo com sua língua nativa e não apresentaram o mesmo desempenho para discriminar as duas realizações do som [d] no hindi, enquanto os bebês expostos ao hindi mantiveram perceptivamente esta distinção. Esse resultado favorece a ideia de que os bebês apresentam a capacidade de discriminar um conjunto universal de contrastes fonéticos das línguas e de que existe um aparente declínio ou reorganização nessa sensibilidade fonética universal, em função da aquisição de uma língua em particular. Apesar disso, a pesquisadora reforça que o desenvolvimento da percepção da fala não é tão simples quanto parece⁴.

Infelizmente, existem poucos estudos que têm investigado a percepção de fala de crianças entre um e quatro anos de idade (por exemplo, BARTON, 1980; SVACHKIN, 1973). Os resultados destes estudos indicam que as crianças mais velhas parecem ter mais dificuldade para discriminar diferenças entre sílabas e palavras que bebês. Ressalte-se que existem algumas diferenças metodológicas importantes nos experimentos realizados com bebês, crianças e adultos. Em primeiro lugar, diferentes procedimentos de teste são utilizados e os processos perceptuais dos sujeitos diferem essencialmente de um para outro. Em estudos com bebês, a tarefa de discriminação é realizada com palavras ou sílabas, diferindo em um segmento (por exemplo, ba-ba-ba-ba-da-da-da). Os bebês são treinados a fazer um movimento com a cabeça quando o estímulo muda (*conditioned head turn procedure*, WERKER; POLKA e PEGG, 1997). Já a discriminação com ouvintes adultos normalmente envolve a comparação de dois ou mais estímulos apresentados em uma mesma série ao longo de um *continuum*. Os ouvintes adultos são solicitados a indicar se os dois estímulos são “idênticos” ou “diferentes” ou detectar qual estímulo difere dos demais na série.

A capacidade precoce dos bebês de detectar as diferenças acústicas está supostamente relacionada às representações das categorias fonêmicas em sua memória de longo prazo, porque eles simplesmente ainda não desenvolveram tais representações. No caso de ouvintes

⁴ Ver Aslin, Werker e Morgan (2002) para uma crítica de suas alegações de que cientistas que investigam o desenvolvimento apresentam uma interpretação muito simplificada da percepção da fala dos bebês.

adultos, a hipótese de percepção categórica prevê que a discriminação é principalmente baseada no processo de categorização fonêmica. Para crianças de um a quatro anos de idade, a história é novamente diferente. A resposta de discriminação utilizada com crianças, geralmente, envolve a escolha entre um dos dois objetos que têm um nome sem sentido ou representam palavras familiares, diferindo somente em um som distintivo. Esta tarefa é idêntica à 2AFC (*two alternatives forced choice*), escolha forçada com dois intervalos e duas alternativas, em pesquisas realizadas com adultos, e não analisa em que medida as crianças são capazes de discriminar entre os sons da fala. Isso implica que estes estudos de discriminação realmente relatam a capacidade da criança de classificar ou reconhecer palavras, o que poderia ser mais difícil que detectar diferenças entre dois sinais acústicos, que não apresentam sentido e nem representações lexicalizadas.

Um estudo holandês desenvolvido por Stoep e Verhoeven (2001) sobre o desempenho da classificação de 445 crianças de cinco anos de idade mostrou que existem diversas distinções fonéticas nas quais elas apresentaram mais que 20% de erro. Uma tarefa de seleção de figuras foi realizada com pares mínimos familiares à criança. Novamente, é uma tarefa de classificação, e não de discriminação. A porcentagem de erro para /b-p/ e /d-t/ foi tão alta quanto 43,1% e 32,7%, respectivamente. Esse resultado indica que crianças de aproximadamente quatro e cinco anos de idade ainda não adquiriram completamente um sistema que lhes permite classificar os sons de sua língua nativa.

Outras pesquisas apontaram que isso ocorre até mesmo com crianças de quatro a doze anos de idade. Nestes estudos, a tarefa de classificação tradicional é utilizada para avaliar a percepção categórica: uma simples pista auditiva é sistematicamente variada ao longo de um *continuum* entre dois sons distintivos da língua (KRAUSE, 1982; KUIJPERS, 1996; MANN, SHARLIN e DORMAN, 1985). Estes resultados evidenciaram diferenças significativas relacionadas à idade, voltadas à fronteira fonêmica e ao declive da curva de classificação. Artigos mostram que o desempenho na classificação de crianças é similar ao de adultos entre dois e seis anos de idade (por exemplo, WERKER e POLKA, 1993). Outros afirmam que somente entre as idades de dez e doze anos a fronteira fonêmica e o declive das funções de classificação tornam-se como nos adultos (BURNHAM, EARNSHAW e CLARK, 1991; KUIJPERS, 1996). É provável que essa discrepância entre as idades de aquisição reflita o fato de que alguns sons distintivos são adquiridos mais cedo que outros.

Em suma, diversos estudos auxiliam no esclarecimento de questões direcionadas ao desenvolvimento da percepção da fala e das categorias fônicas, tais como: bebês, crianças e adultos utilizam as mesmas informações do sinal acústico para identificação dos sons da fala? Ainda assim, existe um longo caminho a percorrer com o intuito de alcançar um quadro completo do desenvolvimento dessas habilidades.

A seguir, será abordado o tema central de investigação deste estudo - a ponderação de pistas auditivas por crianças e adultos. Por fim, será realizada uma análise sobre as hipóteses disponíveis no meio científico que buscam explicar as diferenças nas estratégias de ponderação de pistas entre crianças e adultos.

2.4 A ponderação perceptiva de pistas auditivas

Conforme descrito nas seções anteriores, o sinal acústico fornece diversas informações auditivas aos ouvintes quando eles necessitam discriminar os sons da fala. Por exemplo, o contraste /ba/-/da/ fornece duas principais pistas auditivas que são a amplitude da fase de soltura da oclusão (*burst*) e os valores iniciais das frequências dos formantes da vogal que acompanha a plosiva. Estas pistas encontram-se disponíveis ao longo do tempo.

A maioria dos investigadores interessados na percepção de fala, durante a segunda metade do século XX, direcionou os estudos para os correlatos acústicos dos segmentos fonéticos, ou seja, as pistas auditivas.

Alguns investigadores acreditavam que um conjunto de propriedades do espectro seria encontrado no sinal de cada elemento linguístico e seria o mesmo ao longo de todos os contextos fonéticos (STEVENS e BLUMSTEIN, 1978). Outros postularam que regiões espectrais específicas (por exemplo, formantes individuais e transição dos formantes) seriam encontradas no sinal de cada elemento linguístico (MANN e REPP, 1980). Infelizmente, ambas as linhas de investigação têm falhado em explicar como é que ouvintes derivam a estrutura fonética do sinal acústico de fala.⁵

⁵ Cabe ressaltar que nas décadas de 70 e 80 o processamento digital dos sinais de fala era mais primitivo em comparação com o processamento digital realizado atualmente. Os pesquisadores não dispunham de tecnologia avançada para criar os estímulos utilizados no experimentos de percepção da fala.

Nos últimos 30 anos, observa-se um crescimento dos estudos direcionados a investigação do desenvolvimento da ponderação de pistas por crianças e adultos (MORRONGIELLO, ROBSON, BEST *et al.*, 1984; NITTROUER, 1992, 1996, 2002). É bem estabelecido que ouvintes não prestam atenção em todas as pistas auditivas disponíveis no sinal acústico de fala. Ao invés disso, os ouvintes prestam mais atenção em algumas pistas em relação às outras, o que caracteriza o processo de ponderação de pistas. Ponderar uma pista auditiva significa atribuir valor ou dar importância a determinada pista relativamente às outras presentes no sinal. A aquisição de um esquema de ponderação apropriado é essencial para a classificação categórica e, então, para a compreensão da fala.

Estudos mostram que o padrão de ponderação de pistas auditivas muda ao longo do desenvolvimento, ou seja, as crianças possuem estratégias de ponderação de pistas auditivas diferentes dos adultos. Nittrouer e colaboradores verificaram que para a classificação do contraste fricativo “sue” e “shoe”, do inglês americano, as crianças mais jovens atribuem maior importância à pista de transição dos formantes do que as crianças mais velhas e os adultos (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 1992; NITTROUER e MILLER, 1997a).

Basicamente, os experimentos de percepção da fala envolvem a manipulação de múltiplas propriedades acústicas. Usualmente, uma pista é manipulada ao longo de um *continuum*, de uma categoria fonética à outra, tal que as duas categorias formam um par mínimo ou um contraste. Uma segunda pista é manipulada dicotomicamente. O estímulo resultante é oferecido aos ouvintes em tempos múltiplos. A tarefa dos ouvintes é executar uma decisão de categorização, de escolha binária. O ouvinte que classifica os dois continua de maneira idêntica não é influenciado pela pista que muda entre os continua.

Com relação ao desenvolvimento, as questões que vêm sendo investigadas por estes experimentos são: 1^a) o que exatamente muda quando a criança passa a ouvir a língua dos que a rodeiam e adquire experiência com a língua nativa?, 2^a) como ocorrem aquelas mudanças relacionadas à habilidade de recuperar a estrutura fonética do sinal acústico de fala? e 3^a) qual é a explicação para as diferenças no padrão de ponderação de pistas auditivas entre crianças e adultos?

Para responder a estas questões, uma série de experimentos foi realizada, utilizando estímulos fricativa-vogal (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 1992), plosiva-vogal (GERRITS, 2001; MAYO E TURK, 2004) e vogal (GERRITS, 2001).

Em um dos estudos pioneiros desenvolvidos por Nittrouer e Studdert-Kennedy (1987), os pesquisadores investigaram as estratégias de ponderação, de crianças (idade três, cinco e sete anos) e adultos, todos falantes nativos do inglês americano, para duas pistas do par mínimo “shoe-sue” e do contraste “sa-cha”. As pistas manipuladas foram as características espectrais do ruído fricativo (altura da frequência) e a transição dos formantes vocálicos. O estímulo era composto por dois continua para /su/-/fu/ e dois continua para /sa/-/fa/, cada um contendo variações sistemáticas no ruído sintético de [s] e [ʃ].

O *continuum* de ruído consistiu em nove porções de ruído aperiódico de 210ms, com centro de frequência de 2.2kHz (apropriado para [ʃ]) a 3.8 kHz (apropriado para [s]), em passos de 200 Hz. Em um *continuum* as porções de ruído foram concatenadas com as porções vocálicas de /su/ e /sa/ e em outro *continuum* com a vogal de /fu/ e /fa/. A diferença entre as vogais estava nos valores dos formantes na região de transição (entre a fricativa e a vogal) e na região estável da vogal.

Os resultados desse estudo, mostrados na figura 1, são semelhantes aos achados publicados em outros artigos de Nittrouer (NITTROUER, 2000, 2002).

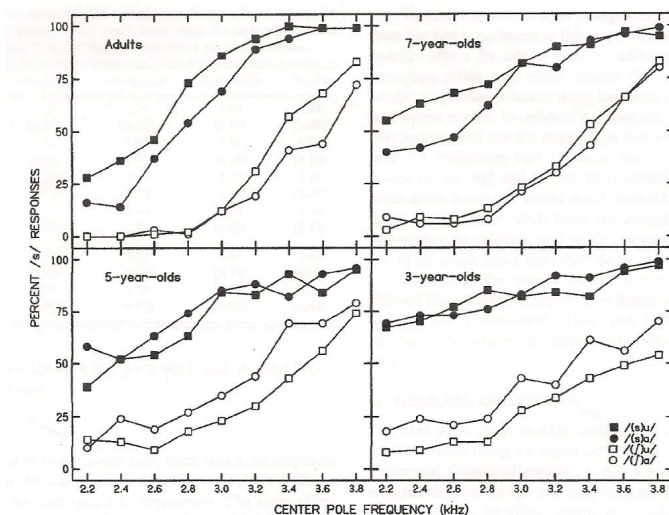


FIGURA 1 - Resultados da tarefa de classificação das crianças de três, cinco, sete anos de idade e adultos para os contrastes /fa/-/sa/ e /fu/-/su/
Fonte: NITTROUER e MILLER (1997a, p. 2262)

As curvas de classificação visualizadas na figura 1 são diferentes em dois aspectos: 1º) as curvas das crianças são menos íngremes do que as dos adultos; e 2º) existe uma separação maior entre as duas curvas das crianças. Isso indica que: a) a pista de transição dos formantes foi mais ponderada nas decisões das crianças do que a dos adultos sobre a identidade das fricativas; e b) as características espectrais do ruído fricativo tiveram menor peso (NITTROUER e STUDDERT-KENNEDY, 1987). Estes resultados de ponderação têm sido replicados diversas vezes com as mesmas combinações /sa/-/ʃa/ e /su/-/ʃu/ (NITTROUER, 1992, 1996, 2000; NITTROUER e MILLER, 1997a). A ponderação destas pistas também foi testada com fricativas em combinação com a vogal [i]: /si/-/ʃi/. Em todos os contextos vocálicos, as crianças ponderaram a pista de transição dos formantes mais fortemente do que os adultos (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 1992, 1996, 2000; NITTROUER e MILLER, 1997b). Tal estratégia seria utilizada como uma tentativa tanto de facilitar a análise linguística como de indicar os gestos articulatórios necessários para produzir sua língua nativa.

As pesquisadoras admitiram o declive da curva de classificação como um índice de peso para o parâmetro representado na abscissa, neste caso, o espectro do ruído fricativo. Quanto mais íngreme a curva, maior peso é atribuído ao espectro do ruído fricativo e mais categórica é a resposta. O peso atribuído à transição dos formantes pode ser observado pela separação entre as curvas. Em todos os estudos de ponderação de pistas auditivas, a inclinação das curvas de classificação é interpretada como a representação de como o índice do peso atribuído para pista é representado pelo continuum de estímulos ao longo da abscissa (NITTROUER e STUDDERT-KENNEDY, 1987; CROWTHER e MANN, 1992; HAZAN e BARRETT, 2000; GERRITS, 2001).

Em estudos de classificação mais tradicionais, em que uma simples pista é variada a cada vez, diferenças relacionadas à idade são também medidas por meio da inclinação das curvas de classificação. Normalmente, as curvas das crianças são menos íngremes do que aquelas de ouvintes adultos, indicando um desempenho de classificação menos consistente, com respostas menos categóricas. É claro que a ponderação de uma pista auditiva está correlacionada ao desempenho da classificação: se crianças ponderam uma pista menos fortemente que ouvintes adultos, é esperado que sua função de classificação seja menos íngreme do que a de adultos.

De acordo com Nittrouer (2002), a transição dos formantes forma a espinha dorsal da sílaba, por assim dizer, ligando o núcleo da sílaba com sua margem. Então, prestar atenção à transição dos formantes poderia proporcionar um caminho para a criança analisar o sinal de entrada quando ela ainda não tem conhecimento pleno da língua. Além disso, por prestar atenção àquelas transições, a criança poderia também obter informação sobre o local das constrictões no trato vocal que deveriam ser feitas para produzir as palavras que seriam ouvidas, apesar da falta de informação sobre a forma de constrictão necessária. Resumindo, a pressão para reconhecer a estrutura interna da sílaba e produzir gestos articulatórios mais precisos levaria as crianças a focalizar nos detalhes espectrais sobre a transição dos formantes.

É claro que poderia existir outra explicação para as tendências relacionadas à idade exibidas na figura 1. Por exemplo, as crianças poderiam ter mais dificuldades que adultos para desempenhar a tarefa utilizada para coletar os dados. Talvez as crianças não possam cumprir tarefas longas o bastante para proporcionar respostas confiáveis ou pode ser que elas tenham dificuldades em categorizar estímulos ambíguos, tais como os que são utilizados em experimentos perceptuais. Para responder a esta questão, Nittrouer desenvolveu estudos com outros sons que evidenciaram que os resultados obtidos com as fricativas estão relacionados com as estratégias de ponderação de pistas, e não com a atenção ou demanda da tarefa de percepção.

Nittrouer, Manning e Meyer (1993) denominaram as diferenças relacionadas à idade na ponderação de pistas fricativas [s]-[ʃ] de “Mudança evolutiva da ponderação de pistas” (*Developmental Weighting Shift* ou DWS). Eles propõem que o desenvolvimento pela criança de um esquema de ponderação de pistas apropriado faz parte do aprendizado da linguagem. Para as fricativas, isso significa que a transição vocálica tem mais efeito sobre o desempenho da classificação de crianças jovens em idade escolar do que no desempenho da classificação de adultos, ao passo que o ruído fricativo tem menos efeito no desempenho da classificação de crianças que adultos. Tal efeito muda com o aumento da experiência da criança com sua língua nativa. Outras hipóteses desenvolvidas para explicar as diferenças de ponderação de pistas por crianças e adultos serão descritas adiante.

Em resumo, parece que para a categorização de [ʃ]-[s] existem evidências convincentes de que crianças jovens ponderam mais a transição dos formantes e menos o espectro do ruído fricativo do que adultos.

A seguir realizou-se uma retrospectiva dos principais estudos que investigaram o desenvolvimento da ponderação de pistas auditivas utilizando pares mínimos ou contrastes compostos pelas fricativas e plosivas, sons investigados nesta tese.

2.4.1 A ponderação perceptiva de pistas auditivas nos estudos desenvolvidos com sons fricativos e plosivos

Basicamente, os resultados dos estudos desenvolvidos por Nittrouer e colaboradores mostraram que as crianças ponderam mais fortemente a pista de transição dos formantes quando necessitam classificar os contrastes compostos pelas fricativas (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 1992 e NITTROUER E MILLER, 1997a). Durante anos de pesquisa, Nittrouer investigou as estratégias de ponderação de pistas de falantes nativos do inglês americano, com foco no contraste fricativo [ʃ]-[s] e buscou sustentar a hipótese DWS, descrita anteriormente. Somente a partir do ano de 2004, a pesquisadora desenvolveu estudos utilizando pares mínimos compostos pelas plosivas. Além de Nittrouer, outros pesquisadores investigaram o tema utilizando contrastes compostos pelas plosivas, pelas vogais e pelas nasais, em outras línguas e os resultados nem sempre estão de acordo com o que a hipótese DWS pressupõe, ou seja, nem sempre as crianças mais jovens ponderaram mais fortemente a pista dinâmica do que a pista estática quando comparadas com os adultos. Estes estudos serão apresentados a seguir.

Ohde e Haley (1997) foram uns dos pesquisadores que estenderam a investigação dos esquemas de ponderação de pistas auditivas de crianças e de adultos, falantes nativos do inglês americano, utilizando contrastes compostos pelas plosivas. Eles realizaram um estudo para determinar se as crianças de três e quatro anos de idade ponderavam as propriedades dinâmicas das consoantes plosivas e das vogais de uma maneira diferente dos adultos. Os estímulos utilizados foram sílabas sintetizadas compostas pelas plosivas [b, d, g] e pelas vogais [i, u, a]. As pistas auditivas manipuladas foram as amplitudes do *burst* (soltura da oclusão) e a transição dos formantes vocálicos. Ohde e Haley (1997) observaram que os adultos apresentaram respostas mais consistentes na tarefa de classificação dos estímulos do

que as crianças, sendo que a percepção das crianças de quatro anos de idade foi significativamente melhor quando comparada com a percepção das crianças de três anos. Além disso, as crianças ponderaram mais a pista de transição dos formantes somente diante dos contrastes compostos pela plosiva velar [g]. Desta forma, nem sempre as crianças utilizaram a pista dinâmica conforme é postulado pela hipótese DWS.

Outro estudo desenvolvido no inglês, porém britânico, foi o de Hanzan e Barrett (2000). As pesquisadoras verificaram o desenvolvimento da categorização fonêmica, de crianças entre 6 e 12 anos de idade e adultos, falantes nativos do inglês britânico. O objetivo do estudo foi investigar a idade na qual as crianças alcançam a competência adulta, em termos tanto da consistência na categorização de contrastes fonêmicos quanto de sua habilidade de categorizar estímulos com informação limitada de pistas auditivas. As pesquisadoras utilizaram estímulos sintéticos. No par mínimo “goat-coat”, as pistas manipuladas foram o tempo de início de vozeamento (VOT) e a frequência inicial do primeiro formante (F1). No par “date-gate”, manipulou-se a amplitude relativa do *burst* e a transição do F2 e do F3. No par composto pelas fricativas, “sue-zue”, as pistas variadas foram a amplitude de vozeamento e a duração da fricção. Por fim, no par “sue-shoe”, as pistas manipuladas foram a altura da frequência do ruído fricativo e a transição do F2. Os estímulos foram apresentados em condições de pistas simples e pista combinada. Por exemplo, no caso das fricativas, as duas pistas auditivas foram manipuladas ao longo do *continuum* na condição de pista combinada. Na condição de pista simples somente uma pista foi manipulada, ora a altura da frequência do ruído fricativo, ora a transição dos formantes.

Os resultados do estudo de Hanzan e Barret (2000) forneceram evidência de que as crianças foram menos consistentes do que adultos ao categorizar um *continuum* contendo informação limitada de pistas auditivas. As crianças com idade entre seis e doze anos parecem mostrar menos flexibilidade do que adultos em suas estratégias perceptuais. Além disso, as pesquisadoras revelaram que a categorização das crianças mais jovens nem sempre foi mais consistente quando somente a pista dinâmica foi disponibilizada, o que contradiz a hipótese DWS, proposta por Nittrouer, Manning e Meyer (1993).

Mais tarde, Gerrits (2001) realizou um estudo sobre o tema de ponderação de pistas auditivas de crianças (idade de quatro, seis e nove anos) e adultos, todos falantes nativos do holandês, utilizando contrastes fricativos, plosivos e vogais. A pesquisadora utilizou

estímulos naturais, compostos por pares de palavras, nos quais foram manipuladas duas pistas auditivas, que especificaram o contraste fonêmico. As palavras originais (“sjok-sok”, “pop-kop” e “zak-zaak”) foram produzidas por um falante nativo feminino do holandês. No contraste fricativo, duas pistas auditivas foram manipuladas: a altura da frequência do ruído fricativo e a transição do segundo e do terceiro formantes (F2 e F3). Os resultados de ponderação de pistas para fricativas geralmente suportaram os achados de Nittrouer e colaboradores (por exemplo, 1992, 1996, 2000). Já aos seis anos de idade o esquema de ponderação das crianças para a pista de transição dos formantes foi similar ao dos adultos. Além do mais, as crianças entre seis e nove anos de idade e os adultos utilizaram somente a informação da pista de ruído das fricativas. Esses achados são diferentes daqueles obtidos por Nittrouer, em cujos estudos as crianças de sete anos de idade ponderaram mais fortemente a pista de transição dos formantes do que os adultos. Gerrits (2001) relata que essa discrepância pode ser explicada pela diferença entre os estímulos, uma vez que a pesquisadora utilizou ruído e porção vocálica naturais e Nittrouer, ruídos sintéticos combinados com porções vocálicas naturais (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 1992, 1996, 2000; NITTROUER e MILLER, 1997a).

Ao investigar os sons plosivos, Gerrits (2001) manipulou as seguintes pistas auditivas: as amplitudes da fase de soltura da oclusão (*burst*) e a transição do F2 e do F3. Neste contraste, a pesquisadora observou que apenas as crianças de quatro anos de idade atribuíram mais peso à pista de transição dos formantes. Os outros grupos etários utilizaram tanto a pista de transição dos formantes quanto a pista de amplitude relativa do *burst* para classificação das plosivas. E, além disso, os adultos mostraram um amplo efeito da pista de transição dos formantes. Esses resultados não corroboraram com os resultados alcançados para as fricativas. Uma das explicações da pesquisadora é que a duração da pista de amplitudes da fase de soltura da oclusão (30ms) foi menor quando comparada com a duração da pista de altura da frequência do ruído fricativo (190ms), o que proporciona aos ouvintes informação fonética mais limitada.

Por fim, para as vogais, Gerrits (2001) manipulou as pistas de duração vocálica e de frequência dos formantes (F1, F2 e F3). A pesquisadora observou que todas as faixas etárias investigadas utilizaram tanto o espectro vocálico quanto a duração vocálica para decidir sobre a identidade da vogal. As crianças de quatro e de seis anos atribuíram menos peso à informação de frequência do formante vocálico do que as crianças mais velhas e os adultos.

Os resultados do estudo de Gerrits (2001) indicaram que houve uma diferença clara entre a ponderação da pista de transição dos formantes pelos ouvintes na percepção das fricativas. Com o intuito de tomar uma decisão sobre a categoria das fricativas, as crianças de seis e de nove anos de idade e os adultos utilizaram somente a informação da pista de ruído. Este não foi o caso para classificação das plosivas, na qual as crianças de seis e de nove anos de idade e os adultos dependeram tanto da pista de ruído quanto da pista de transição dos formantes e, ainda, os adultos mostraram um amplo efeito da pista dinâmica.

Cabe ressaltar que a qualidade dos estímulos utilizados nos estudos desenvolvidos por Nittrouer e colaboradores e Gerrits (2001) é diferente. Enquanto Gerrits (2001) utilizou apenas estímulos naturais, Nittrouer (1992) e Nittrouer e Studdert-Kennedy (1987), em seus experimentos iniciais, utilizaram porções de ruído sintético concatenado com porções vocálicas naturais. Como consequência, pode ser que o maior peso atribuído à transição dos formantes reflita a possibilidade de que as crianças encontrem mais dificuldade para processar a informação do ruído sintético do que dos componentes da fala natural. Esta questão foi abordada em Nittrouer (1996), Nittrouer e Miller (1997a) e Nittrouer e Miller (1997b). Em Nittrouer (1996), a ponderação da pista de transição dos formantes e do ruído foi testada com dois conjuntos de estímulos fricativa-vogal (/sa/-/ja/ e /su/-/ju/). Um conjunto de estímulos consistiu de ruído sintético com porção vocálica natural (estímulo híbrido). Para o estímulo híbrido, foi verificado o mesmo padrão de resultados de estudos prévios: crianças ponderaram a pista de transição dos formantes mais fortemente do que adultos. No entanto, esse não foi o caso do estímulo sintético: não houve diferença significativa na ponderação da pista de transição dos formantes entre crianças (três anos de idade) e adultos. De acordo com Nittrouer (1996), essa ausência de efeito para a transição é causada pelo fato de que somente o F2 (segundo formante) foi variado, e não o F3 (terceiro formante), como ocorreu com o estímulo natural. A ideia de que a transição do F3 também participa da tarefa de classificação foi confirmada por Nittrouer e Miller (1997a). No experimento III deste estudo, novamente, estímulos completamente sintéticos foram utilizados, mas desta vez F2 e F3 foram variados. Os resultados mostraram que, neste caso, crianças atribuíram maior peso do que adultos à informação de transição dos formantes.

Os estudos desenvolvidos por Nittrouer e colaboradores impulsionaram tanto a investigação do desenvolvimento da ponderação de pistas auditivas por crianças e adultos, quanto o levantamento das hipóteses que pudessem explicar as diferenças de ponderação verificadas

entre estas faixas etárias. Em 2001, Sussman realizou um estudo de ponderação de pistas auditivas para a classificação da vogal na estrutura silábica CVC, com crianças de quatro anos de idade e com adultos. As pistas manipuladas foram as transições dos formantes no início e no fim da vogal e as frequências dos formantes vocálicos na região estável. Os resultados mostraram que, embora adultos e crianças fossem capazes de classificar as vogais quando somente a informação dinâmica foi fornecida, as crianças não obtiveram sucesso como os adultos. Além disso, na classificação dos estímulos onde as duas pistas (dinâmica e estática) estavam disponíveis, as crianças fizeram mais uso da informação estática do que os adultos. Sussman (2001) propôs que as diferenças de ponderação de pistas ocorrem devido às diferenças no processamento sensorial geral entre adultos e crianças, ao invés de diferenças nas estratégias entre estes grupos, conforme o modelo DWS postula.

Assim como a comparação entre os resultados dos estudos de Nittrouer e Gerrits (2001) torna-se dificultada em decorrência da qualidade dos estímulos, existem alguns aspectos do estudo de Sussman (2001) que também impedem uma comparação direta com os resultados obtidos por Nittrouer e colaboradores (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER e MILLER, 1997; NITTROUER, MANNING e MEYER, 1993). A pesquisadora produziu os estímulos diferentemente de Nittrouer e colaboradores (1987, 1993) nos estudos que formaram a base do modelo DWS.

Em 2002, Jones pesquisou o desenvolvimento das categorias fonológicas para a percepção do vozeamento final, de crianças de quatro anos de idade e de adultos, falantes nativos do inglês australiano. O par mínimo utilizado neste estudo foi “heart-heard” e as pistas auditivas manipuladas foram o valor da frequência do F1 da vogal que precede as plosivas e a duração vocálica. Jones (2002) verificou que tanto os adultos quanto as crianças fizeram uso das duas pistas auditivas para classificar o vozeamento final das plosivas. Porém, a contribuição relativa que as duas pistas forneceram aos ouvintes para a classificação do par mínimo foi diferente para as crianças quando comparadas com os adultos. As crianças utilizaram mais a pista relativa ao valor da frequência do F1, fizeram um uso menor da pista de duração vocálica e apresentaram menor flexibilidade para lidar com estímulos compostos pelas pistas auditivas mais ambíguas dos continua. Já os adultos apresentaram maior flexibilidade para classificar o par mínimo. Quando a duração vocálica apresentou valores mais ambíguos, os adultos fizeram maior uso da informação disponível pela pista de frequência do F1. Por fim, Jones (2002) concluiu que os resultados de seu estudo

corroboraram com o que Nittrouer e colaboradores propõem, uma vez que as crianças fizeram uso relativamente maior da pista de frequência do F1 quando comparado com a pista de duração vocálica, para classificar o vozeamento final das plosivas [t] e [d] (NITTROUER e STUDDERT-KENNEDY, 1987; NITTROUER, 2002).

Ainda no inglês, porém britânico, Mayo e Turk (2004) verificaram as estratégias de ponderação de pistas de crianças e de adultos para a classificação dos contrastes /səl-/ʃəl/, /de-/be/, /ta-/da/ e /ti-/di/. As pistas auditivas manipuladas, no contraste composto pelas fricativas, foram a altura da frequência do ruído fricativo e a transição dos formantes vocálicos. No contraste /de-/be/, as pistas manipuladas foram as amplitudes do *burst* e a transição dos formantes, enquanto em /ta-/da/ e /ti-/di/, manipulou-se as pistas de duração do tempo de início de vozeamento (VOT) e de transição dos formantes vocálicos. As pesquisadoras observaram que as crianças utilizaram mais a pista de transição dos formantes, quando comparadas com os adultos, para classificar os contrastes compostos pelas fricativas (/səl-/ʃəl/) e pela plosiva alveolar /ta-/da/. Esta tendência não foi observada para a classificação dos demais contrastes (/de-/be/ e /ti-/di/). Mayo e Turk (2004) concluíram que nem sempre as crianças tendem a utilizar a pista dinâmica para classificar todos os contrastes e que os esquemas de ponderação de pistas podem diferir de acordo com o contexto segmental. Desta forma, parte dos resultados desse estudo contradiz a hipótese DWS.

O primeiro estudo de Nittrouer que investigou o papel do componente dinâmico na percepção do vozeamento da plosiva, em posição final de palavra, por crianças e adultos foi realizado no ano de 2004. O objetivo deste estudo foi verificar se os adultos ponderam mais a pista de duração vocálica e se as crianças utilizam mais a pista referente aos valores finais de frequência do F1 em suas decisões perceptuais sobre o vozeamento das plosivas, em posição final de palavra. A pesquisadora desenvolveu duas tarefas de classificação compostas pelos pares mínimos do inglês americano *pot/pod* e *buck/bug* sintetizados e pelos pares naturais *buck/bug*, *cop/cob*, *boot/booed*, *feet/feed* e *pick/pig*. Crianças de sete e de oito anos de idade e adultos participaram dos experimentos. As pistas auditivas manipuladas foram a duração da porção vocálica da sílaba e os valores finais de frequência do primeiro formante vocálico (F1). Os resultados mostraram que os adultos, em ambos os pares sintetizados, basearam as suas decisões utilizando a pista de duração vocálica, enquanto as crianças utilizaram mais intensamente a pista dinâmica de frequência do F1, conforme

esperado por Nittrouer. Diante de todos os estímulos naturais, tanto os adultos quanto as crianças utilizaram mais a pista de frequência do F1 para classificar os pares, embora as crianças tenham ponderado menos a duração vocálica quando comparadas com os adultos. O estudo de Nittrouer (2004) mostrou que houve uma modificação das estratégias de ponderação de pistas de acordo com a qualidade dos estímulos (estímulos sintetizados ou naturais). A pesquisadora relata que esta tendência também foi observada em seu estudo desenvolvido em 2002, com estímulos compostos pelos ruídos fricativos naturais [f]-[θ] associados às vogais sintéticas [a] e [u]. Aqui, Nittrouer (2002) verificou que tanto as crianças (quatro, seis e oito anos) quanto os adultos utilizaram mais intensamente a pista de transição dos formantes para classificar estes contrastes. A pesquisadora concluiu que, quando a informação fonética das propriedades temporais e espectrais estática está limitada pelas condições naturais, os adultos aumentam o peso que eles atribuem aos componentes dinâmicos do sinal de fala. Por exemplo, os ruídos naturais [f]-[θ] diferem muito menos do ponto de vista espectral quando comparados com os ruídos sintéticos [s]-[ʃ], e assim os adultos pesam mais as transições dos formantes vocálicos para classificar as fricativas [f]-[θ] do que as fricativas [s]-[ʃ] (NITTROUER, 2002). Os resultados do estudo de Nittrouer parecem indicar que os adultos possuem uma habilidade maior para lidar com as informações acústicas mais limitadas.

No ano de 2007, Riki, Minka e Liat desenvolveram uma pesquisa com o objetivo de avaliar a ponderação de pistas auditivas (tempo de início de vozeamento (VOT), *burst* inicial e transição dos formantes vocálicos) para a percepção do vozeamento inicial das plosivas com falantes nativos do hebraico. O estudo foi realizado com crianças entre quatro e cinco anos de idade, entre seis e sete anos e entre nove e dez anos, além dos adultos. O contraste /bil-/pil/ foi utilizado para a coleta dos dados. Os pesquisadores verificaram uma hierarquia semelhante de ponderação de pistas entre os grupos, sendo que a pista de VOT foi a mais ponderada em relação às demais pistas. Os adultos deram mais peso à pista de transição dos formantes e ao *burst* quando comparados com as crianças. Além disso, as crianças mais jovens apresentaram respostas menos categóricas e a inclinação das curvas de classificação foi menor em relação aos adultos. Como conclusão Riki, Minka e Liat (2007) descreveram que estes achados realçam a habilidade do sistema auditivo em focar nas pistas importantes desde muito cedo e relatam que é possível que isso seja uma habilidade inata ou adquirida nos primeiros anos de vida.

Por fim, o único estudo sobre o tema realizado com falantes nativos do português brasileiro foi o de Penido (2012). A pesquisadora investigou a ponderação de pistas auditivas com o par mínimo /aʃaR/-/asaR/ e com o contraste /ʃu/-/su/ por crianças (quatro e sete anos de idade) e adultos. As pistas auditivas manipuladas foram a transição do segundo formante (F2) e a altura da frequência do ruído fricativo.

Diante do par mínimo /aʃaR/-/asaR/, tanto as crianças quanto os adultos foram influenciados pela pista de altura da frequência do ruído fricativo. No entanto, as crianças de quatro anos de idade foram necessariamente influenciadas pela pista de transição do F2 quando comparadas com os demais grupos de idade. Além disso, o único grupo que demonstrou uma interação entre as pistas auditivas de altura da frequência do ruído fricativo e de transição do F2 foi o de quatro anos. Isso quer dizer que para este grupo de idade as pistas não foram independentes e a alteração no valor de uma refletiu em alteração no valor da outra. Este resultado está de acordo com os resultados dos estudos de Nittrouer e Studdert-Kennedy (1987), Nittrouer e Miller (1997a) e Mayo e Turk (2004).

Os resultados de ponderação de pistas auditivas do estudo de Penido (2012) alcançados para o contraste /ʃu/-/su/ demonstraram discordância dos achados encontrados nos estudos de Nittrouer e Studdert-Kennedy (1987) e Nittrouer e Miller (1997), uma vez que os ouvintes modificaram suas estratégias de ponderação quando comparadas com o par mínimo /aʃaR/-/asaR/. A faixa etária mais influenciada pela pista de altura da frequência do ruído fricativo foi a das crianças de quatro anos. Esta faixa etária também foi fortemente influenciada pela pista de transição do F2 para a classificação deste contraste. Já as crianças de sete anos ponderaram mais fortemente a pista de transição do F2, enquanto os adultos demonstraram maior influência pela pista de transição do F2, mas também utilizaram fortemente a pista de altura da frequência do ruído fricativo. Esta divergência observada nos resultados para o contraste /ʃu/-/su/ poderia ser explicada pelo seguinte motivo: no português brasileiro os ouvintes não utilizam a pista de altura da frequência do ruído fricativo tanto quanto os ouvintes ingleses para a classificação deste contraste. Ou seja, a pista de altura da frequência do ruído fricativo para a vogal [u] não oferece informações fonéticas suficientes para a classificação de /ʃu/-/su/ no português.

Esta variação nas estratégias de ponderação de pistas entre crianças e adultos foi identificada em outros estudos, conforme descrito nesta seção (GERRITS, 2001; MAYO e TURK,

2004). Assim, para que a mudança no desenvolvimento da ponderação de pistas auditivas entre crianças e adultos seja considerada um fenômeno universal entre as línguas, conforme propõem Nittrouer e colaboradores no modelo DWS, ela deveria ocorrer em qualquer contraste investigado, de qualquer língua. Direcionando ao PB, o estudo de Penido (2012) sugeriu que nesta língua existem indícios de que a pista de frequência do ruído fricativo fornece informações fonéticas mais favoráveis para a classificação do contraste /f-vogal/-/s-vogal/ desde muito cedo no desenvolvimento da criança. Pode-se atribuir esta característica à língua, o que aponta para o fato de que as pistas são dependentes da língua.

Na próxima seção será realizada uma reflexão sobre as hipóteses que buscam explicar as diferenças na ponderação de pistas auditivas entre crianças e adultos.

2.4.2 As hipóteses que buscam explicar as diferenças na ponderação de pistas auditivas entre crianças e adultos

Existem duas principais hipóteses que buscam elucidar as diferenças na ponderação de pistas auditivas entre crianças e adultos: o modelo DWS, proposto por Nittrouer, Manning e Meyer (1993) e as explicações auditivas oriundas do estudo de Sussman (2001).

O modelo DWS apóia-se em teorias que propõem que o nível de detalhes exigido para representar os itens lexicais se modifica com o crescimento do léxico. Neste tipo de estrutura, um léxico restrito exige somente um nível de detalhe global para acomodar adequadamente e diferenciar entre todos os itens armazenados. Assim, o léxico restrito da criança pode ser representado em termos de sílabas ou palavras monossílabas. Com o crescimento do léxico e com a experiência linguística, um nível mais sofisticado de detalhes é exigido para diferenciar entre todos os itens. Um léxico mais amplo, tal como o de crianças mais velhas e o de adultos, deve ser representado em termos de unidades muito menores, isto é, com base em segmentos ou, possivelmente, traços. Estudos da produção da fala de crianças proporcionam suporte para este ponto de vista. Padrões nos enunciados precoces de crianças parecem indicar que para elas “a palavra é uma entidade, armazenada e acessada como um bloco” (MENN, 1971, p. 247; NITTROUER, STUDDERT-KENNEDY e MCGOWAN, 1989).

Nittrouer e colaboradores (2000) relatam que correlatos físicos de características mais globais seriam as transições dos formantes. As pesquisadoras ressaltam que pistas como estas são “perceptivamente salientes e delimitam porções do sinal correspondentes às sílabas” (NITTROUER, MILLER, CROWTHER *et al.*, 2000, p. 268). Assim, as crianças jovens fazem relativamente mais uso da informação dinâmica que os adultos porque percebem e processam a fala de forma mais global, enquanto eles fazem relativamente menos uso da informação dinâmica, visto que percebem e processam a fala mais analiticamente. Além disso, as crianças ponderam mais fortemente a transição dos formantes, uma vez que esta pista auditiva está relacionada a modulações relativamente lentas na estrutura espectral. Desse modo, a transição dos formantes atrairia a atenção da criança precocemente no desenvolvimento, uma vez que elas são dinâmicas.

Outra hipótese que busca esclarecer o motivo pelo qual as crianças focam fortemente na transição de formantes emerge da explicação que propõe que as diferenças de ponderação de pistas ocorrem devido às diferenças no processamento sensorial geral entre adultos e crianças, e não às diferenças nas estratégias entre esses grupos (SUSSMAN, 2001). Sobre essa explicação, a maturação de alguns aspectos do sistema auditivo central origina as diferenças de percepção da fala. Em consequência disso, as crianças tendem a utilizar as pistas mais informativas no sinal acústico de fala, ou seja, as pistas mais altas, mais longas e que possuem maior quantidade de informação espectral.

Oponente à hipótese de Sussman (2001), os pesquisadores Eisenberg *et al.* (2000) sugerem que, como a maturação das habilidades perceptuais auditivas ocorre durante os primeiros dez, doze meses de idade, os padrões centrais de reconhecimento da fala podem ser muito menos robustos para as crianças do que para os adultos. Assim, elas podem ser incapazes de lidar perceptivamente com a informação acústica insuficiente ou incompleta e podem, então, necessitar de pistas auditivas mais distintas do que os adultos para identificar ou discriminar os estímulos.

Conforme mencionado na seção anterior, Sussman (2001) desenvolveu um estudo de ponderação de pistas para a classificação da vogal, na estrutura silábica CVC, manipulando as pistas de transição dos formantes no início e no final da vogal (pista mais curta e menos informativa) e as frequências dos formantes vocálicos na região estável (pista mais longa e mais informativa), para dar suporte a hipótese sensorial. No entanto, algumas questões entre

o estudo de Sussman (2001) e os estudos desenvolvidos por Nittrouer e colaboradores devem ser consideradas, tais como as diferenças na qualidade dos estímulos (fricativas x vogais) e as diferenças relacionadas à metodologia dos estudos (NITTROUER, MANNING e MEYER, 1993; NITTROUER e MILLER, 1997a; NITTROUER e STUDDERT-KENNEDY, 1987). Tais questões impedem uma comparação direta entre os resultados alcançados. Além disso, Sussman (2001) não manipulou a saliência da pista auditiva de forma clara e não foi desenvolvido um estudo que buscasse comparar o padrão de ponderação de pistas de crianças e de adultos diante de pistas longas, intensas e mais informativas versus pistas curtas e menos informativas.

Nos anos de 1996 e 1998, Nittrouer e colaboradores testaram explicitamente o papel do sistema auditivo sobre os esquemas de ponderação das crianças em dois estudos. No primeiro estudo, Nittrouer (1996) demonstrou que as crianças necessitam de diferenças acústicas mais amplas, tanto no espectro de ruído fricativo quanto na transição do F2, para discriminar o contraste /su/-/ʃu/ do que os adultos. No segundo estudo, Nittrouer e Crowther (1998) investigaram o papel da sensibilidade auditiva em uma série de experimentos de discriminação com estímulos não pertencentes à fala. Assim, limiares de diferenciação⁶ (do inglês, *difference limens* - DLs) foram obtidos para crianças (cinco e sete anos de idade) e adultos, para três tipos de propriedades acústicas: espectrais-dinâmicas, espectrais-estáticas e temporais. As crianças obtiveram DLs mais amplos que os adultos para as propriedades tanto estáticas quanto dinâmicas. Não existiram diferenças relacionadas à idade entre os DLs para as propriedades temporais. Os resultados deste estudo indicaram que as diferenças na sensibilidade auditiva não podem explicar as diferenças anteriormente relacionadas à idade na percepção da fala. Em 2007, Nittrouer e Lowenstein realizaram outro estudo que também investigou se as diferenças na ponderação de pistas entre crianças e adultos podem ser explicadas pelas diferenças na sensibilidade auditiva. As pesquisadoras verificaram o mesmo padrão de resultados do estudo desenvolvido em 1996 e em 1998.

Em estudo mais recente, Mayo e Turk (2005a) objetivaram testar a viabilidade da explicação sensorial de Sussman (2001) como uma alternativa para a teoria DWS de Nittrouer (1993). As pesquisadoras testaram a hipótese de Sussman (2001) a partir da investigação da

⁶ Em inglês, *difference limens*. Refere-se à menor variação em um estímulo que permita a uma pessoa detectar a diferença entre dois estímulos.

ponderação das pistas dinâmicas pelas crianças e pelos adultos, na condição tanto espectralmente informativa quanto menos informativa. Se os resultados do estudo indicassem que as crianças atribuíram mais peso que os adultos para transições mais informativas e menos peso que os adultos para as transições menos informativas, a afirmação de Sussman (2001) seria suportada.

O estudo de Mayo e Turk (2005a) acompanhou a mesma metodologia de Nittrouer e Studdert-Kennedy (1987). A ponderação de pistas das crianças (idades de três a quatro anos, cinco e sete anos) e dos adultos foi investigada utilizando contrastes compostos por transições dos formantes vocálicos semelhantes (/ni-/mi/, /de-/be/ e /ti-/di/) e diferentes (/no-/mo/, /do-/bo/ e /ta-/da/), do ponto de vista espectral. Por exemplo, nos contrastes compostos pelas nasais /no-/mo/ e /ni-/mi/, a transição dos formantes vocálicos em /no/ foi bastante distinta de /mo/ (principalmente o F2), fazendo com que esta pista seja mais informativa para determinar a consoante deste contraste. Já a transição dos formantes em /ni-/mi/ foi semelhante, o que deixou esta pista muito menos informativa para a classificação das consoantes. Para os contrastes descritos acima, as pistas estáticas foram: murmúrio nasal para os contrastes /no-/mo/ e /ni-/mi/, frequência inicial da plosiva para /do-/bo/ e /de-/be/ e tempo de início de vozeamento (VOT) para /ta-/da/ e /ti-/di/. Já as pistas dinâmicas foram as frequências dos formantes (F1, F2 e F3) no início da vogal ou no final do vozeamento, no caso do contraste /t-/d/.

Os resultados indicaram que, embora crianças e adultos não tenham apresentado diferenças em suas ponderações quanto à pista espectralmente mais informativa para /no-/mo/, as crianças jovens deram muito menos peso do que as crianças mais velhas e os adultos para a pista menos informativa em /ni-/mi/. Do mesmo modo, crianças e adultos diferiram ligeiramente em suas ponderações da pista mais informativa para /do-/bo/, mas as crianças deram muito menos peso que os adultos para a pista menos informativa em /de-/be/. Segundo Mayo e Turk (2005a), tal fato proporciona algum suporte para as afirmações de Sussman (2001), em que as diferenças na ponderação de pistas poderiam ser causadas pela dificuldade da criança em utilizar as pistas menos informativas e mais curtas do sinal acústico.

Os resultados para o contraste [t]-[d], apesar de proporcionarem mais evidência das diferenças na ponderação de pistas entre crianças e adultos, não suportaram a hipótese

sensorial de Sussman (2001). Mayo e Turk (2005a) observaram que nenhuma diferença foi encontrada entre crianças e adultos no peso dado à pista menos informativa em /ti-/di/. No entanto, a hipótese baseada nas diferenças auditivas gerais entre crianças e adultos prediz que as crianças deveriam dar menos peso para esta pista que os adultos. Finalmente, no contraste /ta-/da/, as crianças também pesaram menos fortemente a pista mais informativa que os adultos. De acordo com as pesquisadoras, estes resultados para os dois contrastes [t]-[d] também não são previstos pela hipótese DWS, com base nas estratégias de ponderação de Nittrouer, Manning e Meyer (1993): as crianças não pesaram a informação dinâmica mais fortemente que os adultos para /ti-/di/ e /ta-/da/.

Mayo e Turk (2005a, 2005b) concluíram que qualquer explicação para a ponderação de pistas que seja baseada somente nas diferenças sensoriais ou nas diferenças de estratégias entre crianças e adultos não é capaz de dar conta de todas as diferenças de ponderação entre estes dois grupos. Além disso, as pesquisadoras sugerem a realização de mais pesquisas para determinar se uma associação das explicações estratégicas e sensoriais pode justificar todas as diferenças na ponderação de pistas ou se estas diferenças seriam decorrentes de outros fatores perceptivos ou cognitivos gerais.

Penido (2012) também verificou que os resultados de seu estudo não corroboraram em sua totalidade com o que é proposto pelo modelo DWS (Nittrouer, Manning e Meyer, 1993), uma vez que os sujeitos modificaram suas estratégias de ponderação de pistas auditivas do par mínimo /aʃaR-/asaR/ para o contraste /ʃu-/su/. Assim, considerando a importância e interesse em pesquisas de percepção da fala, o presente estudo se propõe a investigar de forma mais aprofundada o desenvolvimento da ponderação de pistas auditivas nos falantes nativos do português brasileiro (PB) e, conseqüentemente, responder à seguinte questão: as crianças atribuem os mesmos pesos às pistas dinâmicas que os adultos para a classificação das fricativas coronais não vozeadas e das plosivas vozeadas [b] e [d] do português brasileiro? Aqui, buscou-se desenvolver um estudo mais controlado, do ponto de vista metodológico, utilizando pares mínimos e pseudopalavras compostos pelas fricativas [s] e [ʃ] e pelas plosivas [b] e [d] em posição inicial de palavra, mantendo o mesmo contexto vocálico (vogal [a]). Espera-se que os resultados indiquem que nem sempre as crianças mais jovens ponderam mais fortemente a pista dinâmica do que a pista estática. Presume-se que as crianças de quatro anos de idade, que possuem menos experiência na língua, sejam menos estáveis e consistentes do que os adultos para classificar os estímulos e utilizem qualquer

pista auditiva que esteja mais disponível no sinal acústico de fala, seja ela dinâmica ou estática. Se tal fato for verificado, a hipótese proposta por Nittrouer e colaboradores não será sustentada.

A seguir, será apresentado o primeiro capítulo referente à análise das propriedades acústicas dos estímulos utilizados no experimento de ponderação de pistas auditivas.

3 CAPÍTULO 1 - ESTUDO PILOTO DAS PROPRIEDADES ACÚSTICAS DAS FRICATIVAS CORONAIAS NÃO VOZEADAS [ʃ] E [s] E DAS PLOSIVAS VOZEADAS [b] E [d] DO PORTUGUÊS BRASILEIRO EM POSIÇÃO INICIAL DE PALAVRA

Este capítulo foi desenvolvido com o intuito de descrever e analisar previamente as propriedades acústicas das fricativas coronais não vozeadas e das plosivas bilabial e alveolar vozeadas, que foram utilizadas no experimento de ponderação de pistas adiante. Tal estratégia foi extremamente importante para criação dos estímulos por meio da síntese de fala, uma vez que estes foram criados com base nas características acústicas da fala natural de um dos informantes. Desta forma, foi realizado um breve levantamento teórico das propriedades acústicas das fricativas ([s] e [ʃ]) e das plosivas ([b] e [d]), bem como um estudo piloto, com adultos do sexo masculino, falantes nativos do PB. Pelo fato de este estudo se limitar às fricativas [s] e [ʃ] e às plosivas [b] e [d], serão descritos somente os estudos que investigaram as propriedades acústicas destes sons.

Além disso, neste capítulo, encontram-se descritos a metodologia e os resultados obtidos com o estudo piloto. A descrição detalhada das propriedades acústicas das fricativas e das plosivas foi realizada somente nas primeiras sílabas das palavras /ʃapa/, /sapa/, /ʃaba/, /saba/, /bata/ e /data/, uma vez que no experimento de ponderação, as pistas auditivas dinâmicas e estáticas foram manipuladas somente na posição inicial de palavra. Ainda no caso das fricativas, analisou-se as propriedades acústicas somente da primeira sílaba das palavras /ʃapa/ e /sapa/, já que as palavras /ʃaba/ e /saba/ são compostas pelos mesmos segmentos fricativos e pela vogal [a].

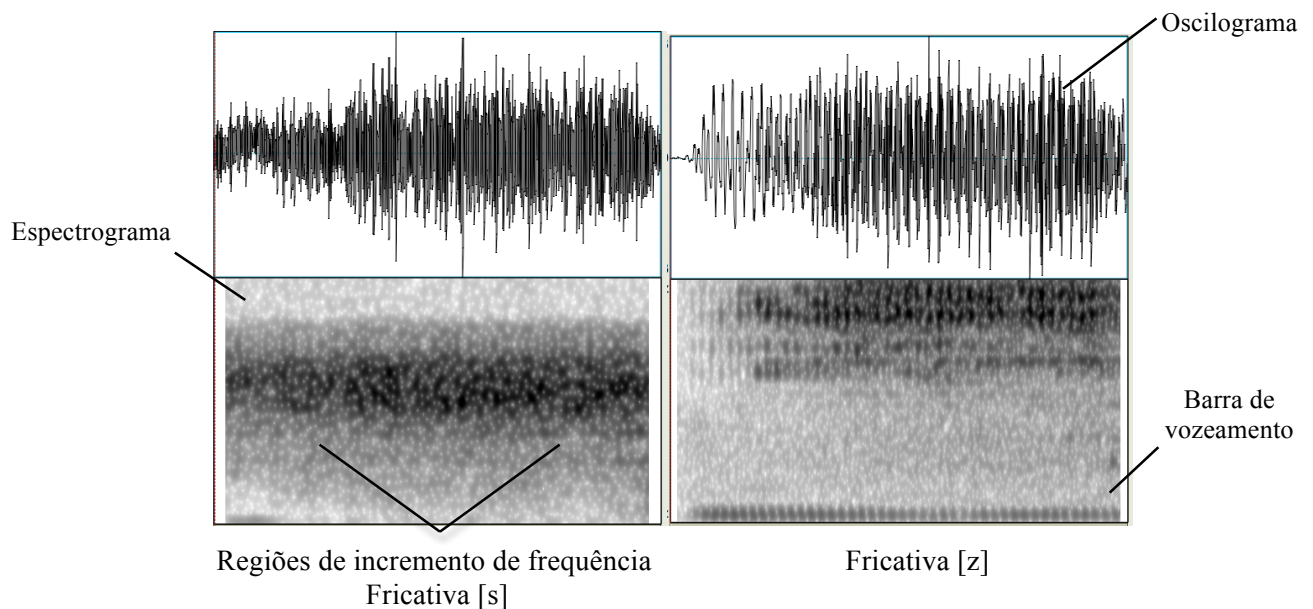
3.1 Descrição das propriedades acústicas das fricativas, com foco nas fricativas [s] e [ʃ]

Os sons fricativos representam a segunda categoria de sons em nível de ocorrência na maior parte das línguas (CARTERETTE e JONES, 1974) e são produzidos por constrições muito estreitas no trato vocal, que causam turbulência ou fricção da corrente aérea (PETERSON e SHOUP, 1966).

Foneticamente, o Português brasileiro apresenta cinco consoantes fricativas não vozeadas /f, s, ʃ, X, h/ e cinco vozeadas /v, z, ʒ, ʝ, ã/, que são classificadas quanto ao ponto de articulação em labiodental, alveolar, alveolopalatal, velar e glotal (SILVA, 2001).

Diversos autores deram importante impulso à compreensão das propriedades acústicas dos sons fricativos (FANT, 1960; HEINZ e STEVENS, 1961; HIXON, 1966; SHINDLER, 1974; BORZONE DE MANRIQUE e MASSONE, 1981; LACERDA, 1982; SANTOS, 1987). Potter, Kopp e Green Kopp (1947) são os grandes pioneiros na utilização da espectrografia acústica no estudo dos sons da fala.

Através da análise espectrográfica é possível observar que as fricativas não vozeadas são formadas por uma fonte de ruído resultante da constrição do trato vocal, enquanto as vozeadas são formadas por duas fontes: a fonte de ruído e a fonte glótica, responsável pelo vozeamento (POTTER, KOPP e GREEN KOPP, 1947).



FIGURAS 2 e 3 – Oscilograma e espectrograma da fricativa coronal não vozeada [s] (à esquerda) e da fricativa coronal vozeada [z] (à direita)

As figuras 2 e 3 mostram tanto o oscilograma quanto o espectrograma das fricativas coronais [s] (à direita) e [z] (à esquerda). O registro do espectro da fonte glótica torna-se visível no espectrograma pela barra de vozeamento. Já a fonte de fricção pode ser observada pelas regiões de incremento de frequências no espectrograma da fricativa coronal não vozeada [s].

De acordo com a necessidade fonética, o uso da fonte de ruído e o da glótica pode ser sucessivo ou simultâneo. Quando o falante necessita produzir uma consoante sonora torna-se necessário o uso simultâneo dessas duas fontes, gerando um espectro misto, resultado da inserção da vibração glótica na turbulência aérea, revelada pela fricativa [z] na figura 3 (TOSI, 1979).

Santos (1987), em seu estudo sobre análise espectrográfica dos sons fricativos surdos e sonoros do português brasileiro, verificou que as características espectrais do som fricativo dependem do lugar da constrição no trato vocal, do formato do orifício na constrição e da queda de pressão nesta região.

Laderfoged e Maddieson (1993) também evidenciaram que as diferenças de natureza articulatória entre [s] e [ʃ], objetos deste estudo, irão influenciar no espectro da fonte de ruído, uma vez que este depende da forma do trato vocal pela qual a corrente de ar passa, além da velocidade do fluxo de ar através da constrição. Durante a produção do [s] temos a formação de um canal mais estreito. O canal mais largo para o [ʃ] faz com que o ar tenha menos velocidade. Devido à diminuição da velocidade da corrente de ar, da extensão maior do trato vocal e do arredondamento dos lábios, o [ʃ] tem maior concentração de energia espectral em regiões mais baixas quando comparado com [s].

Russo e Behlau (1993), em seus estudos das consoantes do português brasileiro, verificaram que estas diferenças articulatórias na produção das fricativas [s] e [ʃ] geram diferenças espectrográficas nestes sons. Desta forma, de acordo com as pesquisadoras, a fricativa alveolar [s] é um som forte, agudo, com uma faixa de frequência acima de 4500Hz, chegando a 8000Hz. Enquanto a alveolopalatal [ʃ] também apresenta uma faixa de frequência ampla, semelhante a [s], porém mais grave, entre 2500 e 6000Hz. Outros pesquisadores confirmaram essa diferença espectrográfica dos sons fricativos alveolares e alveolopalatais.

Stevens e Klatt (1968) constataram que o espectro acústico das fricativas não vozeadas e vozeadas do inglês apresenta energia mais concentrada nas altas frequências. O espectro do som [ʃ] aparece com a principal concentração de energia na faixa de frequência de 2000 a 3000Hz. Para o som [s] o espectro mostra um aumento de energia que começa ao redor de 3560 a 4400Hz.

Shindler (1974) também descreveu as características dos sons fricativos do italiano e observou que o espectro do som [s] se localiza em torno dos 4000Hz e o [ʃ] em torno dos 2000Hz.

Borzzone de Manrique e Massone (1981) em seu estudo sobre análise acústica e perceptual dos sons fricativos do espanhol argentino, verificaram que o som [s] apresenta picos ao redor de 5000 e 8000Hz e [ʃ] por volta de 2500 e 5000Hz.

Além dos autores supracitados, Lacerda (1982) realizou um estudo sobre a análise perceptual dos sons fricativos surdos do português europeu e concluiu, através de testes perceptuais, que o som [s] é melhor percebido quando o estímulo tem altos níveis de intensidade e picos espectrais na região de 5000Hz, e o som [ʃ] é normalmente associado com altos níveis de intensidade juntamente com picos espectrais na região dos 3000Hz.

Santos (1987) e Haupt (2007), após realizarem a análise acústica dos sons fricativos do português brasileiro, verificaram resultados semelhantes aos demais autores em relação aos valores das frequências dos sons [s] e [ʃ].

Outro parâmetro investigado nos sons fricativos é a duração. Diversos estudos evidenciaram que os sons fricativos alveolares [s, z] são mais longos que os alveolopalatais [ʃ, ʒ], e os sons surdos são mais longos que os sonoros (STEVENS e KLATT, 1968; BORZONE DE MANRIQUE e MASSONE, 1981; SANTOS, 1987; HAUPT, 2007).

Muitos estudos acústicos sobre os sons fricativos já foram realizados, porém não se verificou estudos que analisaram a propriedade acústica de amplitude do ruído fricativo no PB, fundamental para a criação das fricativas sintetizadas. Desta forma, o estudo piloto busca analisar principalmente quatro parâmetros dos sons fricativos: os valores dos formantes⁷ e as amplitudes do ruído de fricção, a duração desses sons e a região de maior concentração de energia no espectro. Tais parâmetros foram relevantes para a síntese destes

⁷ De acordo com Barbosa e Madureira (2015) somente uma classe de sons não produz formantes: as dos sons bilabiais, uma vez que estes sons são produzidos nos lábios e são emitidos sem passar por um tubo. Pode soar estranho falar de formantes para as fricativas, porém os formantes não se limitam às vogais. Barbosa e Madureira (2015, p. 95) relatam que “os formantes são, assim, propriedades de um tubo acústico: havendo tubo ressoante há possibilidade de ondas estacionárias ou ressoantes”. Assim, o ruído fricativo produz formantes.

sons no *software* Sensyn (Sensimetrics Org.) e são considerados pistas auditivas importantes para que os ouvintes façam a distinção entre as fricativas [s] e [ʃ]. Já a análise dos momentos espectrais dos sons fricativos (variância, assimetria e curtose) não foi realizada, uma vez que estes parâmetros não serão utilizados para a síntese.

3.2 Descrição das propriedades acústicas das plosivas, com foco nas plosivas [b] e [d]

As consoantes plosivas são segmentos produzidos com o bloqueio total da corrente de ar em algum ponto do trato vocal, com posterior soltura deste bloqueio.

O português brasileiro possui três segmentos plosivos não vozeados /p, t, k/ e três segmentos vozeados /b, d, g/, que são classificados quanto ao ponto de articulação como bilabiais ([p] e [b]), alveolares ([t] e [d]) e velares ([k] e [g]). A plosiva bilabial vozeada [b] é produzida por meio da constrição dos lábios e da vibração simultânea das pregas vocais, enquanto a plosiva [d] é produzida com a constrição da ponta da língua contra a face interna dos dentes incisivos superiores. Neste caso, também ocorre a vibração das pregas vocais (LADEFOGED e MADDIESON, 1996).

Do ponto de vista articulatorio, a produção das plosivas é identificada por duas fases distintas: a fase de oclusão e a fase de soltura (ISTRE, 1983). A fase de oclusão corresponde ao intervalo em que os articuladores interrompem completamente a passagem de ar. Esta fase é caracterizada somente pelo período de silêncio, no caso das plosivas não vozeadas, ou pelo período de silêncio associado à sonorização, no caso das plosivas vozeadas. A sonorização pode ser visualizada no espectrograma da figura 4 pela barra de vozeamento, que corresponde à vibração das pregas vocais. A barra de vozeamento pode ser observada pela concentração de energia em frequências baixas ao longo da base do espectrograma.

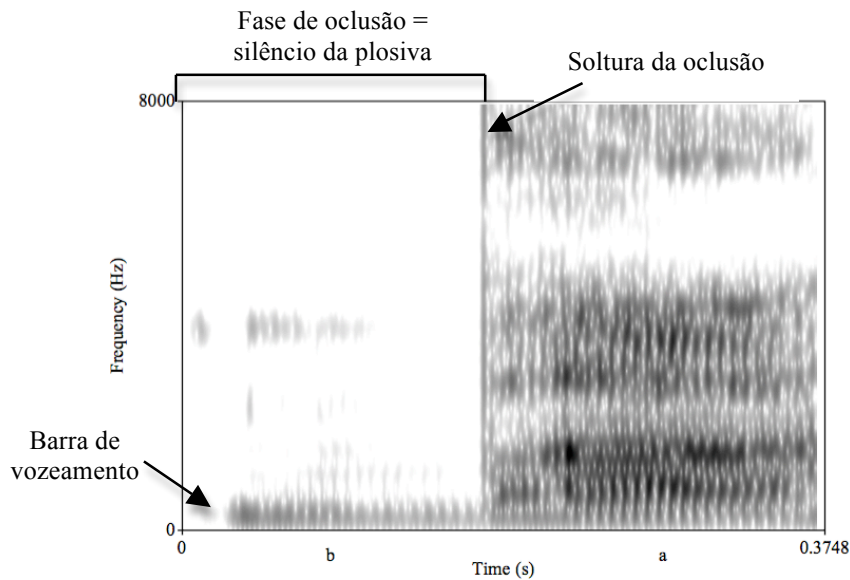


FIGURA 4 - Barra de vozeamento na plosiva [b] da sílaba [ba], visualizada pela região mais escurecida na base do espectrograma

A fase de soltura, que corresponde à liberação da corrente de ar previamente bloqueada em algum ponto do trato vocal, é caracterizada pela espícula de plosão ou *burst*, seguida da região de transição dos formantes para a vogal seguinte. Em geral, observa-se uma espícula mais intensa para as plosivas não vozeadas /p, t, k/ do que para as plosivas vozeadas /b, d, g/.

Stevens e Blumstein (1978) observaram que o *burst* pode ser considerado uma pista importante para a distinção entre as consoantes plosivas. Os pesquisadores verificaram que o *burst* das plosivas bilabiais possui energia concentrada em regiões de frequência mais baixas, quando comparado com o *burst* das plosivas alveolares, que apresentam energia concentrada em regiões de frequência mais altas. Segundo Russo e Behlau (1993), os segmentos plosivos [p] e [b] são os mais graves e possuem energia em torno de 500 a 1500Hz. Os segmentos [t] e [d] são considerados os mais agudos, com energia concentrada por volta de 4000Hz, com um pico secundário fraco, em torno de 500Hz.

Acusticamente, além dos parâmetros de barra de vozeamento e do *burst*, outros aspectos são considerados importantes para a distinção das consoantes plosivas vozeadas e não vozeadas, tais como a duração da fase de oclusão e o tempo de início de vozeamento (VOT).

As plosivas não vozeadas são mais longas do que as vozeadas. Lousada, Jesus e Hall (2010) verificaram tal fato no português europeu. A duração média da fase de oclusão das plosivas

em posição inicial de palavra foi: 155ms para [p], 146ms para [t], 128ms para [k], 108ms para [b], 90ms para [d] e 75ms para [g]. No PB, Barbosa (1999) também verificou valores de duração da fase de oclusão maiores para as plosivas não vozeadas quando comparada com as plosivas vozeadas. O pesquisador observou os seguintes valores: 120ms para [p], 113ms para [t], 121ms para [k], 86ms para [b], 71ms para [d] e 67ms para [g].

Por fim, será abordado o parâmetro de tempo de início de vozeamento (VOT), que pode ser definido como o intervalo de tempo entre a soltura do segmento plosivo e o início do vozeamento. Os segmentos vozeados apresentam o início do vozeamento antes da soltura da oclusão. Neste caso, teremos o valor de VOT negativo. Já os segmentos não vozeados possuem o valor de VOT positivo ou nulo, uma vez que o vozeamento inicia-se logo após a soltura da oclusão (LISKER e ABRAMSON, 1964). O VOT é um parâmetro relevante na discriminação do traço de sonoridade dos segmentos plosivos em diversas línguas. Os valores do VOT para o PB são bem menores quando comparado com os valores do inglês (RUSSO e BEHLAU, 1993).

Como este estudo utilizou dois segmentos plosivos vozeados, [b] e [d], para investigar o desenvolvimento da ponderação de pistas auditivas, o VOT não será abordado detalhadamente.

Conforme relatado, o estudo piloto teve como objetivo analisar as propriedades acústicas das fricativas [s] e [ʃ], das plosivas [b] e [d] e da vogal [a] para a posterior síntese destes segmentos. Desta forma, o parâmetro acústico mais importante para a síntese das fricativas foram as amplitudes do ruído de fricção e para as plosivas foram os valores das amplitudes do *burst*. Além destas pistas, ainda temos a pista de transição dos formantes vocálicos, que é considerada relevante quando os ouvintes necessitam discriminar os segmentos. Esta pista será discutida adiante.

3.3 Descrição das propriedades acústicas da vogal [a], com foco na transição dos formantes (F1, F2 e F3)

Os formantes vocálicos podem ser definidos como regiões de maior concentração de energia em determinadas faixas de frequências. Fant (1960) definiu os formantes como frequências de ressonâncias do trato vocal.

No espectrograma da figura 5 observa-se regiões mais escurecidas na vogal [a]. Estas são as regiões que definem os formantes.

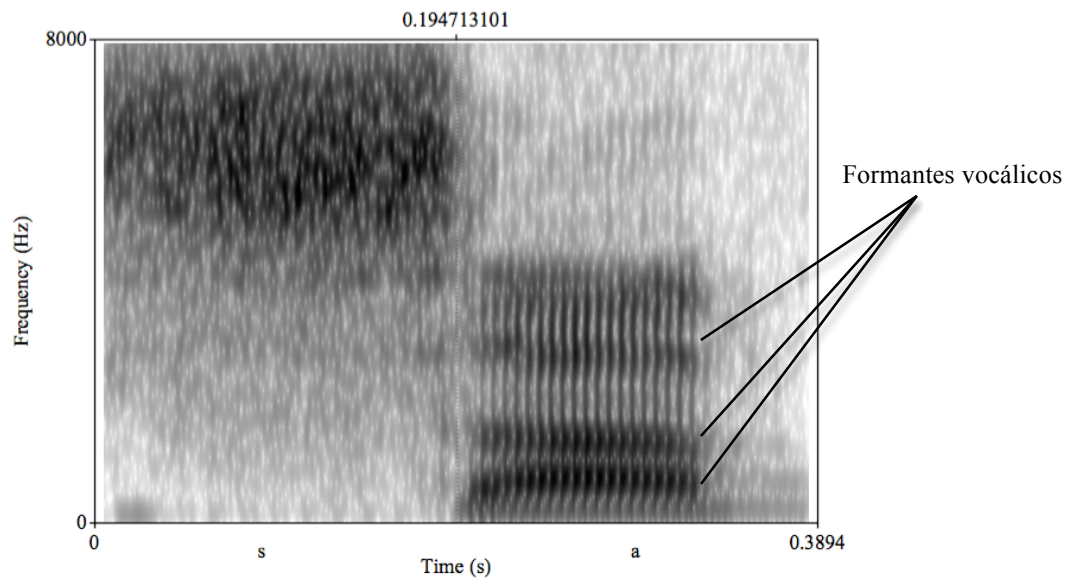


FIGURA 5 - Três primeiros formantes (F1, F2 e F3) da vogal [a] seguida da fricativa [s], na palavra [sapa]

Os três primeiros formantes (F1, F2 e F3) são os mais importantes e considerados suficientes para diferenciar as vogais.

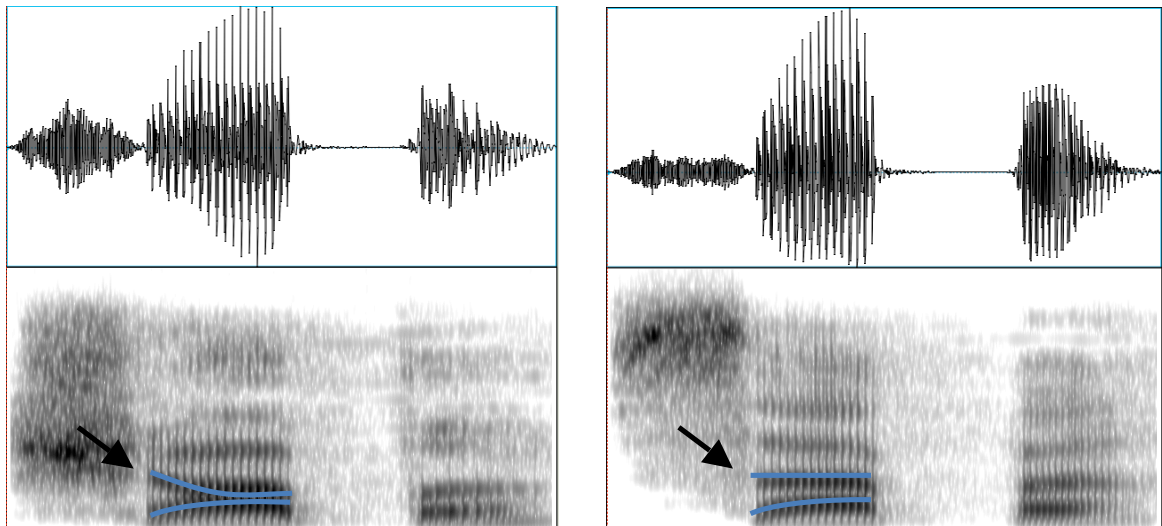
O PB possui sete vogais orais, que são classificadas de acordo com um plano horizontal e vertical, ou seja, conforme a abertura da boca, altura e anteriorização da língua. De acordo com o plano vertical, as vogais podem ser classificadas como baixa ([a]), médias baixas ([ɛ] e [ɔ]), médias ([e] e [o]) e altas ([i] e [u]) (CÂMARA JÚNIOR, 1970).

Delattre (1951) observou que o grau de abertura de uma vogal e altura da língua interfere nos valores do primeiro formante (F1) e o grau de anteriorização influencia nos valores do segundo formante (F2). Assim, a vogal [a], que é produzida com maior abertura da boca, possui valores maiores do F1 quando comparada com as demais vogais.

Behlau *et al.* (1988) analisaram os valores dos três primeiros formantes das sete vogais orais do PB, de homens, de mulheres e de crianças, da cidade de São Paulo. Os pesquisadores verificaram que os valores dos formantes são mais baixos para os homens, uma vez que

estes apresentam os maiores tratos vocais. Porém, os pesquisadores obtiveram os valores dos formantes na região estável das vogais.

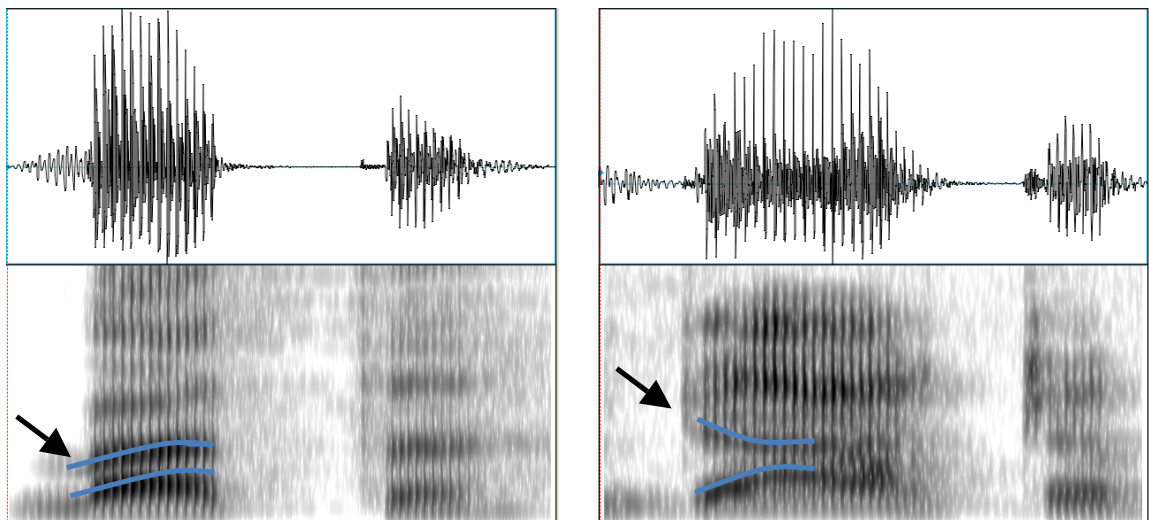
O estudo piloto que será descrito a seguir analisou os valores dos formantes (F1, F2 e F3) da região estável da vogal [a], mas principalmente da região de transição dos formantes vocálicos. Ou seja, da região entre a consoante e a vogal. A mudança nos valores dos formantes desta região é considerada uma pista auditiva importante para que os ouvintes façam a distinção entre os sons. Por exemplo, espera-se observar diferenças nos valores do segundo formante vocálico, da vogal [a] que segue as fricativas [s] e [ʃ], com valor maior do F2 da vogal que segue a fricativa [ʃ] quando comparado com [s]. Tal fato é esperado porque, conforme foi descrito, o valor do F2 é influenciado pela anteriorização da língua. Quanto mais posterior a língua se encontra durante a articulação de determinado som, maior será o valor do segundo formante. Como durante a produção do segmento [ʃ] a língua permanece em uma posição mais posterior do que em [s], o valor do F2 vocálico é maior para [ʃ].



FIGURAS 6 e 7 - F1 e F2 da vogal [a] seguinte à fricativa [ʃ], na palavra /ʃapa/ (à esquerda) e F1 e F2 da vogal [a] seguinte à fricativa [s], na palavra /sapa/ (à direita) – Informante do sexo masculino de 33 anos de idade

As figuras acima mostram o movimento do F1 e do F2 na região de transição dos formantes entre as fricativas [ʃ] e [s] e a vogal [a], nas palavras /ʃapa/ e /sapa/. Na figura da esquerda observa-se um decréscimo do valor do segundo formante (F2) entre a fricativa [ʃ] e a vogal [a], conforme esperado.

A diferença das transições dos formantes vocálicos entre [ba] e [da] se concentra principalmente na região de transição do segundo formante (F2). Desta forma, no estudo piloto, espera-se verificar valores do F2 para [da] maiores quando comparado com [ba], na região de transição dos formantes. As figuras abaixo mostram o movimento do F1 e do F2 na região de transição dos formantes entre as plosivas [b] e [d] e a vogal [a], nas palavras /bata/ e /data/.



FIGURAS 8 e 9 - F1 e F2 da vogal [a] seguinte à plosiva [b], na palavra /bata/ (à esquerda) e F1 e F2 da vogal [a] seguinte à plosiva [d], na palavra /data/ (à direita) – Informante do sexo masculino de 33 anos de idade

Os resultados apresentados adiante com o estudo piloto serão extremamente importantes para a posterior síntese dos estímulos, uma vez que não se verificou estudos na literatura que investigaram os valores dos formantes vocálicos seguintes às fricativas [s] e [ʃ] e às plosivas [b] e [d] do PB, na porção inicial e estável.

A seguir, será apresentada a metodologia do estudo piloto bem com os resultados alcançados. O estudo foi desenvolvido com a aprovação do Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, Parecer CAAE – 62617616.0.0000.5149.

3.4 Metodologia do estudo piloto

3.4.1 Participantes do estudo

Os participantes deste estudo foram sete adultos do sexo masculino, falantes nativos do português brasileiro, com faixa etária entre 20 e 45 anos de idade. Os participantes assinaram o Termo de Consentimento Livre e Esclarecido, após serem informados sobre os aspectos gerais da pesquisa (ANEXO 1).

3.4.2 Procedimento

Os estímulos analisados foram compostos pelos pares mínimos /ʃapa-/sapa/ e /bata-/data/ e pelo par de pseudopalavras /ʃaba-/saba/ do PB.

As gravações dos estímulos foram realizadas em uma sala acusticamente isolada, com a utilização de um notebook (Macbook Air) e de um microfone de cabeça (Philips), a partir do *software* Praat (versão 5.2.35) (BOERSMA e WEENINK, 2011), com a taxa de amostragem de 22050Hz. Após a gravação, os dados de fala foram etiquetados e analisados. Cada palavra foi produzida 3 vezes por cada um dos sete participantes. As palavras foram inseridas na seguinte frase veículo: “Eu digo _____ para ela”.

3.4.3 Análise das propriedades acústicas dos estímulos

A análise acústica dos dados de fala iniciou-se pela inspeção visual da forma de onda, do espectrograma e do oscilograma. Em seguida todas as produções dos informantes foram etiquetadas conforme mostra a figura 10.

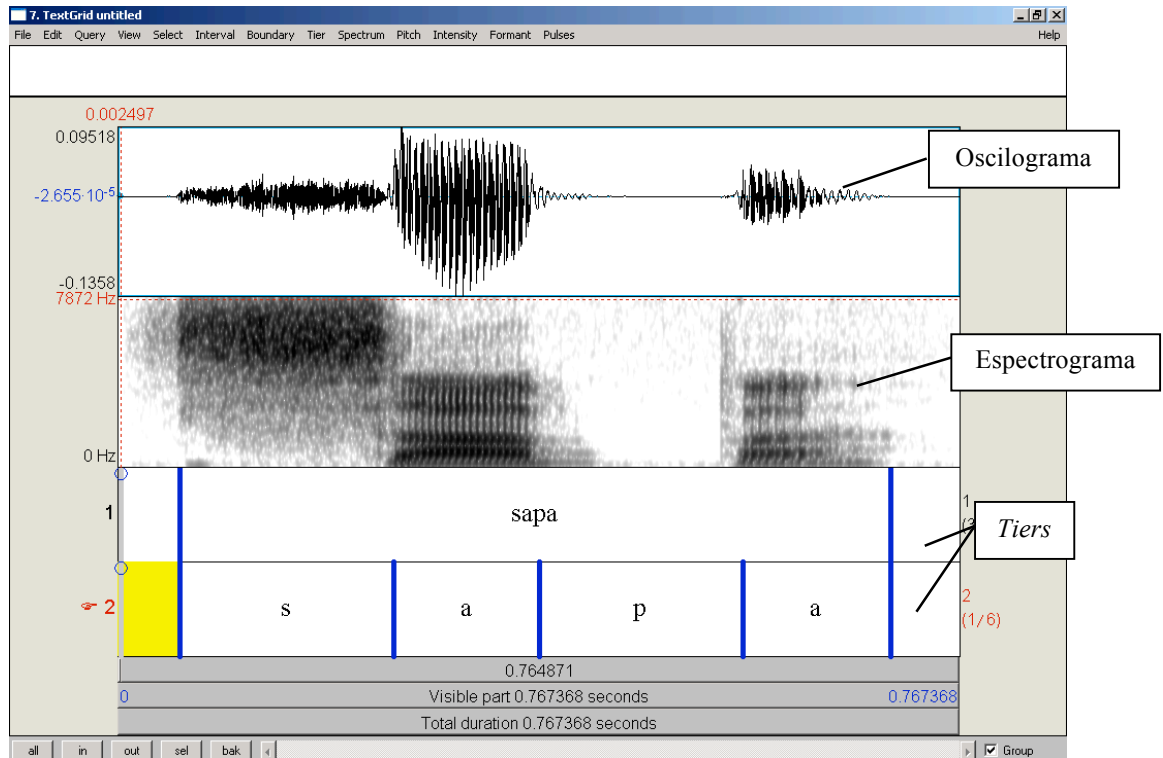


FIGURA 10 - Exemplo da anotação realizada nos dados de fala dos informantes; informante masculino, 34 anos

Na figura 10, a primeira camada (*tier*) representa a etiquetagem de uma das três repetições produzidas por um dos informantes, neste caso a palavra /sapa/. A segunda camada relaciona-se à etiquetagem de cada som separadamente. Após a etiquetagem das palavras em duas camadas, os quatro sons da palavra foram extraídos e analisados separadamente.

Basicamente, a avaliação acústica contemplou a análise de dois tipos de medidas, nos três pares de palavras descritos: as medidas referentes à frequência dos formantes vocálicos e as medidas relativas ao ruído fricativo e às plosivas. O primeiro conjunto de medidas incluiu os valores de frequência dos três primeiros formantes vocálicos (F1, F2 e F3) na porção inicial (transição dos formantes entre a consoante e a vogal) e na porção estável da vogal e os valores de duração destes sons. O segundo conjunto incluiu medidas que quantificam a duração e as características espectrais dos segmentos fricativos e plosivos: picos de amplitudes e formantes do ruído fricativo, picos de amplitudes e formantes da barra de oclusão (do inglês, *burst*), duração total das fricativas e das plosivas, duração da fase de pré-sonorização e da fase de soltura da oclusão das plosivas e centro de gravidade do ruído fricativo. Os valores destas medidas das produções dos informantes foram obtidos de forma manual por meio da utilização do *software* Praat. As configurações do espectrograma, selecionadas no Praat, serão apresentadas nas próximas seções, separadamente.

3.4.3.1 Medidas dos formantes e da duração vocálica

As medidas temporais e espectrais dos três primeiros formantes (F1, F2 e F3) foram obtidas manualmente através da análise do espectrograma e do oscilograma. A análise espectral dos formantes vocálicos foi realizada por meio da seleção dos seguintes parâmetros no Praat, na configuração de espectrograma (“Spectrogram settings”): a) View range (Hz) = 0 a 8000Hz; b) Window length (s) = 0.005 e c) Dynamic range (dB) = 70dB.

Os valores dos formantes vocálicos foram extraídos automaticamente pelo Praat, a partir do método LPC⁸, em dois principais momentos: 1) na porção inicial da vogal (*onset* vocálico; porção de transição dos formantes entre a consoante e a vogal) e na porção estável da vogal, ou seja, no momento em que se observou pequena variação entre os valores dos formantes.

O ponto inicial da vogal, ou seja, de transição dos formantes, foi selecionado com base nas informações oferecidas pelo oscilograma e pelo espectrograma. No oscilograma é possível visualizar uma mudança abrupta na amplitude entre a consoante e a vogal, conforme sinaliza a seta em preto na figura 11.

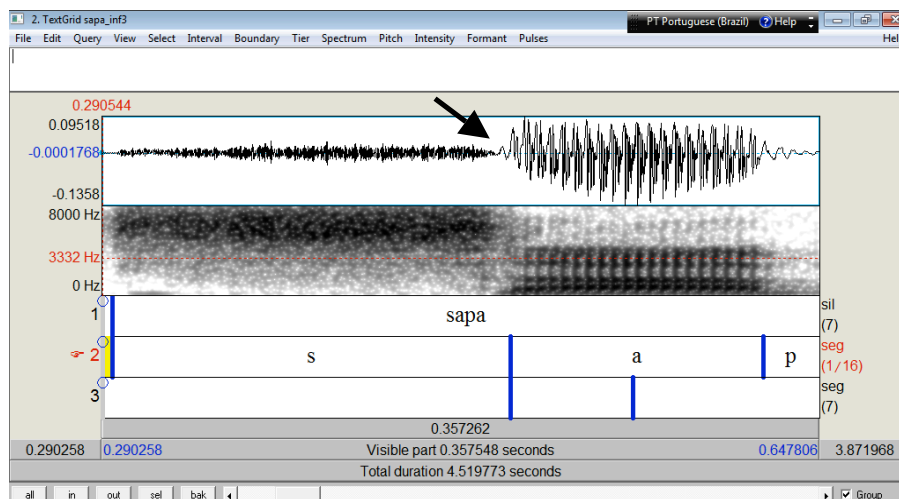


FIGURA 11 - Região de mudança abrupta da amplitude visualizada no oscilograma entre a consoante [s] e a vogal [a], na palavra /sapa/

⁸ LPC (*Linear Prediction Coding*) relaciona-se a uma metodologia de análise do sinal acústico que baseia-se no fato de que qualquer ponto de uma amostra de fala digitalizada é parcialmente previsto a partir do que imediatamente o precedeu, uma vez que a fala não varia amplamente de um ponto a outro. Essa metodologia de decomposição da onda sonora foi escolhida por oferecer valores mais precisos dos formantes vocálicos, das amplitudes do ruído fricativo e das amplitudes da fase de oclusão das plosivas em comparação com a metodologia FFT (*Fast Fourier Transform*).

Após a identificação deste ponto no oscilograma, marcou-se o ponto no qual se observou o início do primeiro pulso glótico e o início do segundo formante vocálico (seta preta) visualizado no espectrograma (figura 12). Neste ponto obteve-se os valores das frequências dos formantes (F1, F2 e F3), que representaram a porção inicial da vogal.

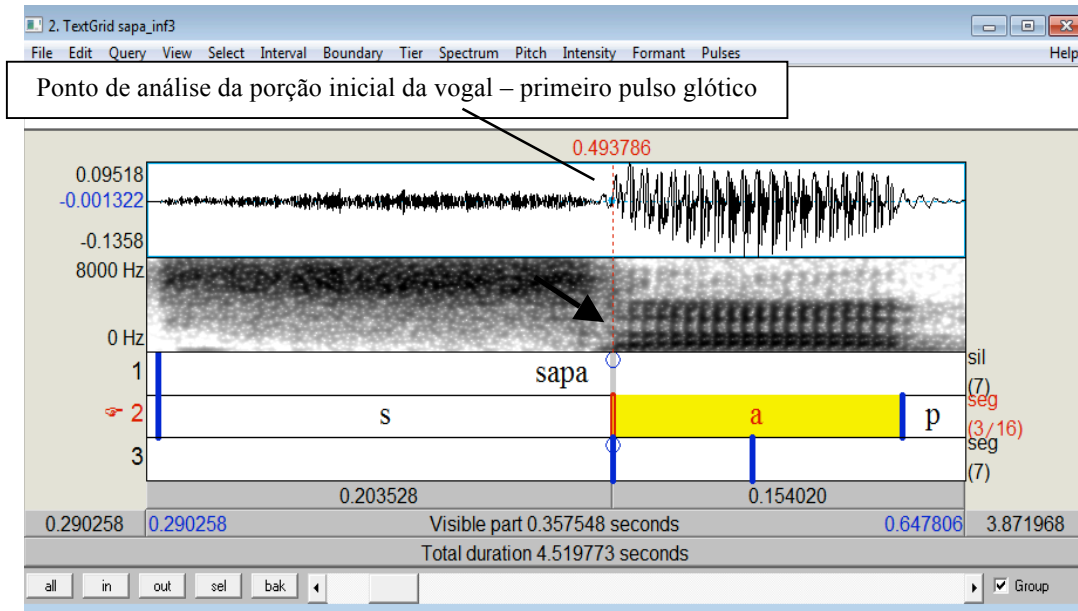


FIGURA 12 - Ponto inicial da vogal [a] na palavra /sapa/. No oscilograma observa-se o primeiro pulso glótico e no espectrograma nota-se o ponto inicial do segundo formante vocálico (seta preta)

No caso da vogal precedida pelas fricativas selecionou-se uma porção de 0.03s de ruído e 70ms de vogal, com o intuito de auxiliar no estabelecimento do ponto inicial da porção vocálica para extração dos valores dos formantes. Em seguida, solicitou-se o “*Formant listing*” disponibilizado pelo Praat, nas configurações em “*Formant*” (figuras 13 e 14).

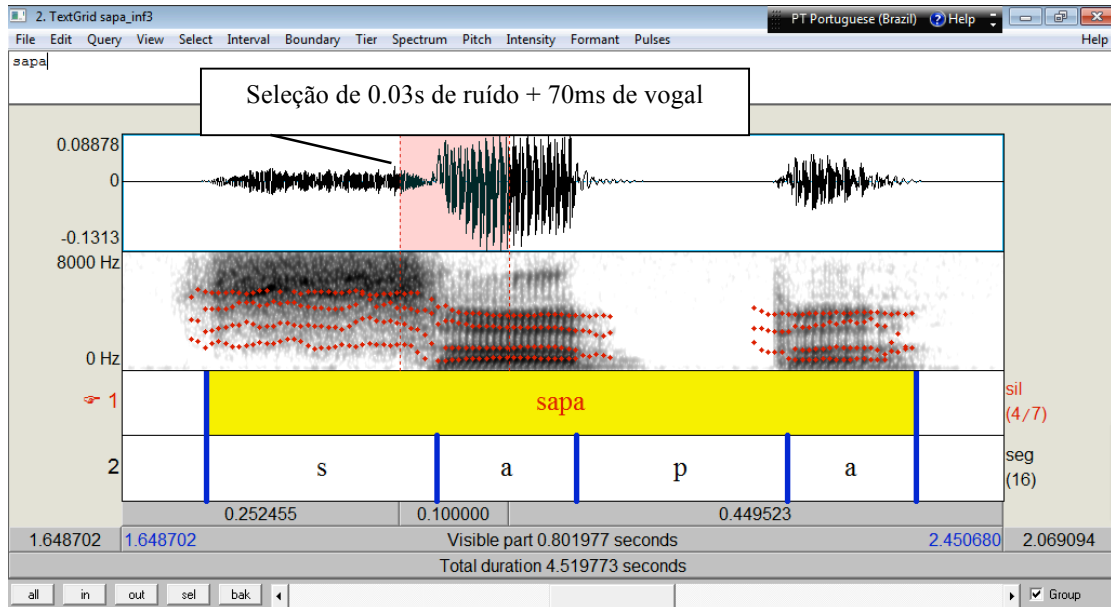


FIGURA 13 - Porção de 0.03s de ruído e 70ms da vogal [a] na palavra /sapa/

Praat Info						
File	Edit	Search	Convert	Font	Help	
Time_s	F1_Hz	F2_Hz	F3_Hz	F4_Hz		
0.226950	540.447533	1356.421671	2948.061515	3726.426185		
0.233200	591.017803	1372.796137	2988.793489	3760.267352		
0.239450	629.360351	1367.175360	2923.145525	3715.496574		
0.245700	645.582470	1350.602860	2783.672989	3707.388175		
0.251950	660.372869	1346.787075	2778.456446	3695.645509		
0.258200	672.901180	1347.673768	2783.389688	3689.106847		
0.264450	677.088245	1355.158827	2748.964160	3631.492513		
0.270700	682.785232	1357.537238	2740.252312	3571.793421		
0.276950	685.254986	1338.376605	2702.110091	3481.355409		
0.283200	686.276567	1347.700515	2670.120319	3469.941007		
0.289450	686.414327	1356.414328	2677.100812	3509.731955		
0.295700	685.148461	1367.684321	2693.599867	3592.485388		
0.301950	679.915353	1370.294534	2678.970271	3613.303733		
0.308200	670.015087	1347.859085	2679.864007	3639.691671		
0.314450	657.056219	1332.192185	2673.968987	3531.318867		
0.320700	650.139619	1323.209642	2672.478843	3593.386579		
0.326950	642.319668	1307.063866	2640.747550	3637.779786		
0.333200	623.979719	1261.534805	2615.461984	3600.544072		
0.339450	600.092306	1206.065767	2577.710964	3637.270452		
0.345700	578.836122	1193.952497	2595.624688	3605.149797		

FIGURA 14 - Tela do Praat que disponibiliza o *Formant listing*, da região selecionada no espectrograma da figura 13

Os valores dos formantes da porção estável da vogal foram definidos pela visualização do movimento dos formantes no espectrograma (pontos vermelhos) e pelos pulsos glóticos regulares no oscilograma. Além disso, com o auxílio do “*Formant listing*”, selecionou-se 20ms da parte estável da vogal para obtenção do F1, F2 e F3.

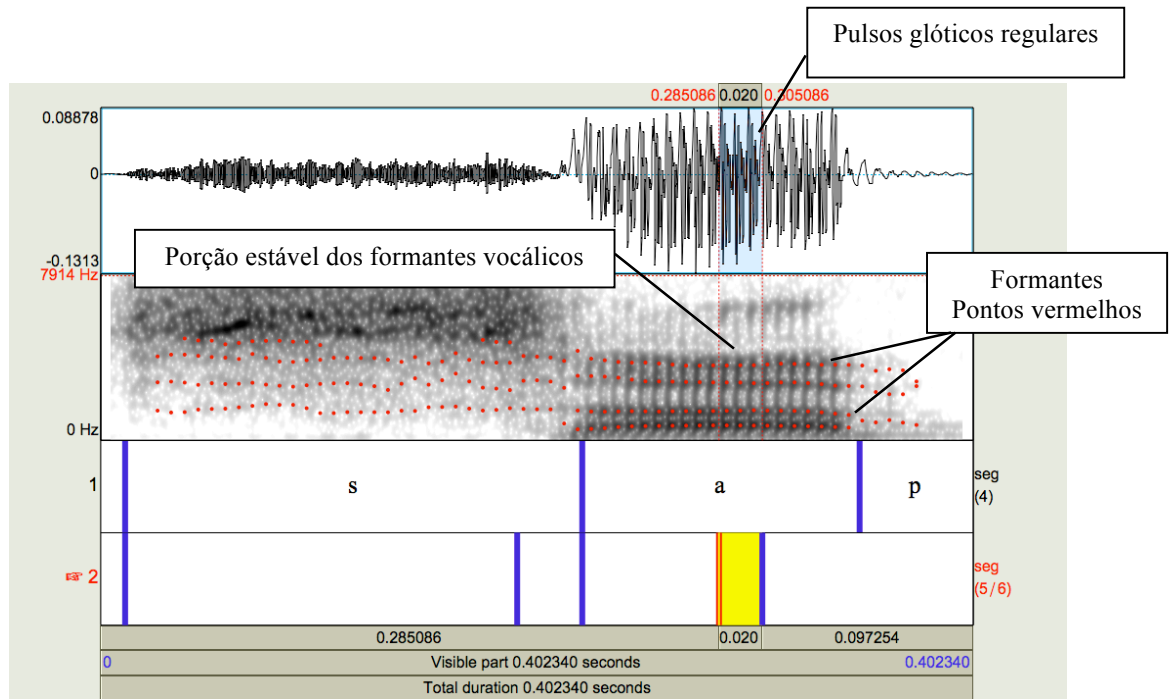


FIGURA 15 - Formantes da primeira vogal [a] da palavra /sapa/ (pontos vermelhos), com indicação dos pulsos glóticos regulares no oscilograma e da porção estável de 20ms da vogal

No caso da vogal precedida pelas plosivas, os valores dos formantes foram obtidos da mesma forma do que nas fricativas. O ponto inicial da vogal foi estabelecido a partir do primeiro pulso glótico vocálico, logo após a barra de soltura da oclusão da consoante plosiva, e com base na identificação do segundo formante (F2). Já os valores dos formantes da porção estável da vogal foram definidos conforme a descrição acima, de acordo com a figura 16.

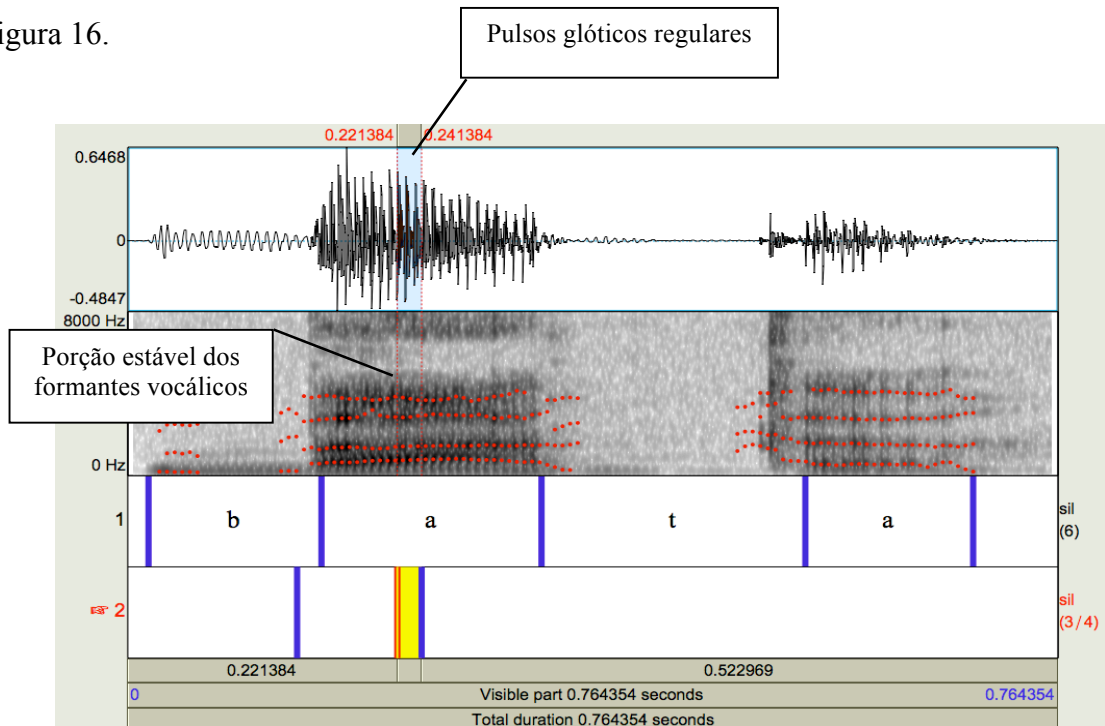


FIGURA 16 - Formantes da primeira vogal [a] da palavra /bata/ (pontos vermelhos), com indicação dos pulsos glóticos regulares no oscilograma e da porção estável de 20ms da vogal

As medidas de duração das vogais foi feita com o auxílio visual da forma de onda juntamente com o espectrograma. Tais medidas foram obtidas selecionando a vogal a partir do primeiro pulso glótico regular até o último pulso regular desta mesma vogal (figura 17).

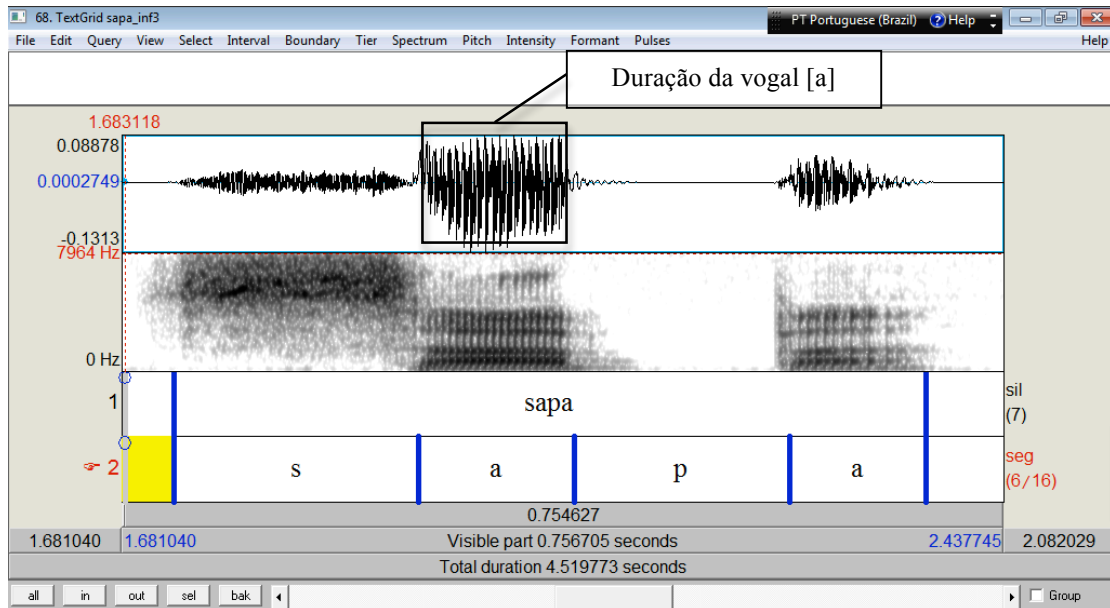


FIGURA 17 - Seleção da vogal tônica [a] da palavra /sapa/

3.4.3.2 Medidas dos segmentos fricativos e plosivos

As medidas temporais e espectrais das fricativas [s] e [ʃ] e das plosivas [b] e [d] também foram obtidas manualmente por meio da inspeção do espectrograma e do oscilograma disponibilizados pelo *software* Praat.

A avaliação acústica dos segmentos fricativos e plosivos foi composta pela análise dos seguintes parâmetros:

- (a) Duração total dos segmentos fricativos e plosivos;
- (b) Centro de gravidade do ruído fricativo;
- (c) Picos de amplitude e formantes do ruído fricativo;
- (d) Duração da pré-sonorização do segmento plosivo vozeado;
- (e) Duração da soltura da oclusão (do inglês, *burst*), representada pela barra de plosão, nos segmentos plosivos;
- (f) Picos de amplitude e formantes da soltura da oclusão.

A descrição detalhada dos parâmetros acústicos analisados nos segmentos fricativos e plosivos supracitados, bem como as configurações selecionadas no Praat será realizada adiante, separadamente.

3.4.3.2.1 Propriedades acústicas dos segmentos fricativos

O ruído fricativo foi selecionado excluindo a região de transição dos formantes conforme mostra a figura 18.

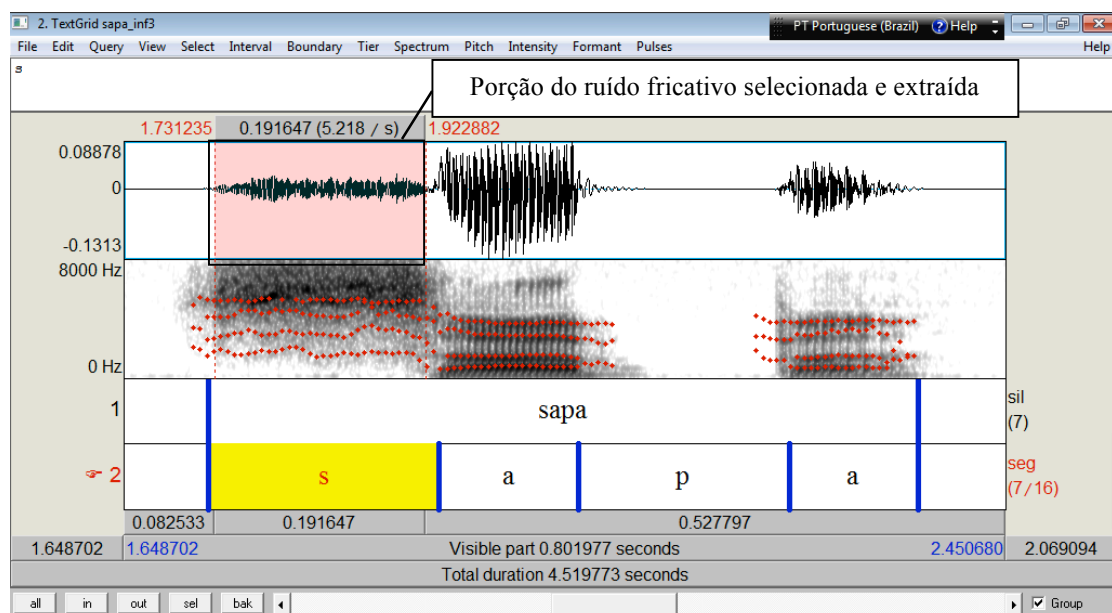


FIGURA 18 - Porção do segmento fricativo [s] na palavra /sapa/ que foi selecionada e extraída para análise

Selecionado o ruído, foram obtidos os valores dos formantes no Praat, utilizando o método LPC, com base nos seguintes parâmetros: *Analyse Spectrum*; *To Formant (Max. Number of formants: 6; Maximum formant (Hz): 8500Hz; Window length(s): 0.025); Tabulate; List*. Com este procedimento, o Praat disponibiliza uma lista composta pelos valores das frequência dos formantes de toda a porção selecionada do ruído fricativo. Após a obtenção destes valores foi criada uma tabela e posteriormente obtidos os valores médios das seis frequências proeminentes do ruído fricativo (F1 a F6).

Feito isso, retornou-se para a janela de objetos do Praat, selecionou-se o ruído fricativo e as seguintes configurações: *Analyse Spectrum*; *To LPC (burg)*; *Extract To Spectrum slice*;

ViewEdit. A figura 19 mostra a fatia do espectro do ruído fricativo com os seus respectivos picos de amplitude.



FIGURA 19 - Fatia do espectro do ruído fricativo [s] com a indicação dos picos de amplitude do quarto e do quinto formante (F4 e F5)

A obtenção dos valores das amplitudes dos formantes do ruído fricativo foi realizada com base nos valores dos formantes. Por exemplo, na figura acima, o valor da amplitude do quarto formante (A4F) foi de 44.5dB. Este valor foi obtido pela média dos valores do quarto formante do ruído, ou seja, 5058Hz.

As fricativas não possuem formantes bem definidos. Desta forma, a localização exata das frequências dos formantes não é muito importante para a síntese dos segmentos fricativos, que será descrita no próximo capítulo. O principal componente na diferenciação da fricativa [s] para [ʃ] é a região do espectro em que o ruído é predominante: frequências mais altas para [s] quando comparado com [ʃ]. No entanto, o que define as regiões do espectro que serão excitadas pelo ruído de fricção, no *software* de síntese de fala Sensyn, são as amplitudes dos ressoadores do trato vocal em paralelo (A2F-A6F)⁹. Independentemente do valor exato dos formantes, a ativação de um ressoador garante a presença de ruído na faixa de frequência em que atua. Assim, a análise dos parâmetros de frequências dos formantes

⁹ Principal parâmetro manipulado no *software* Sensyn para criação do ruído fricativo, nomeado como amplitudes de fricção dos formantes

bem como dos seus valores de amplitude foi realizada para verificar as diferenças mais gerais entre os segmentos fricativos [s] e [ʃ].

Outro parâmetro acústico investigado no ruído fricativo foi o centro de gravidade (do inglês, *center of gravity*, CG). O CG pode ser definido como a frequência abaixo (ou acima) da qual estão concentradas 50% da energia de determinado som.

A investigação dos valores deste parâmetro foi importante para posterior construção do continuum fricativo, de uma faixa de frequência apropriada de [ʃ] para [s].

O arquivo *Spectrum slice* do ruído fricativo foi selecionado e as seguintes configurações foram adotadas na janela de objetos do Praat: *Query*, *Get centre of gravity* (conforme demonstrado na figura 20).

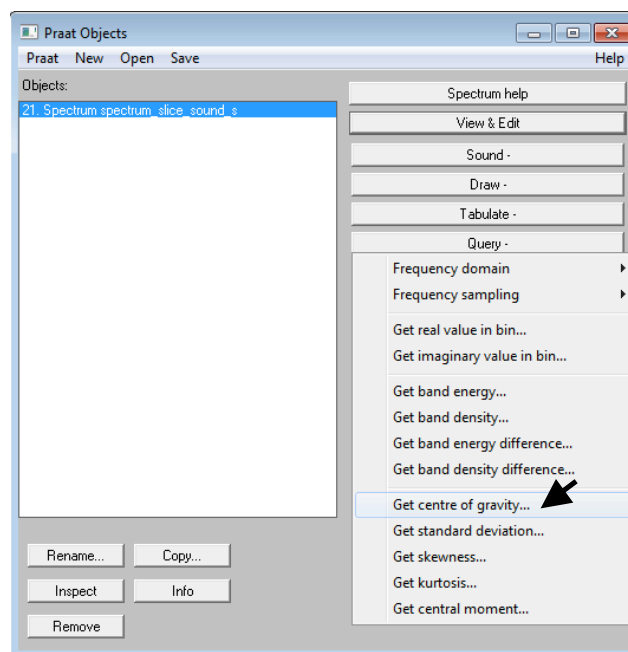


FIGURA 20 - Janela de objetos do Praat para a obtenção do valor referente ao centro de gravidade (seta preta)

O valor do centro de gravidade foi obtido somente com a seleção de todo o ruído fricativo excluindo a região de transição dos formantes. Este parâmetro não foi investigado separadamente, ou seja, porção inicial, medial e final do ruído, devido à constatação de uma pequena variação nos seus valores.

Por fim, a duração total do ruído fricativo foi investigada. Tal medida foi obtida

selecionando o ruído fricativo, sem a porção de transição dos formantes entre o ruído e a vogal subsequente.

3.4.3.2.2 Propriedades acústicas dos segmentos plosivos

O parâmetro acústico mais relevante nos segmentos plosivos, para criação dos estímulos sintetizados, foram as amplitudes do *burst*. Tal parâmetro determina a forma espectral do *burst* nas plosivas, importante para a realização da síntese de fala.

Nos segmentos [b] e [d], as amplitudes foram obtidas selecionando a soltura da oclusão. Para a seleção do *burst* utilizou-se como referência as informações disponibilizadas pelo oscilograma e pela barra de explosão, quando presente (figura 21). Diante da ausência da barra de explosão utilizou-se como referência as informações disponibilizadas pelo oscilograma. Desta forma, selecionou-se o ponto onde se observa a modificação do traçado no oscilograma e o início da produção vocálica com a identificação do segundo formante (F2).

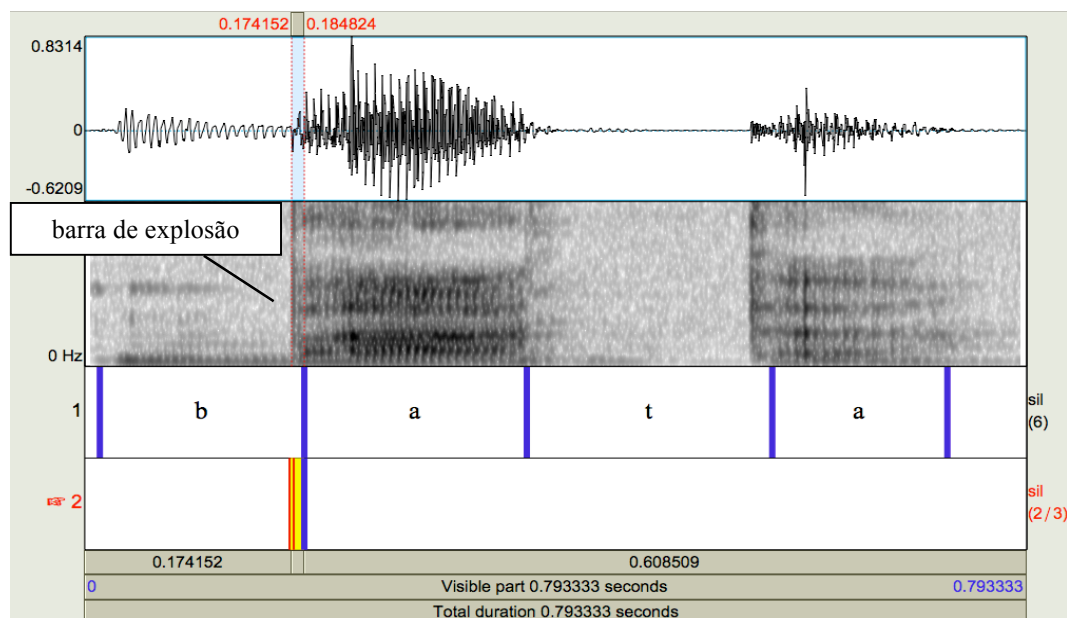


FIGURA 21 - Barra de explosão selecionada na plosiva [b], da palavra /bata/

Após a identificação e a seleção do *burst*, os valores das amplitudes foram obtidos no Praat com base nos seguintes parâmetros: *Analyse Spectrum, To LPC (burg), Extract, to Spectrum Slice*.

Os valores dos formantes do *burst* foram obtidos da mesma forma das fricativas, seguindo os seguintes passos: *Analyse Spectrum; To Formant (Max. Number of formants: 6; Maximum formant (Hz): 5000Hz; Window length(s): 0.025); Tabulate; List.*

A figura 22 mostra as amplitudes dos formantes da plosiva [b] na palavra /bata/. No caso da plosiva bilabial [b] observa-se uma queda nos valores da amplitude.

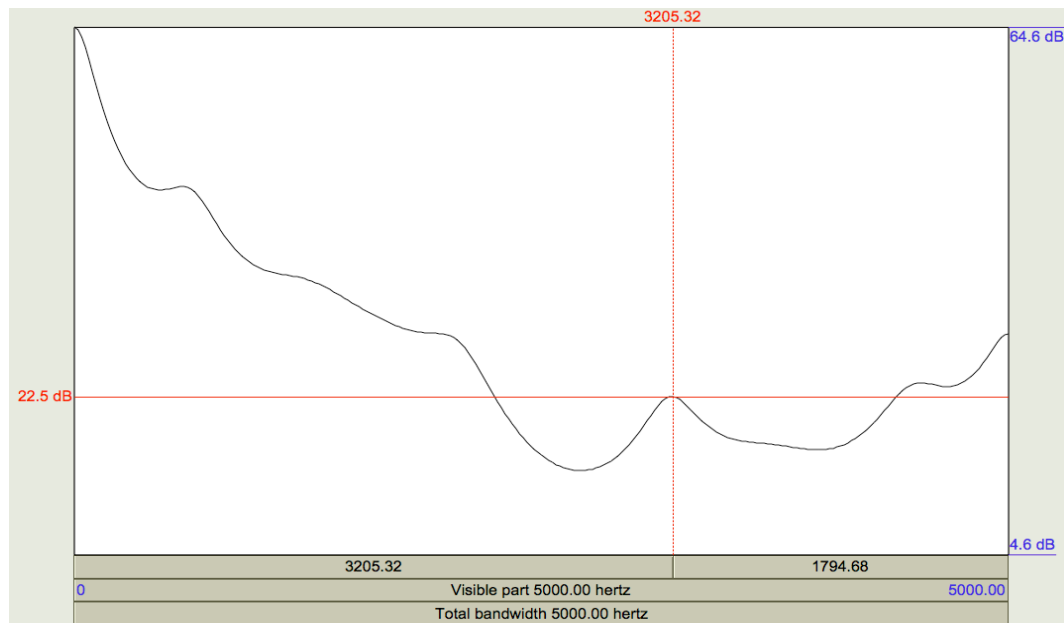


FIGURA 22 - Forma espectral do *burst* da plosiva [b], na palavra /bata/ com a indicação dos picos de amplitude do quarto formante (F4)

Outro parâmetro analisado foi a duração da fase de pré-sonorização e do *burst* dos segmentos plosivos [b] e [d], importantes para a criação dos estímulos sintetizados. A figura 23 mostra a barra de vozeamento, que corresponde à vibração das pregas vocais durante a articulação desses sons e a região selecionada para obtenção da duração do *burst*.

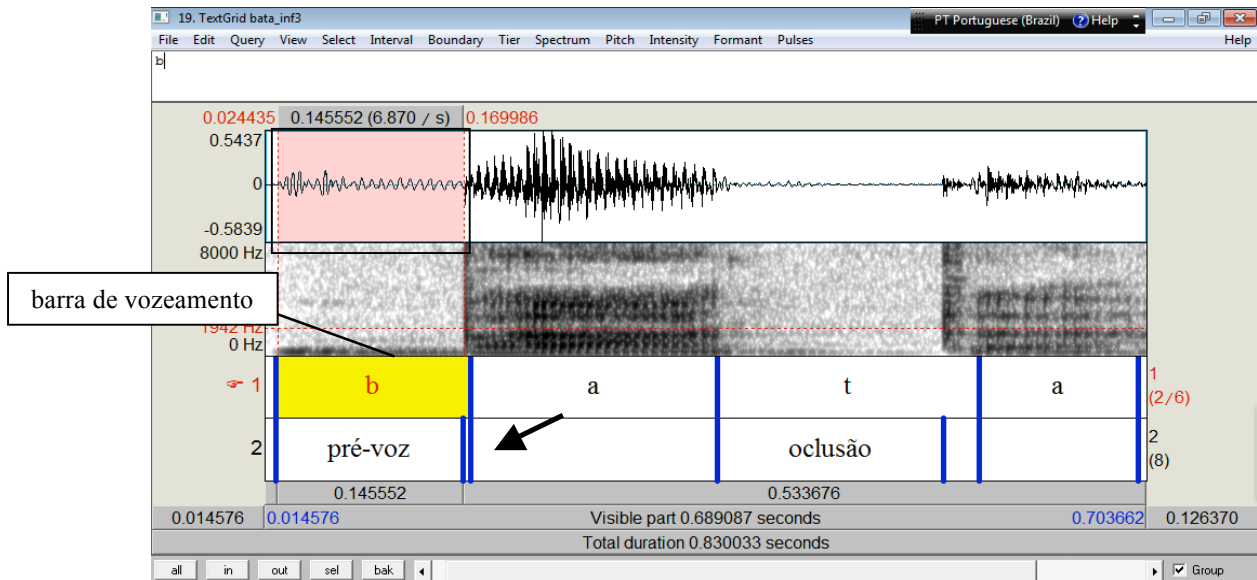


FIGURA 23 - Duração da fase de pré-sonorização do segmento plosivo [b], na palavra /bata/. A seta indica a duração do *burst*, sinalizada pelas duas barras azuis

A seguir serão apresentados os resultados do estudo piloto da análise das propriedades acústicas das fricativas ([s] e [ʃ]) e das plosivas ([b] e [d]) seguidas da vogal [a].

3.5 Resultados e considerações do estudo piloto

Os resultados do estudo piloto serão apresentados por meio de tabelas e gráficos. Foram calculadas as médias para cada parâmetro acústico, considerando todos os sujeitos e, depois, obteve-se os valores do desvio padrão (dp). O desvio padrão é uma medida de dispersão, que mostra como os valores da amostra se distanciaram em relação à média encontrada.

3.5.1 Propriedades acústicas das fricativas [s] e [ʃ]

3.5.1.1 Duração das fricativas

A duração total dos segmentos fricativos pode ser observada na tabela abaixo.

TABELA 1
Duração das fricativas [s] e [ʃ] em (ms)

Duração dos segmentos fricativos		
Informante	Fricativa [s]	Fricativa [ʃ]
1	192	194
2	160	145
3	176	189
4	183	177
5	158	171
6	141	142
7	178	197
Média	169.7	173.5
dp	17.4	22.4

dp = desvio padrão

A duração das fricativas é considerada um parâmetro acústico robusto para diferenciar as fricativas vozeadas das fricativas não vozeadas. Em geral, as fricativas não vozeadas são mais longas do que as vozeadas. Tal fato foi observado nos trabalhos de Samczuk e Gama-Rossi (2004) e Haupt (2007), no português brasileiro.

Como este estudo utilizou apenas as fricativas não vozeadas [s] e [ʃ], a análise dos valores de duração destes sons foi realizada com o objetivo de buscar um valor de duração médio para a posterior síntese. Os valores médios de duração das fricativas [s] e [ʃ] foram 169ms e 173ms; respectivamente. No entanto, as fricativas sintetizadas apresentaram uma duração de 160ms. Este valor foi obtido por meio da média dos valores de duração de [s] e [ʃ] de uma das produções do informante 5, participante desse estudo piloto.

3.5.1.2 Formantes e picos de amplitude do ruído fricativo

A tabela 2 mostra os valores dos seis formantes (F1-F6) das fricativas [s] e [ʃ]. Estes valores não apresentam grande importância para a síntese desses sons. No entanto, só é possível obter os valores dos picos de amplitude dos sons fricativos a partir dos formantes. Para a síntese das fricativas o que importa são os valores dos picos de amplitude (A2F-A6F), que irão definir as regiões do espectro que serão excitadas pelo ruído de fricção. Desta forma, a discussão deste tópico irá concentrar-se nesses valores e a divulgação da tabela será apenas informativa.

TABELA 2
 Frequência dos formantes das fricativas [s] e [ʃ] em (Hz)

Formantes dos segmentos fricativos							
Informante	Fricativa	F1	F2	F3	F4	F5	F6
1	[s]	1904	3369	4635	5396	6224	7116
	[ʃ]	2282	2999	3753	4913	5791	6810
2	[s]	1562	2719	3951	5063	5752	6390
	[ʃ]	1721	2360	3108	4399	5307	6225
3	[s]	1430	2612	3856	5104	5762	6377
	[ʃ]	1943	2554	3627	5430	5541	6270
4	[s]	1547	2813	4148	5120	5837	6440
	[ʃ]	2248	2846	3410	4439	5490	6242
5	[s]	1734	2908	4106	5120	5742	6406
	[ʃ]	1976	2781	3460	4722	5397	6153
6	[s]	1416	2973	4126	4723	5401	6180
	[ʃ]	1993	2451	3287	4432	5519	6167
7	[s]	1674	2980	4337	5061	5892	6505
	[ʃ]	1885	2662	3426	4594	5462	6208
Média	[s]	1609.5	2910.5	4165.5	5083.8	5801.4	6487.7
	[ʃ]	2006.8	2664.7	3438.7	4704.1	5501	6296.4
dp	[s]	174.2	242.9	257.2	196.4	243.5	294.5
	[ʃ]	198.1	226.9	211.3	369.7	150.6	230

dp = desvio padrão

A tabela 3 revela os picos de amplitude das fricativas [s] e [ʃ]. Como a síntese de fala foi realizada por meio da análise dos parâmetros acústicos de um informante e não existem estudos que investigaram os valores dos picos de amplitude das fricativas no PB, esta análise foi desenvolvida com o objetivo de verificar a tendência geral deste parâmetro.

TABELA 3
Amplitude dos formantes das fricativas [s] e [ʃ] em (dB)

Amplitudes dos formantes dos segmentos fricativos						
Informante	Fricativa	A2	A3	A4	A5	A6
1	[s]	17	29	39	38	29
	[ʃ]	43	40	35	34	24
2	[s]	33	34	44	50	41
	[ʃ]	61	50	37	38	30
3	[s]	35	39	44	44	34
	[ʃ]	44	44	39	35	27
4	[s]	31	34	45	44	36
	[ʃ]	55	48	39	37	28
5	[s]	17	23	46	43	28
	[ʃ]	49	41	35	36	26
6	[s]	36	47	52	43	39
	[ʃ]	59	52	46	44	40
7	[s]	30	41	44	47	39
	[ʃ]	57	58	45	40	37
Média	[s]	28.4	35.2	44.8	44.1	35.1
	[ʃ]	52.5	47.5	39.4	37.7	30.2
dp	[s]	8	7.9	3.8	3.7	5
	[ʃ]	7.3	6.4	4.4	3.4	5.9

dp = desvio padrão

A tabela 3 também mostra que as fricativas [s] e [ʃ] diferem em relação aos valores dos picos de amplitude, sendo que a fricativa [s] apresenta os picos de amplitudes mais proeminentes localizados em regiões mais altas de frequência quando comparada com [ʃ]. Assim, a fricativa [s] apresentou valores mais altos da quarta e da quinta amplitude (A4F e A5F), enquanto a fricativa [ʃ] apresentou valores mais altos da segunda e da terceira amplitude (A2F e A3F).

Os informantes 1 e 5 apresentaram valores mais baixos da segunda e terceira amplitudes (A2F e A3F) para a fricativa [s] em relação aos demais informantes. Desta forma, o desvio padrão foi maior nesse caso. Porém, tal fato não foi relevante, uma vez que se analisou a tendência geral do grupo. Se estes informantes fossem excluídos da análise teríamos os valores de 33dB para A2F e 39dB para A3F e, ainda assim, a quarta e a quinta amplitude apresentariam os maiores valores. Tanto informante 1 quanto o informante 5 apresentaram A4F e A5F mais altos em relação aos demais valores de amplitude (A2F, A3F e A6F). O mesmo ocorreu com a fricativa [ʃ]. Estes informantes apresentaram valores mais baixos da

segunda e terceira amplitudes quando comparados com o grupo e, conseqüentemente, o desvio padrão foi maior.

Comparando as tabelas acima observa-se que, como a fricativa [s] apresentou valores maiores da quarta e da quinta amplitude (A4F e A5F), os picos espectrais mais proeminentes deste segmento estão entre 5KHz e 5.8KHz. Estes valores de frequência foram extraídos da tabela 2 e representam os valores médios da quarta e da quinta frequência (F4 e F5) da fricativa [s] (5083.8Hz e 5801.4Hz; respectivamente). Analisando a fricativa [ʃ] observa-se que os picos espectrais mais proeminentes concentram-se entre 2910.5Hz e 3438.7Hz, ou seja, em regiões de frequências mais baixas quando comparada com [s]. Esta questão será confirmada adiante com a extração dos valores do centro de gravidade.

A figura abaixo mostra o espectro acústico das fricativas do árabe egípcio, com as regiões de frequências mais proeminentes. No eixo horizontal temos os valores de frequência em kHz e no eixo vertical temos a amplitude em dB. Na figura 24, assim como ocorreu no PB, observa-se que a fricativa [s] possui dois picos espectrais mais proeminentes, em regiões de frequências mais altas, quando comparada com a fricativa [ʃ]. Conforme descrito, tal fato também foi observado com as fricativas do inglês americano, do português europeu, do espanhol argentino e do italiano (STEVENS e KLATT, 1968; LACERDA, 1982; BORZONE DE MANRIQUE e MASSONE, 1981 e SCHINDLER, 1974).

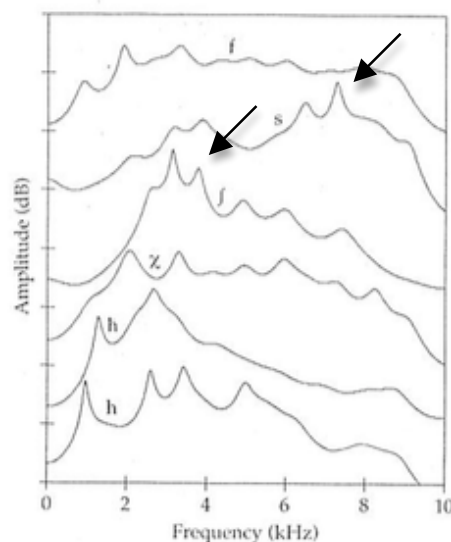


FIGURA 24 - Espectro acústico das fricativas do árabe egípcio. O espectro foi produzido pela análise LPC selecionando a porção média do ruído fricativo. As setas mostram os picos espectrais da fricativas (Johnson, 2011, pp. 130)

A seguir a análise dos valores do centro de gravidade (CG).

3.5.1.3 Centro de gravidade do ruído fricativo

De acordo com a descrição realizada na seção 3.4.3.2.1, o centro de gravidade (CG) pode ser definido como a frequência abaixo (ou acima) da qual estão concentradas 50% da energia de determinado som. A tabela 4 mostra os valores deste parâmetro para as fricativas [s] e [ʃ].

TABELA 4
Centro de gravidade das fricativas [s] e [ʃ] em (Hz)

Centro de gravidade dos segmentos fricativos		
Informante	Fricativa [s]	Fricativa [ʃ]
1	5717	3499
2	5349	2601
3	4994	3200
4	5649	3097
5	5395	3461
6	4401	2745
7	5112	3166
Média	5231	3109.8
dp	449.2	336.2

dp = desvio padrão

Conforme esperado, a fricativa [s] apresentou valor maior do CG do que a fricativa [ʃ] (5231Hz e 3110Hz, respectivamente). Gordon, Barthmaier e Sands (2002) realizaram um estudo acústico das fricativas não vozeadas em sete línguas localizadas na região da América do Norte, da Escócia e da Ásia. Um dos parâmetros analisados foi o CG. Eles concluíram que a fricativa [s] também apresentou valor de CG maior do que a fricativa [ʃ] em seis línguas. O valor do CG da fricativa [ʃ] foi maior em somente uma língua, falada na Índia, chamada Toda.

Quando comparado com os demais estudos, observa-se que os valores do CG obtidos com o estudo piloto são inferiores em relação aos valores divulgados na pesquisa de Jongman, Wayland e Wongs (2000) para o inglês americano. O CG da fricativa [s] foi 6882Hz e da fricativa [ʃ] foi 3712Hz. Os valores obtidos aqui são mais próximos dos valores descritos no trabalho de Gordon, Barthmaier e Sands (2002). O valor maior do CG para a fricativa [s] foi 5463Hz e o menor para a fricativa [ʃ] foi 4134Hz. Para o português brasileiro, Russo e

Behlau (1993), observaram que a fricativa [s] possui faixas de frequência acima de 4500Hz chegando a 8000Hz, enquanto a fricativa [ʃ] apresenta uma faixa de frequência entre 2500 e 6000Hz, valores compatíveis com os alcançados neste estudo.

3.5.2 Propriedades acústicas das plosivas [b] e [d]

3.5.2.1 Duração absoluta dos segmentos

A medida de duração absoluta dos segmentos plosivos foi obtida por meio da análise da fase de pré-sonorização e da fase de soltura da oclusão, ou seja, *burst*, conforme mostram as tabelas 5 e 6.

TABELA 5
Duração da fase de pré-sonorização [b] e [d] em (ms)

Duração da fase de pré-sonorização		
Informante	Plosiva [b]	Plosiva [d]
1	146	96
2	74	52
3	101	56
4	97	80
5	93	75
6	92	72
7	108	73
Média	101.5	72
dp	25.7	28.6

dp = desvio padrão

TABELA 6
Duração da fase de soltura da oclusão [b] e [d] em (ms)

Duração da fase de soltura da oclusão (<i>Burst</i>)		
Informante	Plosiva [b]	Plosiva [d]
1	13	13
2	11	12
3	12	15
4	12	9
5	11	10
6	9	12
7	12	11
Média	11.5	11.8
dp	1.69	2.7

dp = desvio padrão

O valor de duração absoluta da fase de pré-sonorização encontrado neste estudo foi maior para a plosiva [b] quando comparado com a plosiva [d]. Outros trabalhos do PB também revelaram resultados semelhantes. Melo *et al.* (2011) investigaram e compararam as características acústicas das plosivas vozeadas e não vozeadas na fala de crianças com desenvolvimento fonológico típico e de adultos, que pertenciam ao grupo controle. Os pesquisadores analisaram a fala de 11 adultos e obtiveram os valores de duração da fase de oclusão de 98ms para [b] e 91ms para [d], em posição de *onset* medial.

Em geral, observa-se uma diferença no valor de duração dos segmentos plosivos quando se compara as plosivas vozeadas com as plosivas não vozeadas, sendo que estas últimas possuem valores de duração maiores. Como neste estudo utilizou-se as plosivas vozeadas [b] e [d], a análise da duração destes segmentos foi importante somente para a posterior criação dos estímulos sintéticos, com base no valor real de duração dos mesmos. No entanto, Barbosa (1999) também verificou valores da fase de pré-sonorização maiores para a plosiva [b] (86ms) quando comparada com a plosiva [d] (71ms). Da mesma forma, Alves (2015), em seu estudo dos parâmetros acústicos direcionado às plosivas, realizado com cinco informantes do sexo feminino, oriundas da região de Criciúma, no sul do Estado de Santa Catarina, observou que a plosiva [b] apresentou o valor de duração absoluta de 101ms, enquanto a plosiva [d] apresentou a duração de 94ms.

Por fim, a tabela 6 mostra os valores de duração da fase de soltura da oclusão para as plosivas [b] e [d], onde se observa que os valores deste parâmetro foram muito próximos comparando estes dois segmentos (11.5ms para [b] e 11.8ms para [d]; respectivamente). Lieberman e Blumstein (1988) verificaram que a duração do *burst* pode ser de 5 a 15ms, o que corrobora com os achados deste estudo.

A seguir será apresentada a análise de outro parâmetro acústico considerado relevante para a discriminação das plosivas [b] e [d] e para a criação dos estímulos sintetizados, que são as características do *burst*.

3.5.2.2 Picos de amplitude e formantes da soltura da oclusão

Nesta seção realizou-se uma análise qualitativa dos parâmetros de amplitude e dos formantes do *burst* das plosivas [b] e [d]. Tal análise tornou-se necessária pela curta duração

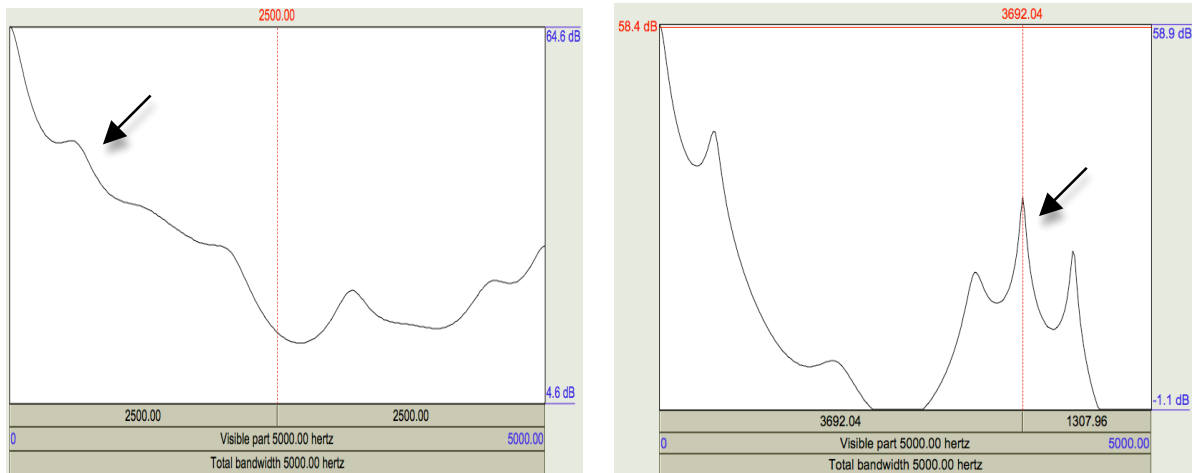
deste parâmetro. Desta forma, em 11 produções da plosiva [b] e em 5 produções da plosiva [d] não foi possível obter o valor exato do sexto formante (F6). No entanto, como as plosivas não possuem valores altos de frequência, que ultrapassem os 6000Hz, o valor do sexto formante não foi considerado relevante para a posterior síntese dos estímulos. Assim como nos sons fricativos, são os picos de amplitude que irão determinar as regiões do espectro das plosivas que serão excitadas pelo ruído de fricção.

Experimentos perceptuais, realizados com as plosivas sintetizadas no inglês, mostraram que o *burst* é considerado uma pista auditiva importante para que os ouvintes façam a distinção entre esta categoria de sons. Estes estudos também verificaram que o pico espectral do *burst* apresentou valores mais altos para as plosivas alveolares (acima de 4000Hz) quando comparado com as plosivas bilabiais (entre 500Hz e 1500Hz) (STEVENS e BLUMSTEIN, 1977; HALLE, HUGHES e RADLEY, 1957).

No português brasileiro, Alves (2015) realizou um estudo das características espectrais do *burst* das plosivas vozeadas e não vozeadas, que envolveu a análise dos picos e dos momentos espectrais (centróide ou média, variância, assimetria e curtose). Direcionando a análise das plosivas vozeadas [b] e [d], assim como no inglês, a pesquisadora também observou que a plosiva bilabial apresentou pico de energia em regiões mais baixas de frequência quando comparada com a plosiva alveolar (1189Hz e 2324Hz; respectivamente). No entanto, apesar da plosiva alveolar apresentar o valor do pico espectral mais elevado do que a plosiva bilabial, esse valor ficou abaixo dos valores obtidos no inglês.

Neste estudo, não realizou-se uma análise detalhada dos momentos espectrais do *burst*, uma vez que estas informações não são relevantes para a posterior síntese dos estímulos. Conforme relatado, a análise do *burst* concentrou-se nos picos de amplitude, parâmetro fundamental para a geração dos estímulos sintéticos. Desta forma, após a extração e a análise das amplitudes do *burst* observou-se que a plosiva bilabial [b] apresentou picos de energia mais intensos em regiões mais baixas de frequência, referente a região do segundo e terceiro formantes (A2F-A3F). Já a plosiva alveolar [d] apresentou picos de energia mais intensos em regiões mais altas de frequência, compatível com a região de frequência do quarto e do quinto formantes (A4F-A5F). A forma espectral da plosiva [b] foi mais plana quando comparada com a plosiva [d]. Tal fato pode ser visualizado nas figuras adiante. À

esquerda observa-se a forma espectral do *burst* da plosiva [b] na palavra /bata/ e da plosiva [d] na palavra /data/.



FIGURAS 25 e 26 - Forma espectral do *burst* da plosiva [b] (esquerda) e da plosiva [d] (direita)

Esta tendência corrobora com os achados descritos anteriormente para o inglês e para o PB.

A seguir, será abordado o último parâmetro analisado no estudo piloto que são as medidas de duração vocálica e dos valores dos formantes da vogal [a] seguinte às fricativas e às plosivas.

3.5.3 Medidas da duração e dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] e às plosivas [b] e [d]

As medidas de duração e dos formantes da vogal [a] seguinte às fricativas e às plosivas também foi importante para a posterior criação dos estímulos sintetizados, uma vez que buscou-se gerar estímulos com os valores reais do PB. Além disso, conforme descrito, não verificou-se estudos na literatura que investigaram os valores dos formantes vocálicos seguinte às fricativas [s] e [ʃ] e às plosivas [b] e [d] do PB, na porção inicial e estável.

3.5.3.1 Medidas da duração e dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ]

A tabela abaixo mostra os valores de duração da vogal [a] seguinte às fricativas [s] e [ʃ].

TABELA 7
Duração da vogal [a] seguinte às fricativas (ms)

Duração da vogal [a]		
Informante	[a] /sa/	[a] /ja/
1	127	129
2	142	155
3	173	178
4	150	155
5	119	131
6	145	149
7	170	177
Média	146.5	153.4
dp	20.1	19.5

Legenda: [a]_/sa/ - valores médios da duração da vogal [a] seguinte a fricativa [s]; [a]_/ja/ - valores médios da duração da vogal [a] seguinte a fricativa [j] e dp = desvio padrão

Observa-se que a vogal [a] seguinte às fricativas [s] e [j] apresentou valores de duração próximos (146ms e 153ms; respectivamente). Desta forma, a vogal [a] sintetizada apresentou valor de duração de 140ms.

A tabela 8 mostra outro parâmetro acústico relevante para a síntese das vogais, que são os valores dos três primeiros formantes da vogal [a], na região de transição dos formantes e na região estável da vogal. Os valores dos formantes na região de transição foram retratados na tabela com a palavra “inicial”, ou seja, os valores do F1, F2 e F3 inicial relacionam-se à região de transição dos formantes entre o segmento fricativo e a vogal. Além disso, a descrição [a]_/sa/, na coluna “vogal [a]” da tabela, refere-se aos valores dos formantes da vogal [a] seguinte a fricativa [s], enquanto a descrição [a]_/ja/ relaciona-se aos valores dos formantes da vogal [a] seguinte a fricativa [j].

TABELA 8
 Frequência dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] em (Hz)

Frequência dos formantes da vogal [a] seguinte às fricativas							
Informante	Vogal [a]	F1 inicial	F1 estável	F2 inicial	F2 estável	F3 inicial	F3 estável
1	[a]_/sa/	526	717	1433	1413	3084	2837
	[a]_/ʃa/	505	687	1550	1396	3077	2844
2	[a]_/sa/	529	739	1211	1210	2491	2447
	[a]_/ʃa/	502	732	1530	1228	2331	2335
3	[a]_/sa/	496	661	1263	1208	2570	2397
	[a]_/ʃa/	441	631	1607	1218	2602	2242
4	[a]_/sa/	554	679	1375	1415	2595	2503
	[a]_/ʃa/	545	670	1694	1536	2775	2767
5	[a]_/sa/	616	768	1389	1366	2863	2584
	[a]_/ʃa/	613	819	1579	1369	2878	2444
6	[a]_/sa/	537	662	1285	1311	2453	2310
	[a]_/ʃa/	543	696	1511	1370	2375	2300
7	[a]_/sa/	582	768	1224	1348	2400	2366
	[a]_/ʃa/	540	797	1641	1362	2984	2379
Média	[a]_/sa/	548.5	713.4	1311.4	1324.4	2636.5	2492
	[a]_/ʃa/	527	718.8	1587.4	1354.1	2717.4	2473
dp	[a]_/sa/	39.7	46.9	87.1	86.7	247.8	176.8
	[a]_/ʃa/	52.6	68.2	64.8	107.8	291.3	236.6

Legenda: [a]_/sa/ - valores médios dos formantes da vogal [a] seguinte a fricativa [s]; [a]_/ʃa/ - valores médios dos formantes da vogal [a] seguinte a fricativa [ʃ] e dp = desvio padrão

Conforme esperado, observa-se uma pequena variação nos valores do primeiro formante (F1), tanto na região de transição dos formantes quanto na porção estável da vogal. Como este estudo analisou somente a vogal [a], tal fato era esperado, uma vez que o F1 relaciona-se ao grau de abertura de uma vogal e a altura da língua.

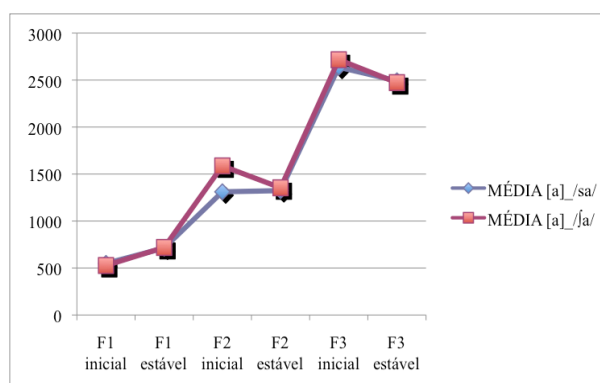
O valor do segundo formante (F2), na região de transição dos formantes, foi maior para a vogal [a] seguinte à fricativa [ʃ] quando comparada com a fricativa [s] (1578Hz e 1311Hz, respectivamente). Tal fato também era esperado porque o valor do F2 está diretamente relacionado com o grau de anteriorização da língua. Como na produção do segmento [ʃ] a língua encontra-se mais posterior na cavidade oral quando comparado com o segmento [s], espera-se o que valor do F2 seja mais alto para [ʃ]. No entanto, observa-se uma pequena variação nos valores médios da frequência do F2 da vogal [a] seguinte a fricativa [s], quando compara-se os valores da região de transição dos formantes e da região estável da vogal (1311-1324Hz).

Nos estudos de ponderação de pistas auditivas desenvolvidos por Nittrouer no inglês americano, a pesquisadora produziu a vogal [a] sintetizada, que seguiu as fricativas [s] e [ʃ], com os seguintes valores (inicial e estável) dos formantes: F1 (450Hz-650Hz), F2 (1250Hz-1130Hz) e F3 (2464Hz-2300Hz) para a vogal [a] seguinte a fricativa [s]; F1 (450Hz-650Hz), F2 (1570Hz-1130Hz) e F3 (2000Hz-2300Hz) para a vogal [a] seguinte a fricativa [ʃ] (NITTROUER, 2002). Em alguns estudos a pesquisadora manipulou somente os valores do F2 e manteve os valores do F1 e do F3 iguais, tanto para a fricativa [s] quanto para fricativa [ʃ] (NITTROUER e MILLER, 1997, experimento II). Desta forma, constata-se uma variação maior dos valores do F2 (inicial e estável) da vogal [a] seguinte a fricativa [ʃ] (440Hz) quando comparado com a fricativa [s] (120Hz), assim como verificou-se no presente estudo. Porém, observa-se que esta variação nos valores do F2 (inicial e estável) da vogal [a] seguinte a fricativa [s], verificada neste estudo (1311Hz-1324Hz), foi menor quando comparada com os estudos desenvolvidos por Nittrouer (1250Hz-1130Hz), para o inglês americano.

O gráfico 1 ilustra os valores médios dos formantes (F1, F2 e F3) da porção inicial e estável da vogal [a] seguinte às fricativas [s] e [ʃ], apresentados na tabela acima.

GRÁFICO 1

Formantes da porção inicial e estável da vogal [a] seguinte às fricativas [s] e [ʃ]



Legenda: média [a]_/sa/ - valores médios dos formantes da vogal [a] seguinte a fricativa [s] e [a]_/ʃa/ - valores médios dos formantes da vogal [a] seguinte a fricativa [ʃ]

O gráfico 1 mostra valores bem próximos do F1 e do F3 da vogal [a] seguinte às fricativas [s] e [ʃ]. No entanto, conforme esperado, observa-se uma variação nos valores inicial e estável do segundo formante, F2, com valor do F2 mais alto para a vogal [a] seguinte a fricativa [ʃ]. Cabe ressaltar que os valores inicial e estável do F2 da vogal [a] seguinte a fricativa [s] apresentou uma pequena variação, conforme mostra a tabela 8 e o gráfico 1.

A seguir serão discutidos os valores de duração e dos formantes da vogal [a] seguinte às plosivas [b] e [d].

3.5.3.2 Medidas da duração e dos formantes da vogal [a] seguinte às plosivas [b] e [d]

A tabela 9 mostra que a vogal [a] seguinte às plosivas [b] e [d] também apresentou valores de duração próximos (166ms e 154ms; respectivamente). O valor de duração da vogal [a] sintetizada, seguinte às plosivas, foi de 140ms, assim como ocorreu com a vogal [a] seguinte às fricativas [s] e [ʃ].

TABELA 9
Duração da vogal [a] seguinte às plosivas em (ms)

Duração da vogal [a]		
Informante	[a]_/ba/	[a]_/da/
1	185	133
2	143	101
3	165	188
4	145	151
5	149	189
6	171	165
7	204	157
Média	166	154.8
dp	22.6	30.9

Legenda: [a]_/ba/ - valores médios da duração da vogal [a] seguinte a plosiva [b]; [a]_/da/ - valores médios da duração da vogal [a] seguinte a plosiva [d] e dp = desvio padrão

Os valores dos três primeiros formantes (F1, F2 e F3) da vogal [a] seguinte às plosivas [b] e [d] podem ser visualizados na tabela 10.

TABELA 10
 Frequência dos formantes da vogal [a] seguinte às plosivas [b] e [d] em (Hz)

Frequência dos formantes da vogal [a] seguinte às plosivas							
Informante	Vogal [a]	F1 inicial	F1 estável	F2 inicial	F2 estável	F3 inicial	F3 estável
1	[a]_/ba/	603	715	1268	1374	2725	2804
	[a]_/da/	419	721	1814	1608	2568	2396
2	[a]_/ba/	513	751	1062	1223	2397	2505
	[a]_/da/	487	669	1717	1596	2297	2183
3	[a]_/ba/	457	625	1045	1016	2502	2417
	[a]_/da/	439	750	1972	1537	2572	2572
4	[a]_/ba/	576	692	1199	1356	2233	2441
	[a]_/da/	422	625	2090	1317	2708	2562
5	[a]_/ba/	524	744	1146	1258	2490	2480
	[a]_/da/	518	803	1826	1742	2436	2546
6	[a]_/ba/	508	739	1111	1265	2352	2474
	[a]_/da/	465	640	1996	1563	2806	2424
7	[a]_/ba/	586	720	1166	1288	2349	2321
	[a]_/da/	544	889	1701	1485	2968	2310
Média	[a]_/ba/	538.1	712.2	1142.4	1254.2	2435.4	2491.7
	[a]_/da/	470.5	728.1	1873.7	1549.7	2622.1	2427.5
dp	[a]_/ba/	52	43.4	77.9	118.1	157.1	150.2
	[a]_/da/	48.2	94.7	148.1	129.8	226	145.8

Legenda: [a]_/ba/ - valores médios dos formantes da vogal [a] seguinte a plosiva [b]; [a]_/da/ - valores médios dos formantes da vogal [a] seguinte a plosiva [d] e dp = desvio padrão

Conforme esperado, a tabela 10 mostra que o valor do F2 para [da] é maior quando comparado com [ba], na região de transição dos formantes (1873Hz e 1142Hz; respectivamente). Isso ocorre, pelo fato da plosiva alveolar [d] apresentar ponto de articulação mais posterior do que a plosiva bilabial [b].

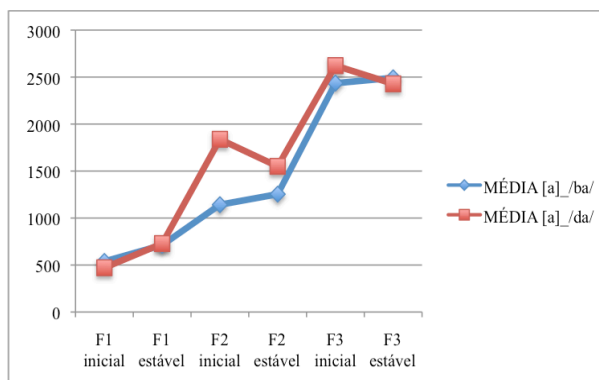
Nos estudos que investigaram a percepção do ponto de articulação das plosivas do inglês americano, desenvolvidos por Stevens e Blumstein (1978) e Walley e Carrell (1983), os pesquisadores produziram a vogal [a] sintetizada, que seguiu as plosivas [b] e [d], com os seguintes valores (inicial e estável) do segundo e do terceiro formantes: F2 (900Hz-1240Hz) e F3 (2000Hz-2500Hz) para a vogal [a] seguinte a plosiva [b]; F2 (1700Hz-1240Hz) e F3 (2800Hz-2500Hz) para a vogal [a] seguinte a plosiva [d]. Estes estudos verificaram valores maiores do F2 para a vogal [a] seguinte a plosiva [d]. Além disso, observa-se um aumento de 340Hz dos valores do F2 (inicial e estável) da vogal [a] seguinte a plosiva [b] e uma diminuição de 460Hz dos valores do F2 da vogal [a] seguinte a plosiva [d].

No presente estudo, nota-se um aumento de 112Hz dos valores médios do F2 (inicial e estável) da vogal [a] seguinte a plosiva [b] e uma diminuição de 324Hz dos valores do F2 da vogal [a] seguinte a plosiva [d]. Por fim, a tabela 10 revela um decréscimo nos valores do terceiro formante (inicial e estável) da vogal [a] seguinte a plosiva [d], o que foi verificado nos estudos desenvolvidos com o inglês (STEVENS e BLUMSTEIN, 1978; WALLEY e CARRELL, 1983).

O gráfico abaixo retrata os valores dos formantes F1, F2 e F3, da porção inicial e estável da vogal [a] seguinte às plosivas [b] e [d], explicitados na tabela 10.

GRÁFICO 2

Formantes da porção inicial e estável da vogal [a] seguinte às plosivas [b] e [d]



Legenda: média [a]/ba/ - valores médios dos formantes da vogal [a] seguinte a plosiva [b] e [a]/da/ - valores médios dos formantes da vogal [a] seguinte a plosiva [d]

O gráfico mostra valores próximos do F1 da vogal [a] seguinte às plosivas [b] e [d]. Copper *et al.* (1952) relata que o primeiro formante da transição pode ser considerado uma pista importante na percepção do vozeamento entre duas categorias de plosivas. Desta forma, espera-se que os valores do F1 sejam diferentes quando comparamos os segmentos plosivos vozeados com os não vozeados, o que não é o caso neste estudo que utilizou somente as plosivas vozeadas [b] e [d].

Conforme esperado, observa-se uma variação nos valores inicial e estável do segundo e do terceiro formantes, com valores do F2 e F3 mais altos para a vogal [a] seguinte a plosiva [d].

3.5.4 Resumo das principais conclusões do estudo piloto

Esta seção apresenta, de forma resumida, os principais resultados do estudo piloto.

1. Duração dos segmentos fricativos: as fricativas [ʃ] e [s] apresentaram valores de duração próximos.
2. Picos de amplitude dos segmentos fricativos: a fricativa [s] apresentou picos de amplitude mais proeminentes em regiões mais altas de frequência quando comparada com a fricativa [ʃ].
3. Centro de gravidade do ruído fricativo: a fricativa [s] apresentou valor médio do CG maior do que a fricativa [ʃ] (5231Hz e 3110Hz; respectivamente).
4. Duração dos segmentos plosivos: a duração absoluta da fase de pré-sonorização foi maior para a plosiva [b] quando comparada com a plosiva [d]. No entanto, observou-se valores próximos de duração da fase de soltura da oclusão.
5. Picos de amplitudes do *burst*: a plosiva alveolar [d] apresentou picos de energia mais intensos em regiões mais altas de frequência quando comparada com a plosiva bilabial [b].
6. Duração da vogal [a] seguinte às fricativas e às plosivas: verificaram-se pequenas diferenças nos valores de duração da vogal [a] seguinte às fricativas [ʃ] e [s] e às plosivas [b] e [d].
7. Medida dos formantes da vogal [a] seguinte às fricativas: observou-se pequenas diferenças nos valores do F1 e do F3 da vogal [a] seguinte às fricativas [ʃ] e [s]. Já o F2 apresentou variação nos valores inicial e estável, com valores mais altos para a vogal [a] seguinte a fricativa [ʃ].
8. Medida dos formantes da vogal [a] seguinte às plosivas: observou-se valores próximos do F1 da vogal [a] seguinte às plosivas [b] e [d]. Porém, verificou-se uma variação nos valores inicial e estável do F2 e do F3, com valores maiores para a vogal [a] seguinte a plosiva [d].

Após descrever e analisar as propriedades acústicas das fricativas [ʃ] e [s] e das plosivas [b] e [d], o capítulo adiante irá abordar o passo a passo do processo de síntese dos estímulos.

4 CAPÍTULO 2 – SÍNTESE DE FALA DOS ESTÍMULOS

O presente capítulo possui o objetivo de descrever o passo a passo da criação dos pares de palavras /sapa/-/ʃapa/, /saba/-/ʃaba/ e /bata/-/data/, utilizados no experimento de ponderação de pistas auditivas, por meio da síntese de fala. Desta forma, serão apresentadas as propriedades acústicas que foram manipuladas no *software* Sensyn. Além disso, será realizada uma breve explanação sobre os tipos de sintetizadores de fala, com enfoque maior no sintetizador de formantes proposto por Klatt (1979).

A análise das propriedades acústicas realizada no capítulo anterior serviu como ponto de partida para a síntese dos pares de palavras. Diante da necessidade, algumas modificações foram executadas nos valores dos parâmetros com o intuito de fornecer maior naturalidade aos estímulos.

4.1 Síntese do sinal de fala

A síntese do sinal de fala é considerada um importante recurso para os estudos de percepção, uma vez que diversas propriedades acústicas podem ser manipuladas para a criação dos estímulos. Grande parte dos estudos direcionados à investigação do desenvolvimento da ponderação de pistas auditivas utiliza a síntese de fala, que é um processo que produz a fala artificialmente.

O sinal artificial de fala é produzido por um sintetizador de fala que, implementado por meio de um *software*, possui como objetivo simular o trato vocal da forma mais próxima da realidade.

De maneira geral, os *softwares* de síntese de fala empregam diferentes configurações para alcançar uma aproximação de alta qualidade com a fala humana. Todos os estímulos deste estudo foram criados com a utilização do sintetizador de formantes, desenvolvido por Klatt (1980), implementado no *software* Sensyn. Por esta razão, este modelo será descrito a seguir de forma mais detalhada.

4.2 Principais tipos de sintetizadores

Historicamente, os sintetizadores de fala podem ser divididos em dois amplos modelos: o modelo de síntese articulatória e o modelo de síntese de formantes.

4.2.1 Sintetizador articulatório

O modelo de síntese articulatória busca reproduzir fielmente os movimentos mecânicos dos articuladores e as distribuições resultantes do volume, velocidade e pressão sonora dos pulmões, laringe, trato vocal e nasal (FLANAGAN, ISHIZAKA e SHIPLEY, 1975). Neste modelo, os mecanismos de produção da fala podem ser reproduzidos por meio dos parâmetros de controle relacionados a abertura e a protrusão labial, a altura da ponta e do dorso da língua, a abertura do véu palatino, dentre outros. Pelo fato do modelo se basear na representação direta das características dinâmicas dos órgãos articulatórios, este modelo de síntese é considerado complexo devido à dificuldade de estimar a configuração exata dos articuladores que produzem um determinado som. No entanto, é possível produzir sons de alta qualidade utilizando este recurso. Além disso, de forma geral, todos os modelos de síntese de fala apresentam os seus benefícios e as suas limitações.

4.2.2 Sintetizador de formantes

O segundo modelo de sintetizador, denominado sintetizador de formantes, busca aproximar diretamente a forma de onda da fala em um conjunto mais simples de regras formuladas no domínio acústico.

O desenho do sintetizador é baseado na teoria acústica da produção da fala desenvolvida por Fant (1960). De acordo com esta teoria, uma ou mais fontes de energia sonora são ativadas pelo aumento da pressão pulmonar. Tratando cada fonte sonora separadamente, é possível caracterizá-las no domínio da frequência por um espectro de fonte. Os sinais gerados pelas fontes sonoras propagam-se através de cavidades com formatos e características variáveis no tempo e são emitidos pelas extremidades oral e nasal para o meio externo.

No modelo de síntese dos formantes, o espectro de um som da fala é considerado como o resultado da combinação linear de três componentes básicos: a) fontes sonoras, b)

característica de filtragem do trato vocal e c) característica de radiação para o meio externo. O espectro $P(f)$ da onda de fala resultante pode ser representado por meio da fórmula:

$$P(f) = S(f) * T(f) * R(f)$$

onde $S(f)$ é o espectro resultante da composição das diversas fontes que excitam o trato vocal; $T(f)$ é a função de transferência linear, composta por pólos e zeros; e $R(f)$ é a característica de radiação, relacionada aos efeitos de propagação dos sons produzidos (KLATT, 1980).

A figura 27 ilustra o processo de produção da fala empregado na síntese de formantes.

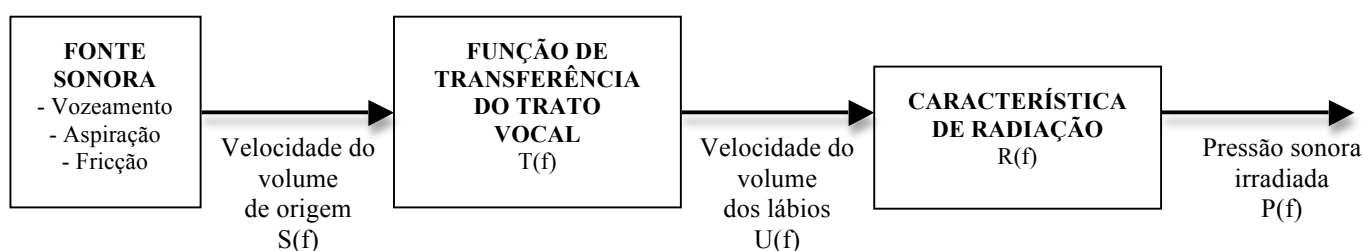


FIGURA 27 - Processo de produção da fala utilizado na síntese de formantes
Fonte: Adaptado de KLATT (1979, p.972)

No sintetizador de formantes proposto por Klatt (1979), existem dois tipos principais de *fonte sonora* que podem ser ativadas durante a produção da fala (Stevens e Klatt, 1972): uma fonte que envolve vibrações quase-periódicas das pregas vocais denominada vozeamento e a outra fonte que relaciona-se à geração de ruído de turbulência caracterizada pela passagem rápida do fluxo de ar em uma constrição estreita. O ruído resultante é denominado aspiração se a constrição está localizada no nível das pregas vocais enquanto o ruído é chamado de fricção se a constrição está localizada acima da laringe como, por exemplo, durante a produção dos sons fricativos. A explosão de uma plosiva também envolve um ruído de fricção (KLATT, 1979).

A estrutura básica do sintetizador de formantes proposto por Klatt (1980) é composta por um conjunto de ressoadores (filtros digitais), que podem ser interligados por dois tipos de configuração: *em série (cascata)* ou *em paralelo*. Estas configurações possuem o objetivo de simular a *função de transferência do trato vocal*, conforme sinalizado na figura 27, e são

ativadas de acordo com a necessidade de reproduzir os sons da fala, ou seja, a seleção dos tipos de configuração do sintetizador pode ser realizada conforme a necessidade do experimentador. Além disso, o sintetizador disponibiliza um anti-ressoador (sistema inverso do ressoador) para síntese dos sons nasais e nasalizados.

4.2.2.1 Ressonadores em cascata

A configuração em cascata caracteriza-se por conter cinco ressoadores de formantes conectados em série e a saída de cada ressoador está ligada a entrada do ressoador seguinte, conforme mostra a figura 28.

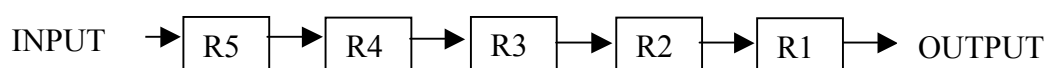


FIGURA 28 - Configuração dos ressoadores em cascata (R5 à R1)
Fonte: Adaptado de KLATT (1979, p.972)

A informação acústica de controle mais importante neste tipo de configuração é a frequência dos formantes, sendo desnecessário inserir a informação referente às amplitudes de cada formante, que são determinadas automaticamente. Esta configuração modela as propriedades ressoantes do trato vocal sempre que a fonte de som está dentro da laringe. Desta forma, a estrutura em cascata normalmente é mais indicada para sintetizar as vogais orais, uma vez que modela corretamente um tubo acústico excitado em uma de suas extremidades, onde as amplitudes de ressonância são impostas pelas frequências e pelas larguras de banda de todos os formantes.

4.2.2.2 Ressonadores em paralelo

Na configuração em paralelo, os ressoadores de formantes que simulam a função de transferência do trato vocal estão conectados em paralelo (figura 29). Cada ressoador de formante é precedido por um controle de amplitude que determina a amplitude relativa do pico espectral (formante) no espectro de saída. O sinal de excitação é aplicado a todos os formantes simultaneamente e suas saídas são somadas.

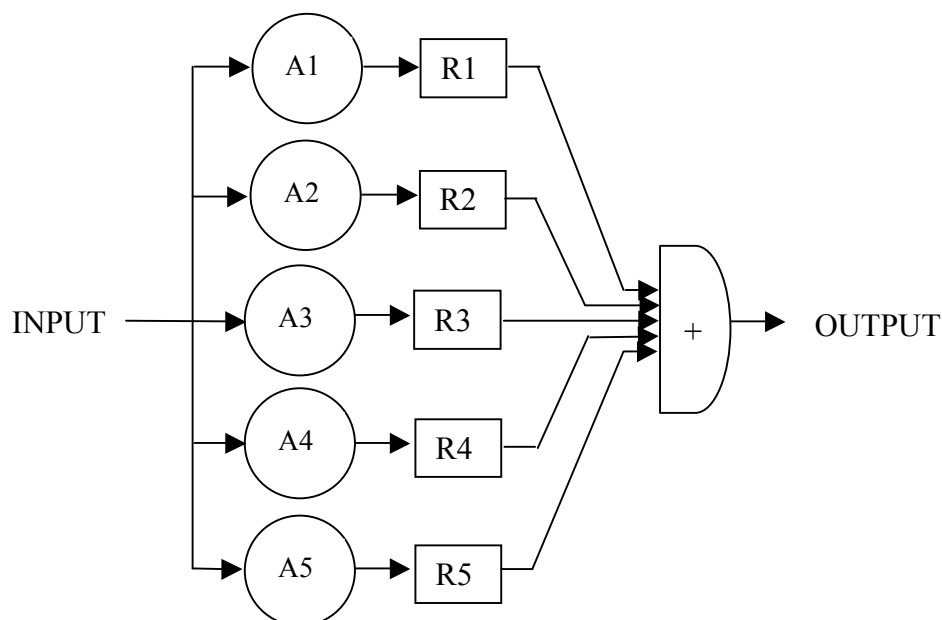


FIGURA 29 - Configuração dos ressoadores em paralelo com as amplitudes dos picos espectrais (formantes)

Fonte: Adaptado de KLATT (1979, p.972)

Normalmente, a configuração em paralelo é mais indicada para a síntese dos sons investigados neste estudo: as fricativas e as plosivas, além das nasais e laterais. Para a síntese das fricativas, que possuem frequências mais altas (por exemplo, [s,z]), o sintetizador disponibiliza seis ressoadores de formantes na configuração em paralelo.

4.3 Parâmetros de controle para síntese dos sons da fala

O sintetizador de formantes proposto por Klatt (1979), implementado no *software* Sensyn, possui dois tipos de parâmetros de controle: 12 parâmetros que permanecem constantes ao longo de todo o enunciado sintetizado e 48 parâmetros variáveis em que os valores podem ser modificados no tempo (ANEXO 2). Os parâmetros variáveis são atualizados a cada 5ms.

Conforme relatado, cinco segmentos foram sintetizados: as fricativas [s] e [ʃ], as plosivas [b], [d] e a vogal [a], nas palavras /sapa/, /ʃapa/, /saba/, /ʃaba/, /bata/ e /data/. Os sons da primeira sílaba de cada palavra foram sintetizados separadamente, uma vez que as pistas auditivas dinâmicas e estáticas do experimento de ponderação de pistas foram manipuladas somente na primeira sílaba de cada palavra. Por exemplo, a fricativa [s] e a vogal [a] da primeira sílaba da palavra /sapa/ foram sintetizadas separadamente, enquanto a segunda sílaba /pa/, da palavra /sapa/, foi sintetizada em conjunto. Feito isso, os sons foram concatenados no Praat. A segunda sílaba das palavras foi sintetizada somente com o intuito de produzir a fala de forma natural, sendo que nenhuma pista auditiva foi manipulada.

Os parâmetros de controle constantes que foram manipulados no *software* Sensyn para todos os estímulos foram: a taxa de amostragem (SR) de 16000Hz; a duração (DU) e a configuração em cascata (CP=0) para a vogal [a] e em paralelo (CP=1) para as fricativas e as plosivas.

Os valores de referência dos demais parâmetros de controle, utilizados para a síntese de todos os estímulos, foram baseados na amostra de fala de um informante do sexo masculino, de 33 anos de idade.

A seguir, será realizada uma descrição detalhada dos parâmetros de controle manipulados para a síntese das fricativas, das plosivas e da vogal [a] no *software* Sensyn.

4.3.1 Síntese da vogal [a] seguinte às fricativas [s] e [ʃ] e às plosivas [b] e [d]

Os parâmetros de controle básicos empregados no *software* Sensyn para a síntese da vogal [a] foram a amplitude de vozeamento (AV), a frequência fundamental (f0) e os três primeiros formantes vocálicos (F1, F2 e F3). O quadro 1 mostra os valores destes parâmetros manipulados ao longo do tempo.

SOUND [a] - Varied Parameters					
time	F0	AV	F1	F2	F3
0	1630	50	415	1522	3110
5	1623	53	443	1509	3070
10	1617	57	471	1496	3030
15	1611	60	499	1484	2990
20	1604	60	527	1471	2951
25	1598	60	555	1459	2911
30	1592	60	584	1446	2871
35	1585	60	612	1433	2832
40	1579	60	640	1421	2792
45	1573	60	668	1408	2752
50	1567	60	696	1396	2713
55	1560	60	696	1396	2713
60	1554	60	696	1396	2713
65	1548	60	696	1396	2713
70	1541	60	696	1396	2713
75	1535	60	696	1396	2713
80	1529	60	696	1396	2713
85	1522	60	696	1396	2713
90	1516	60	696	1396	2713
95	1510	60	696	1396	2713
100	1504	60	696	1396	2713
105	1497	60	696	1396	2713
110	1491	60	696	1396	2713
115	1485	60	696	1396	2713
120	1478	60	696	1396	2713
125	1472	56	696	1396	2713
130	1466	53	696	1396	2713
135	1460	50	696	1396	2713

QUADRO 1 - Parâmetros manipulados ao longo do tempo no *software* Sensyn. Vogal [a] que seguiu a fricativa [ʃ]

Na primeira coluna do quadro 1 visualizamos a duração da vogal que foi de 140ms. A seguir podemos observar os valores da frequência fundamental (f0) inseridos no *software*. A f0 inicial foi de 161Hz, ou seja, 161 ciclos de abertura e fechamento das pregas vocais por segundo. Este parâmetro foi ajustado para decrescer ao longo da emissão da vogal [a], conforme observado na produção do informante representado pela linha azul da figura adiante.

A amplitude de vozeamento (AV) relaciona-se a amplitude, em dB, da fonte de vozeamento, ou seja, da ativação das pregas vocais. Desta forma, a fonte de vozeamento é desativada quando inserimos o valor de 0dB. Tal procedimento deve ser realizado diante dos sons não vozeados, onde não há vibração das pregas vocais. No caso das vogais, sons vozeados, a ativação desse parâmetro torna-se necessária. A amplitude de vozeamento apresentou variações em seus valores, uma vez que este parâmetro não é constante. Na figura 30 observa-se, com o auxílio do oscilograma, um aumento da amplitude durante a emissão da

vogal [a], sua estabilização e posterior redução nos seus valores ao final da produção. Esta variação nos valores da AV foi inserida no *software* Sensyn.

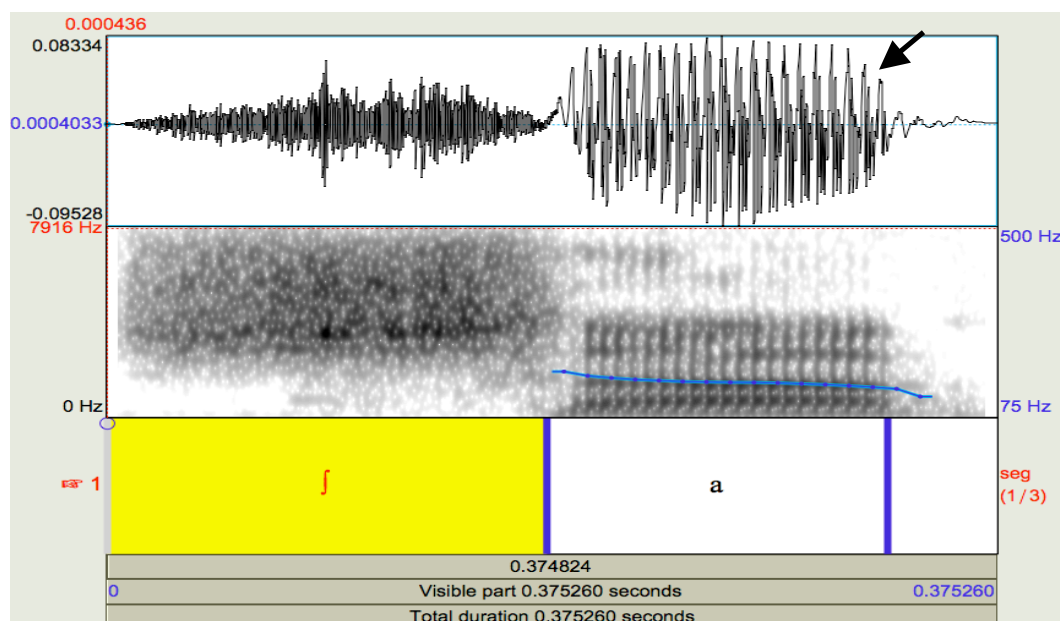


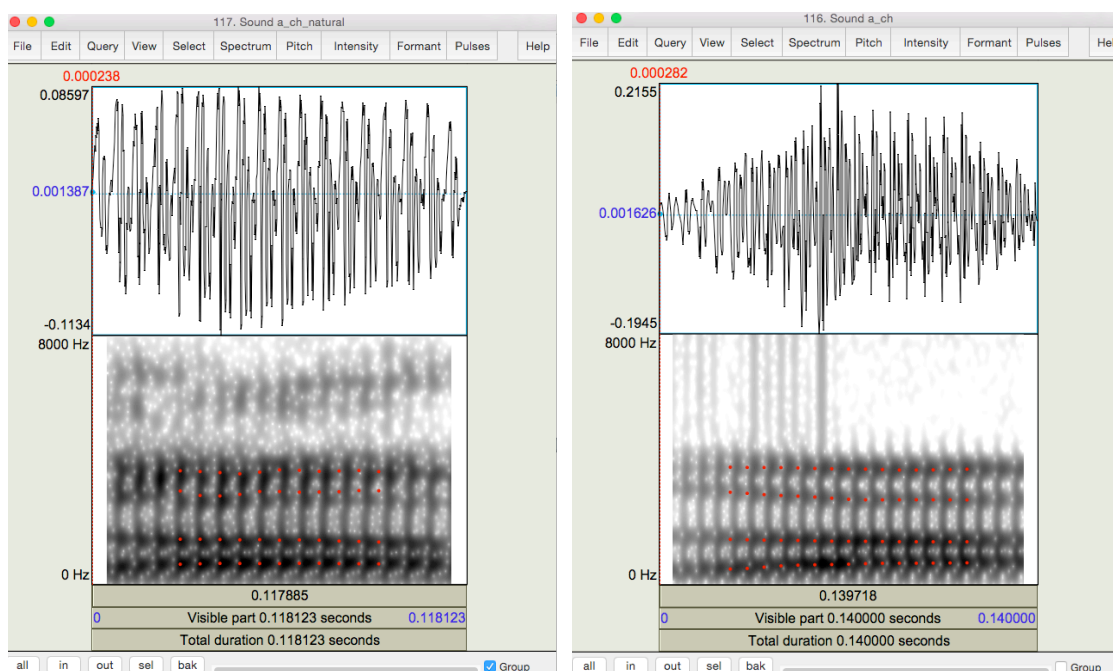
FIGURA 30 - Frequência fundamental da vogal [a] que seguiu a fricativa [ʃ], na palavra /ʃapa/, sinalizada pela linha azul. No oscilograma observa-se o comportamento da amplitude de vozeamento (seta preta)

Por fim, os valores das frequências dos três formantes vocálicos foram manipulados. Este parâmetro pode ser considerado o mais importante para a síntese das vogais, uma vez que a variação dos valores dos formantes na transição da consoante para a vogal constitui uma pista auditiva relevante para que os ouvintes façam a distinção entre os sons das línguas.

Conforme descrito, os valores dos três formantes (F1, F2 e F3) foram obtidos após a análise dos parâmetros acústicos do informante selecionado. Desta forma, o tempo 0ms do quadro 1, representa os valores dos formantes na porção inicial da vogal [a] que seguiu as fricativas [ʃ] e [s] e as plosivas [b] e [d]. A duração da transição dos formantes foi de 50ms, ou seja, os valores dos formantes estabilizaram dentro de 50ms.

As frequências do quarto e do quinto ressonador de formantes, bem como as larguras de banda de todos os formantes foi mantida constante, com pouco decréscimo da qualidade do som de saída, devido a pequena variação em seus valores. Ressalte-se que estes ressonadores contribuem pouco para inteligibilidade das vogais.

As figuras abaixo mostram o espectrograma da vogal [a] seguinte a fricativa [ʃ] do som natural (à esquerda) e do som sintetizado (à direita).



FIGURAS 31 e 32 - Espectrograma da vogal [a] seguinte a fricativa [ʃ]. À esquerda observa-se o espectrograma do som original (vogal natural) e à direita o espectrograma da vogal sintetizada

Basicamente, foram os valores dos três primeiros formantes (F1, F2 e F3) que diferiram entre a vogal [a] seguintes às fricativas e às plosivas. Estes valores serão descritos no capítulo 3.

4.3.2 Síntese das fricativas [ʃ] e [s]

Os parâmetros de controle básicos para criação das fricativas sintetizadas [ʃ] e [s] foram a amplitude de fricção (AF), os seis formantes (F1-F6) e as amplitudes associadas aos ressoadores em paralelo, aos quais é enviado o ruído de fricção (AB, A2F-A6F).

O quadro 2 mostra os parâmetros manipulados ao longo do tempo no *software* Sensyn para a geração da fricativa [s] nas palavras /sapa/ e /saba/. A amplitude de vozeamento (AV) foi desativada durante toda a emissão da fricativa [s] por ser tratar de um som não vozeado. A duração das fricativas sintetizadas foi de 160ms.

SOUND [s] - Varied Parameters														
time	AV	AF	F1	F2	F3	F4	F5	F6	A2F	A3F	A4F	A5F	A6F	AB
0	0	50	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
5	0	51	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
10	0	53	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
15	0	55	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
20	0	57	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
25	0	59	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
30	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
35	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
40	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
45	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
50	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
55	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
60	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
65	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
70	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
75	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
80	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
85	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
90	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
95	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
100	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
105	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
110	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
115	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
120	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
125	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
130	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
135	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
140	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
145	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
150	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30
155	0	60	1871	3385	4500	5058	6201	6868	0	0	45	42	0	30

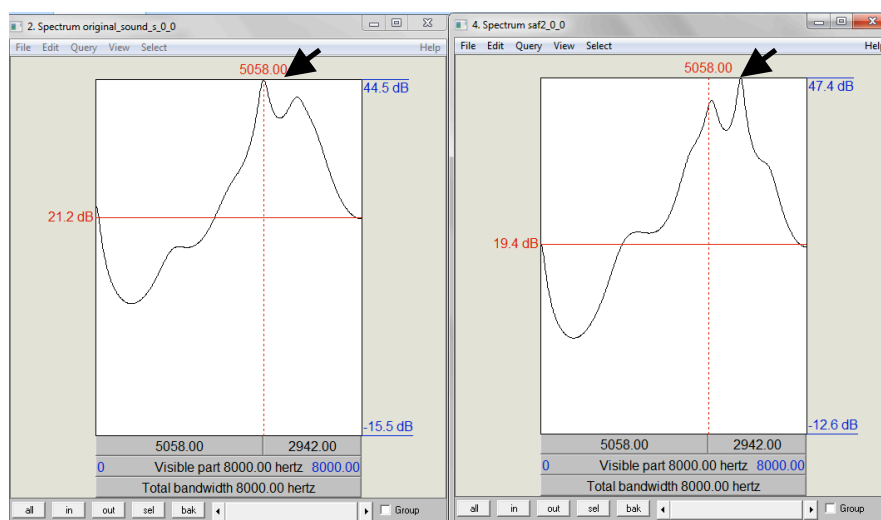
QUADRO 2 - Parâmetros manipulados ao longo do tempo no *software* Sensyn. Fricativa [s], das palavras /sapa/ e /saba/

A amplitude de fricção (AF) determina o nível de ruído fricativo que será enviado aos ressoadores em paralelo (A2F-A6F). No caso das fricativas, este parâmetro deve ser ajustado gradualmente até atingir 60dB. O quadro 2 mostra que a AF atingiu o valor de 60dB em 30ms.

O parâmetro mais importante para a síntese das fricativas é a amplitude de fricção dos formantes em paralelo (A2F-A6F). Este parâmetro determina a forma espectral destes sons e as regiões do espectro que serão excitadas. A principal diferença entre as fricativas [ʃ] e [s] concentra-se nos valores destas amplitudes, sendo que a fricativa [s] apresenta ruído predominante em frequências mais altas quando comparada com a fricativa [ʃ].

Os valores das amplitudes de fricção da fricativa [s] foram ajustados da seguinte forma: A2F, A3F e A6F em 0dB, A4F em 45dB e A5F em 42dB. Ressalte-se que os valores das

amplitudes A2F, A3F e A6F foram inseridos como zero, uma vez que a fatia do espectro desta fricativa revelou dois picos de amplitude mais proeminentes (A4F e A5F), como mostram as setas nas figuras 33 e 34.



FIGURAS 33 e 34 - Fatia do espectro (do inglês, *Spectral slice*) da fricativa [s]. À esquerda observa-se a fatia do espectro do som original (fricativa natural) e à direita a fatia do espectro da fricativa sintetizada

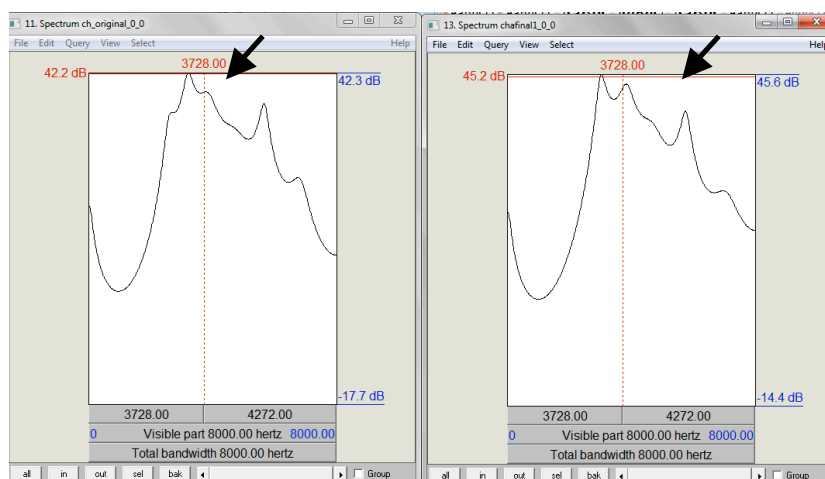
Conforme descrito no capítulo 1, a definição exata dos valores dos formantes e de suas amplitudes não é muito importante para a síntese dos sons fricativos. Desta forma, faz-se um levantamento geral das propriedades acústicas dos sons originais para que posteriormente seja criado o som sintetizado. Neste caso, observou-se a tendência geral do comportamento das amplitudes de fricção para confecção dos estímulos fricativos.

Os valores das amplitudes da fricativa [ʃ] foram ajustados da seguinte maneira: A2F em 41dB, A3F em 39dB, A4F em 32dB, A5F em 36dB e A6F em 23dB. O quadro 3 mostra os valores inseridos no *software* Sensyn para a criação da fricativa [ʃ].

SOUND [ʃ] - Varied Parameters														
time	AV	AF	F1	F2	F3	F4	F5	F6	A2F	A3F	A4F	A5F	A6F	AB
0	0	50	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
5	0	51	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
10	0	53	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
15	0	55	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
20	0	57	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
25	0	59	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
30	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
35	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
40	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
45	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
50	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
55	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
60	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
65	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
70	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
75	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
80	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
85	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
90	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
95	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
100	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
105	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
110	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
115	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
120	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
125	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
130	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
135	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
140	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
145	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
150	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40
155	0	60	1871	3396	3728	4796	5656	6648	41	39	32	36	23	40

QUADRO 3 - Parâmetros manipulados ao longo do tempo no *software* Sensyn. Fricativa [ʃ], das palavras /ʃapa/ e /ʃaba/

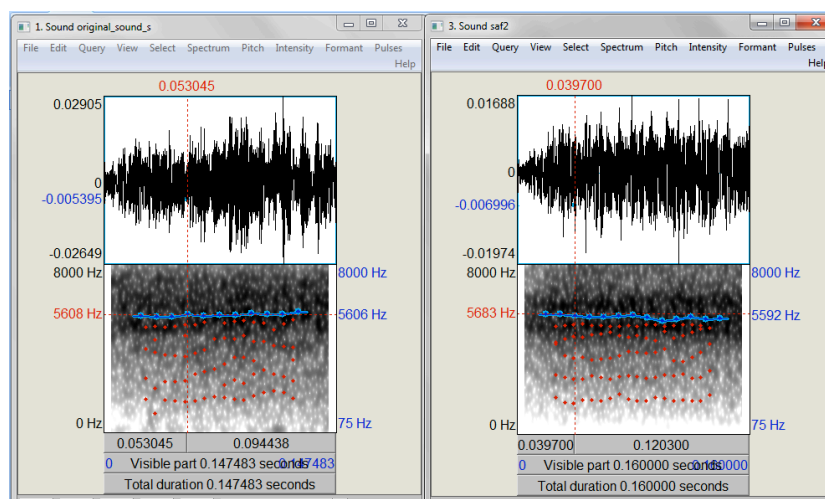
As figuras 35 e 36 revelam a fatia do espectro da fricativa [ʃ]. Neste caso, observa-se que os picos de amplitude (A2F e A3F) deste som estão concentrados em regiões de frequência mais baixas quando comparado com a fricativa [s] e, além disso, não observa-se somente dois picos de amplitudes bem definidos.



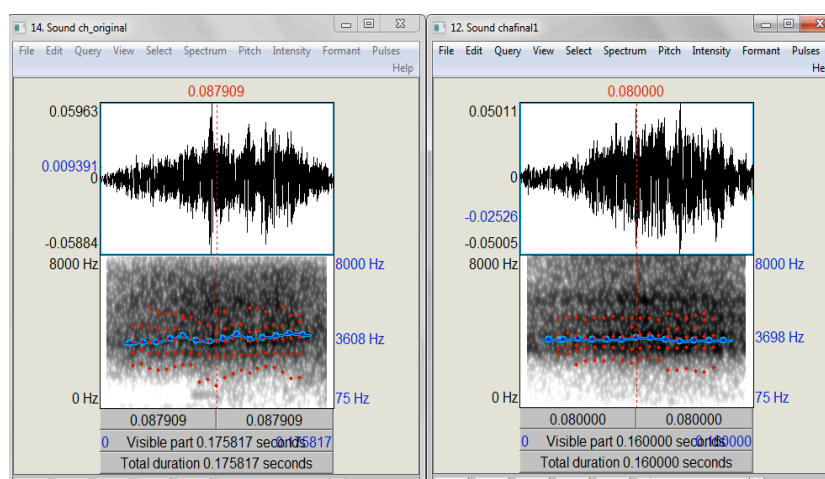
FIGURAS 35 e 36 - Fatia do espectro (do inglês, *Spectral slice*) da fricativa [ʃ]. À esquerda observa-se a fatia do espectro do som original (fricativa natural) e à direita a fatia do espectro da fricativa sintetizada

Em suma, a fricativa [s] apresentou os picos das amplitudes A4F e A5F com valores maiores e bem definidos em regiões de frequências mais altas, enquanto a fricativa [ʃ] apresentou os picos das amplitudes A2F e A3F em regiões de frequências mais baixas. No entanto, a fricativa [ʃ] apresentou picos de amplitudes (A2F-A6F) mais espalhados ao longo do espectro.

Após a inspeção das frequências dos formantes e de suas amplitudes observou-se o centro de gravidade das fricativas verificando este parâmetro no espectrograma do som natural e sintetizado. As figuras 37 e 38 mostram o espectrograma das fricativas [s] e [ʃ] do som natural (à esquerda) e do som sintetizado (à direita).



FIGURAS 37 e 38 - Espectrograma da fricativa [s]. À esquerda observa-se o espectrograma do som original (fricativa natural) e à direita o espectrograma da fricativa sintetizada



FIGURAS 39 e 40 - Espectrograma da fricativa [ʃ]. À esquerda observa-se o espectrograma do som original (fricativa natural) e à direita o espectrograma da fricativa sintetizada

As figuras 39 e 40 mostram que as fricativas naturais [s] e [ʃ] ficaram com os valores do centro de gravidade (CG) bem próximos quando comparadas com as fricativas sintetizadas. Conforme esperado, a fricativa [ʃ] apresentou faixas de concentração de energia em regiões mais baixas quando comparado com a fricativa [s]. O valor do CG da fricativa [s] foi 5633Hz e o da fricativa [ʃ] foi 3830Hz.

4.3.3 Síntese das plosivas [b] e [d]

Os parâmetros de controle básicos para criação das plosivas sintetizadas [b] e [d] foram a amplitude de fricção (AF), os cinco formantes (F1-F5) e as amplitudes associadas aos ressoadores em paralelo, aos quais é enviado o ruído de fricção (AB, A2F-A5F). As plosivas tiveram uma duração total de 100ms, sendo que 85ms representaram a fase de pré-sonorização e os demais 15ms o *burst*. Como estamos lidando com dois sons vozeados, a amplitude de vozeamento (AV) foi ativada automaticamente durante toda a produção das plosivas, diferente das fricativas não vozeadas, nas quais este parâmetro apresentou um valor igual a zero.

O quadro 4 mostra os parâmetros manipulados ao longo do tempo no *software* Sensyn para a geração da plosiva [b] na palavra /bata/.

SOUND [b] - Varied Parameters												
time	TL	AF	F1	F2	F3	F4	F5	A2F	A3F	A4F	A5F	AB
0	30	0	200	900	2000	3500	3850	60	50	20	10	0
5	30	0	200	900	2000	3500	3850	60	50	20	10	0
10	31	0	200	900	2000	3500	3850	60	50	20	10	0
15	32	0	200	900	2000	3500	3850	60	50	20	10	0
20	32	0	200	900	2000	3500	3850	60	50	20	10	0
25	33	0	200	900	2000	3500	3850	60	50	20	10	0
30	34	0	200	900	2000	3500	3850	60	50	20	10	0
35	35	0	200	900	2000	3500	3850	60	50	20	10	0
40	35	0	200	900	2000	3500	3850	60	50	20	10	0
45	36	0	200	900	2000	3500	3850	60	50	20	10	0
50	37	0	200	900	2000	3500	3850	60	50	20	10	0
55	37	0	200	900	2000	3500	3850	60	50	20	10	0
60	38	0	200	900	2000	3500	3850	60	50	20	10	0
65	39	0	200	900	2000	3500	3850	60	50	20	10	0
70	40	0	200	900	2000	3500	3850	60	50	20	10	0
75	40	0	200	900	2000	3500	3850	60	50	20	10	0
80	40	0	200	900	2000	3500	3850	60	50	20	10	0
85	0	60	664	1386	1998	3290	4385	60	50	20	10	60
90	0	40	664	1386	1998	3290	4385	60	50	20	10	40
95	0	20	664	1386	1998	3290	4385	60	50	20	10	20

QUADRO 4 - Parâmetros manipulados ao longo do tempo no *software* Sensyn. Plosiva [b], da palavra /bata/

Conforme relatado, a amplitude de fricção (AF) determina o nível de ruído que será enviado aos ressoadores em paralelo (A2F-A5F). Este parâmetro é considerado importante para a produção do *burst* e foi manipulado de maneira diferente das fricativas. No caso das plosivas, a AF foi igual a 0dB na fase de pré-sonorização e variou de forma abrupta de 60dB para 20dB, nos 15ms da fase de soltura. Esta variação abrupta nos valores da AF é fundamental para a produção do *burst*.

Outro parâmetro fundamental para a produção do *burst* é a amplitude AB. Este parâmetro é utilizado quando os efeitos de ressonância do trato vocal são insignificantes, uma vez que a cavidade anterior a constrição é muito curta. Tal fato ocorre com as plosivas bilabiais e alveolares que são produzidas pela constrição dos lábios e da região alveolar (dentes superiores e ponta da língua), diferente da plosiva [g], que possui uma cavidade anterior a constrição realizada na região do véu palatino maior. A amplitude AB também foi manipulada de forma abrupta de 60dB para 20dB, nos 15ms do *burst*.

O parâmetro mais importante para a síntese das plosivas [b] e [d] e, principalmente, para a reprodução do *burst* é a amplitude de fricção dos formantes em paralelo (A2F-A5F). A principal diferença entre as plosivas [b] e [d] concentra-se nos valores destas amplitudes,

sendo que a plosiva [b] apresenta excitação dos ressoadores em paralelo em regiões mais baixas de frequência quando comparada com a plosiva [d].

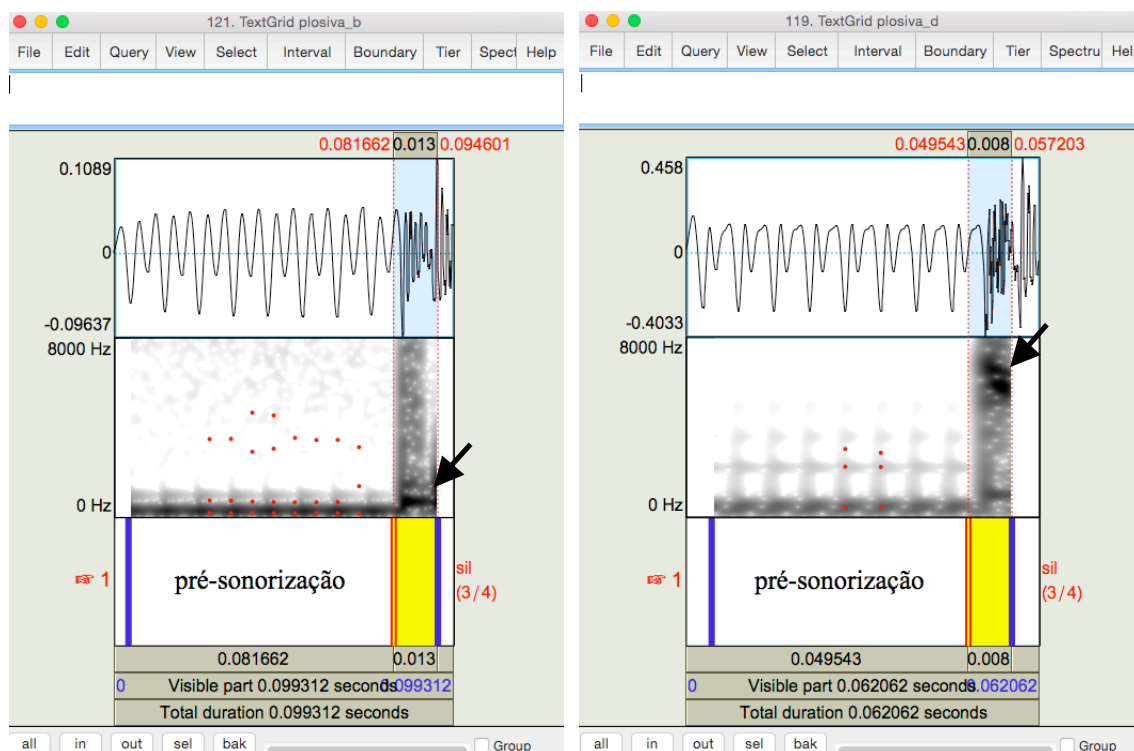
Os valores das amplitudes de fricção da plosiva [b] foram ajustados da seguinte forma: A2F em 60dB, A3F em 50dB, A4F em 20dB e A5F em 10dB. Para produzir picos de energia mais intensos em regiões mais baixas de frequência, os valores da A2F e da A3F foram maiores. No caso da plosiva [d], os valores das amplitudes foram ajustados da seguinte maneira: A2F em 0dB, A3F em 10dB, A4F em 50dB e A5F em 60dB. Os valores maiores da A4F e A5F produzem picos de energia mais intensos em regiões mais altas de frequência. O quadro adiante mostra os valores inseridos no *software* Sensyn para a criação da plosiva [d].

SOUND [d] - Varied Parameters												
time	TL	AF	F1	F2	F3	F4	F5	A2F	A3F	A4F	A5F	AB
0	30	0	292	1389	2245	3074	3670	0	10	50	60	0
5	30	0	292	1389	2245	3074	3670	0	10	50	60	0
10	31	0	292	1389	2245	3074	3670	0	10	50	60	0
15	32	0	292	1389	2245	3074	3670	0	10	50	60	0
20	32	0	292	1389	2245	3074	3670	0	10	50	60	0
25	33	0	292	1389	2245	3074	3670	0	10	50	60	0
30	34	0	292	1389	2245	3074	3670	0	10	50	60	0
35	35	0	292	1389	2245	3074	3670	0	10	50	60	0
40	35	0	292	1389	2245	3074	3670	0	10	50	60	0
45	36	0	292	1389	2245	3074	3670	0	10	50	60	0
50	37	0	292	1389	2245	3074	3670	0	10	50	60	0
55	37	0	292	1389	2245	3074	3670	0	10	50	60	0
60	38	0	292	1389	2245	3074	3670	0	10	50	60	0
65	39	0	292	1389	2245	3074	3670	0	10	50	60	0
70	40	0	292	1389	2245	3074	3670	0	10	50	60	0
75	40	0	292	1389	2245	3074	3670	0	10	50	60	0
80	40	0	292	1389	2245	3074	3670	0	10	50	60	0
85	0	60	611	1201	2360	3537	3984	0	10	50	60	60
90	0	40	611	1201	2360	3537	3984	0	10	50	60	40
95	0	20	611	1201	2360	3537	3984	0	10	50	60	20

QUADRO 5 - Parâmetros manipulados ao longo do tempo no *software* Sensyn. Plosiva [d], da palavra /data/

Finalmente, manipulou-se o parâmetro TL, na fase de pré-sonorização de ambas as plosivas. Este parâmetro produz um amortecimento das frequências mais altas na fase de pré-sonorização, o que auxilia na reprodução da barra de vozeamento.

As figuras 41 e 42 mostram o espectrograma das plosivas [b] e [d] sintetizadas. À esquerda observa-se a plosiva [b] e à direita a plosiva [d], divididas em duas fases: a fase de pré-sonorização e a fase de soltura, *burst*, indicada pela seta.



FIGURAS 41 e 42 - À esquerda observa-se o espectrograma da plosiva [b] sintetizada e à direita observa-se o espectrograma da plosiva [d] sintetizada

As setas também sinalizam que o *burst* da plosiva bilabial [b] apresentou uma maior concentração de energia em regiões de frequências mais baixas quando comparado com o *burst* da plosiva alveolar [d], conforme esperado.

Após a realização da descrição dos parâmetros acústicos e da síntese das fricativas [ʃ] e [s] e das plosivas [b] e [d], desenvolveu-se um estudo piloto com o objetivo de analisar a qualidade dos estímulos sintetizados, que será descrito no próximo capítulo.

5 CAPÍTULO 3 – ESTUDO PILOTO COMPOSTO PELA TAREFA DE CLASSIFICAÇÃO DOS ESTÍMULOS SINTETIZADOS /japa-/sapa/, /jaba-/saba/ E /bata-/data/

Este capítulo foi desenvolvido com o objetivo de analisar, por meio de uma tarefa de classificação, a qualidade dos estímulos sintetizados compostos pelas palavras /japa-/sapa/, /jaba-/saba/ e /bata-/data/. Conforme abordado na seção de revisão de literatura, essa tarefa induz a classificação fonêmica. Desta forma, espera-se que os estímulos sejam classificados de forma categórica e, além disso, espera-se obter curvas de classificação íngremes para os adultos que participaram do estudo piloto. Tal fato indica que as respostas foram consistentes e que os estímulos apresentam boa qualidade.

A seguir, encontram-se descritos a metodologia e os resultados obtidos com este estudo piloto. Este estudo também foi realizado com a aprovação do Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, Parecer CAAE – 62617616.0.0000.5149.

5.1 Metodologia do estudo piloto composto pela tarefa de classificação dos estímulos sintetizados

5.1.1 Participantes do estudo

Os participantes deste estudo foram oito adultos, do sexo feminino, falantes nativos do português brasileiro, com faixa etária entre 20 e 40 anos de idade. Como critério de seleção utilizou-se a condição de que os participantes não poderiam apresentar histórico significativo de otite média, de colocação de tubo de ventilação e de perda auditiva. Todos os participantes passaram previamente por uma triagem auditiva, em ambas as orelhas, que consistiu em tons puros, nas frequências de 0.5, 1, 2, 4 e 8kHz, apresentados em 25dB. Para a realização da triagem, foi utilizado o audiômetro portátil, modelo AD 28.

Os participantes assinaram o Termo de Consentimento Livre e Esclarecido, após serem informados sobre os aspectos gerais da tarefa de classificação (ANEXO 3).

5.1.2 Estímulos

A produção dos estímulos acompanhou a produção dos estímulos realizada por Nittrouer e Miller (1996). Desta forma, realizou-se a criação de dois continua de sons da fala onde duas pistas auditivas foram manipuladas. Uma das duas pistas foi manipulada do mesmo modo ao longo de ambos os continua. Assim, os dois continua foram idênticos em relação a esta pista. Já a outra pista foi manipulada por meio dos dois continua. Portanto, os dois continua diferiram em relação a esta pista.

No caso das fricativas, as duas pistas auditivas manipuladas foram: a altura da frequência do ruído fricativo e a transição do F2 e do F3 no início da vogal. A pista que variou de forma idêntica nos dois continua do par mínimo e do par de pseudopalavras foi a altura da frequência do ruído fricativo, modificada de uma frequência apropriada para [s] (5633Hz) para uma frequência apropriada para [ʃ] (3830Hz). A segunda pista foi a transição do F2 e do F3 no início da vogal e as duas formas foram: um continuum com transições vocálicas apropriadas para acompanhar [s] e outro continuum com transições apropriadas para acompanhar [ʃ].

Já no caso das plosivas, as duas pistas auditivas manipuladas foram: as amplitudes do *burst* e a transição do F2 e do F3 no início da vogal. Diferente das fricativas, a pista que variou de forma idêntica nos dois continuas do par mínimo foi a transição do F2 e do F3, com valores dos formantes apropriados para [b] e valores dos formantes apropriados para [d]. A segunda pista foi a amplitude do *burst* e as duas formas foram: um continuum com amplitudes do *burst* apropriadas para acompanhar [b] e outro continuum com amplitudes do *burst* apropriadas para acompanhar [d].

Este tipo de desenho permite uma investigação do efeito perceptual das duas pistas. Um ouvinte que não é influenciado pela pista que muda entre os continua – ou seja, a pista de transição dos formantes para as fricativas e a pista de amplitudes do *burst* para as plosivas – perceberá os dois continua como sendo o mesmo. Ao contrário, um ouvinte que é influenciado pela pista que muda entre os continua perceberá os dois continua diferentemente.

Conforme descrito, todos os estímulos foram sintetizados por meio do *software* Sensyn, seguindo o padrão de síntese Klatt e Klatt (1990).

5.1.2.1 Continua compostos pelas palavras /japa/-/sapa/ e /jaba/-/saba/

5.1.2.1.1 Ruído fricativo

Para a produção do continuum, foram gerados dois sons de ruído: um som apresentou uma frequência apropriada para [s] (5633Hz) e o outro som apresentou uma frequência apropriada para [ʃ] (3830Hz). Os outros ruídos do continuum (oito) foram gerados por meio do método de interpolação entre as amplitudes relativas dos envelopes espectrais desses dois sons de ruído. O *script* utilizado para geração dos demais ruídos pode ser visualizado no anexo 4.

A figura abaixo mostra a fatia do espectro dos dez ruídos gerados pelo método de interpolação, com os seus respectivos picos de amplitude.

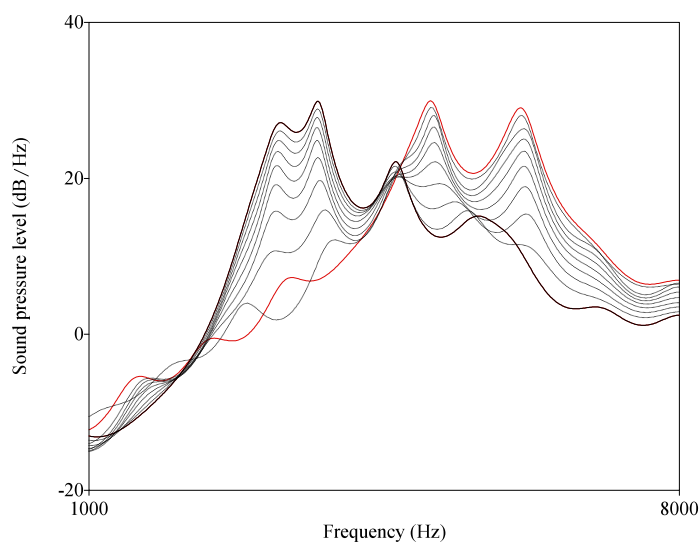


FIGURA 43 - Envolvórios espectrais obtidos em 50 ms em torno do ponto central do primeiro (em preto, linha contínua) ao décimo (em vermelho) ruídos fricativos

O primeiro e o décimo ruído constituíram as extremidades do continuum fricativo. Pela análise da figura 43 observa-se que o ruído 1 (linha preta contínua) apresenta os picos de amplitude em regiões de frequência mais baixas quando comparado com o ruído 10 (linha

vermelha). Desta forma, o ruído 1 representou a fricativa [ʃ], enquanto o ruído 10 representou a fricativa [s].

A duração dos ruídos fricativos foi de 160ms e a intensidade foi ajustada em 55dB com base na análise dos estímulos naturais de fala.

5.1.2.1.2 Porção vocálica seguinte às fricativas

Conforme relatado, a pista auditiva manipulada no caso das vogais foi a transição do segundo e do terceiro formantes (F2 e F3). Para a construção do continuum, criou-se duas vogais [a]: uma vogal apresentou os valores de transição dos formantes apropriados para acompanhar [ʃ] e a outra vogal apresentou os valores de transição dos formantes apropriados para acompanhar [s].

Os valores inicial e estável das frequências dos formantes vocálicos (F1, F2 e F3) podem ser visualizados na tabela 11.

TABELA 11
Valores das frequências dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] em (Hz), informante 1

Valores das frequências dos formantes da vogal [a] sintetizada seguinte às fricativas - Informante 1						
Vogal	F1 inicial	F1 estável	F2 inicial	F2 estável	F3 inicial	F3 estável
[a]_/sa/	491	711	1489	1412	3166	2895
[a]_/fa/	415	696	1522	1396	3110	2713

Cabe ressaltar, que os valores da tabela acima foram extraídos da análise de fala de um informante, falante nativo do PB, do sexo masculino, que não participou do estudo piloto descrito no capítulo 3 (tabela 8). A análise de fala, bem como a criação dos estímulos sintetizados (fricativas e vogais) deste informante foi feita antes da coleta de fala dos participantes do estudo piloto, no momento em que a pesquisadora realizou um estágio sanduíche na Universidade de Edimburgo/Escócia. Desta forma, além do informante 1, selecionou-se um dos informantes que participaram do estudo piloto descrito no capítulo 3 (informante 5, tabela 8), que apresentou os valores dos formantes vocálicos mais próximos da média, como mostra a tabela adiante. No entanto, o informante 1 não apresentou valores dos formantes tão distantes dos valores médios.

TABELA 12

Valores das frequências dos formantes da vogal [a] seguinte às fricativas [s] e [ʃ] em (Hz), informante 2

Valores das frequências dos formantes da vogal [a] sintetizada seguinte às fricativas						
Vogal	F1 inicial	F1 estável	F2 inicial	F2 estável	F3 inicial	F3 estável
[a]_/sa/	616	768	1389	1366	2863	2584
[a]_/ja/	613	819	1579	1369	2878	2444

A duração da vogal foi de 140ms e a intensidade foi ajustada em 70dB. Os valores dos formantes estabilizaram em 50ms. A f0 do informante 1 iniciou em 163Hz e terminou em 146Hz ao longo da vogal. Já a f0 do informante 2 apresentou o valor inicial de 124Hz e o valor final de 112Hz.

Após a análise das respostas da tarefa de classificação serão utilizados os estímulos sintetizados que apresentarem a melhor qualidade para o experimento final de ponderação de pistas auditivas.

A seguir será descrito a forma de criação dos continua compostos pelas palavras /bata/-/data/.

5.1.2.2 Continua compostos pelas palavras /bata/-/data/

Os continua compostos pelo par mínimo /bata/-/data/ foram criados de maneira diferente dos continua compostos pelas fricativas e pela vogal [a]. Aqui, criou-se duas amplitudes do *burst* e nove vogais que diferiram em relação aos valores de transição do segundo e do terceiro formantes (F2 e F3).

5.1.2.2.1 Plosivas [b] e [d]

Para a produção dos continua, foram produzidas duas amplitudes do *burst*: uma amplitude com picos de energia mais intensos em regiões mais baixas de frequência (compatível com [b]) e outra amplitude com picos de energia mais intensos em regiões mais altas de frequência (compatível com [d]), descritos de forma mais detalhada no capítulo 4, seção 4.3.3.

A duração total dos estímulos plosivos foi de 100ms, sendo 85ms de fase de pré-sonorização e 15ms de *burst*. A intensidade das plosivas foi ajustada em 55dB.

5.1.2.2.2 Porção vocálica seguinte às plosivas

Para a construção dos continua, compostos pelas plosivas, foram geradas nove vogais sintetizadas. A tabela 13 mostra os valores inicial e estável dos formantes, F2 e F3, do continuum vocálico [a] seguinte às plosivas [b] e [d].

TABELA 13

Valores das frequências dos formantes do continuum vocálico [a] seguinte às plosivas [b] e [d] em (Hz)

Valores das frequências dos formantes da vogal [a] sintetizada seguinte às plosivas				
Vogal	F2 inicial	F2 estável	F3 inicial	F3 estável
1	1111	1400	2352	2460
2	1199	1400	2386	2460
3	1287	1400	2420	2460
4	1375	1400	2454	2460
5	1463	1400	2488	2460
6	1551	1400	2522	2460
7	1639	1400	2556	2460
8	1727	1400	2590	2460
9	1815	1400	2624	2460

O primeiro formante (F1) apresentou valores iguais para todas as vogais e iniciou em uma frequência de 503Hz, alcançando o valor de 710Hz em 50ms.

Os valores estáveis do F2 e do F3, do continuum vocálico, também foram os mesmos para todas as vogais, 1400Hz e 2460Hz, respectivamente. Estes valores foram obtidos pela média dos valores estáveis do F2 e do F3, descritos no capítulo 3, seção 3.5.3.2.

Os valores iniciais do F2 modificaram em passos uniformes de 88Hz de uma frequência apropriada para acompanhar [b] (vogal 1) para uma frequência apropriada para acompanhar [d] (vogal 9). Da mesma forma do segundo formante, os valores iniciais do F3 alteraram em passos uniformes de 34Hz, de uma frequência apropriada para acompanhar [b] (vogal 1) para uma frequência apropriada para acompanhar [d] (vogal 9).

A duração das vogais [a] foi de 140ms e a intensidade foi ajustada em 70dB. Os valores dos formantes estabilizaram em 50ms. A f_0 iniciou em 132Hz e terminou em 128Hz ao longo da vogal.

5.1.3 Continua

Os ruídos fricativos, as amplitudes do *burst* das plosivas e as vogais sintetizadas foram concatenados no *software* PRAAT (versão 5.2.35) para formação dos continua. A seguir, descrevem-se os seis continua gerados no estudo piloto: dois continua para o par mínimo /*japa*/-/*sapa*/, dois continua para o par de pseudopalavras /*jaba*/-/*saba*/ e dois continua para o par mínimo /*bata*/-/*data*/.

O primeiro continuum do par mínimo /*japa*/-/*sapa*/ apresentou a seguinte estrutura: 10 ruídos fricativos + vogal [a] com transição vocálica apropriada para acompanhar [j] + sílaba final /*pa*/. A figura 44 representa o primeiro estímulo do continuum 1. Nesta figura, observa-se o estímulo de melhor exemplar da categoria [j].

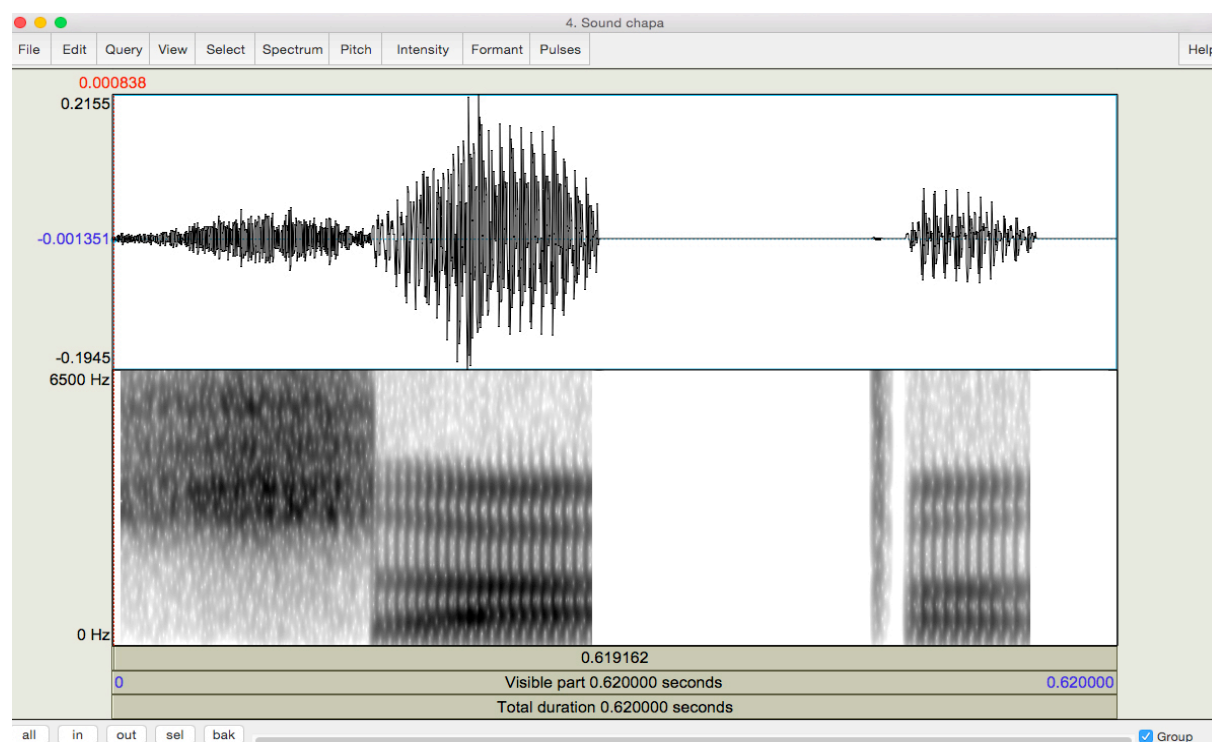


FIGURA 44 – Espectrograma do melhor exemplar da palavra /*japa*/ sintetizada

O segundo continuum foi composto por: 10 ruídos fricativos + vogal [a] com transição vocálica apropriada para acompanhar [s] + sílaba final /pa/. A figura 45 representa o décimo estímulo do continuum 2, considerado o melhor exemplar da categoria [s].

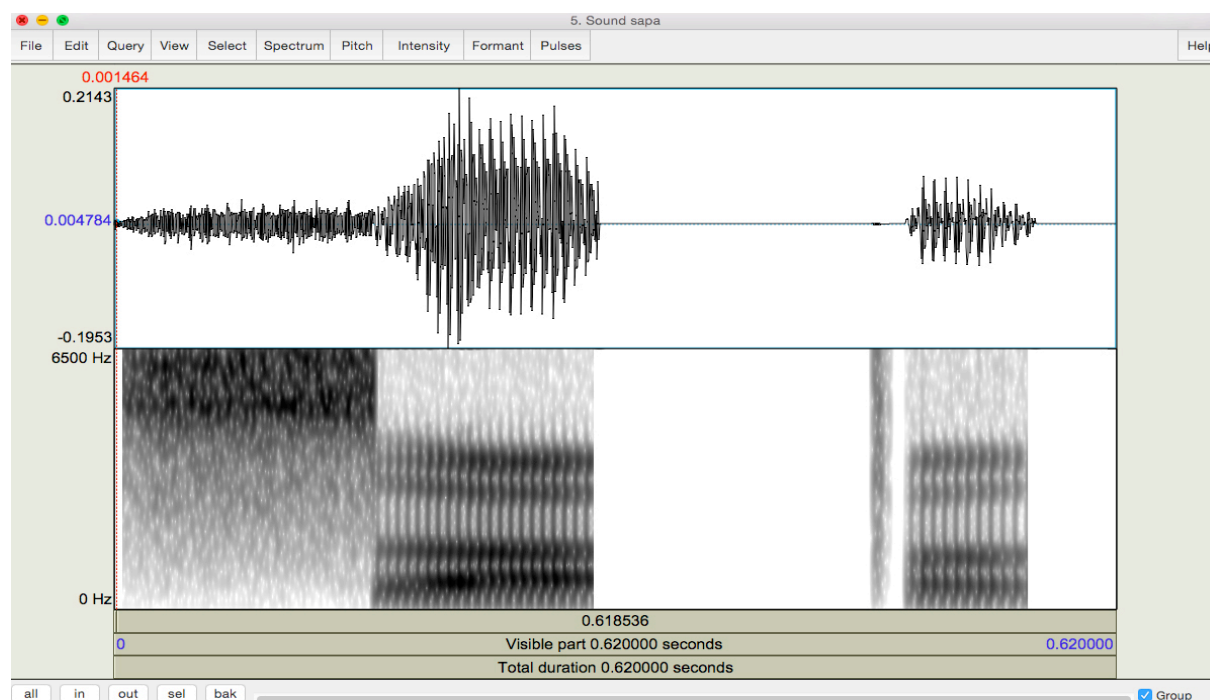


FIGURA 45 - Espectrograma do melhor exemplar da palavra /sapa/ sintetizada

Tanto os estímulos do continuum 1 quanto os do continuum 2 apresentaram duração total de 570ms e cada continuum continha dez estímulos.

Do mesmo modo que no par mínimo /ʃapa-/sapa/, gerou-se dois continua para o par de pseudopalavras /ʃaba-/saba/. Porém, houve diferença na duração da segunda sílaba das palavras. Enquanto a sílaba /pa/ apresentou duração de 270ms, a sílaba /ba/ apresentou duração de 100ms. Além disso, a duração total das pseudopalavras /ʃaba/ e /saba/ do continuum foi de 400ms.

Por fim, o primeiro continuum do par mínimo /bata-/data/ apresentou a seguinte composição: amplitudes do *burst* com picos de energia mais intensos em regiões mais baixas de frequência (apropriado para [b]) + 9 vogais [a] + sílaba /ta/, que apresentou duração de 250ms. A figura 46 representa o primeiro estímulo do continuum 1. Nesta figura, observa-se o estímulo de melhor exemplar da categoria [b].

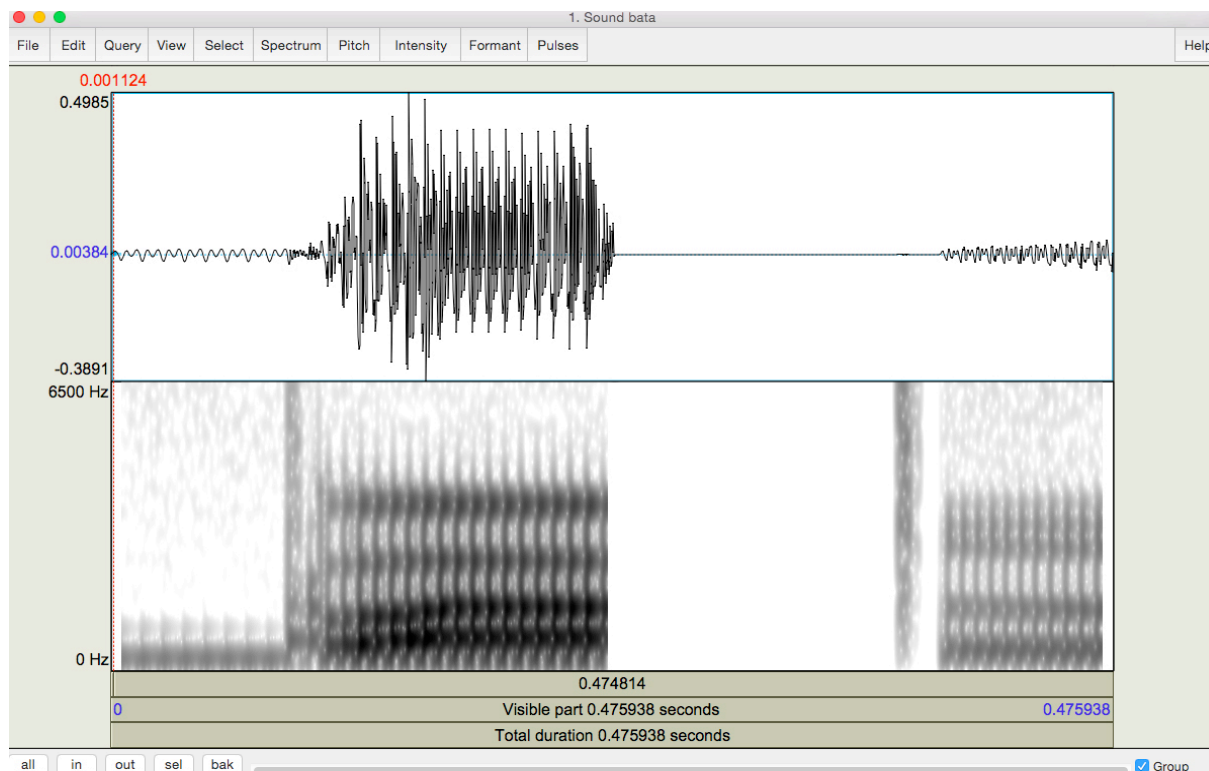


FIGURA 46 - Espectrograma do melhor exemplar da palavra /bata/ sintetizada

O segundo continuum foi composto por: amplitudes do *burst* com picos de energia mais intensos em regiões mais altas de frequência (apropriado para [d]) + 9 vogais [a] + sílaba /ta/. A figura 47 representa o nono estímulo do continuum 2, considerado o melhor exemplar da categoria [d].

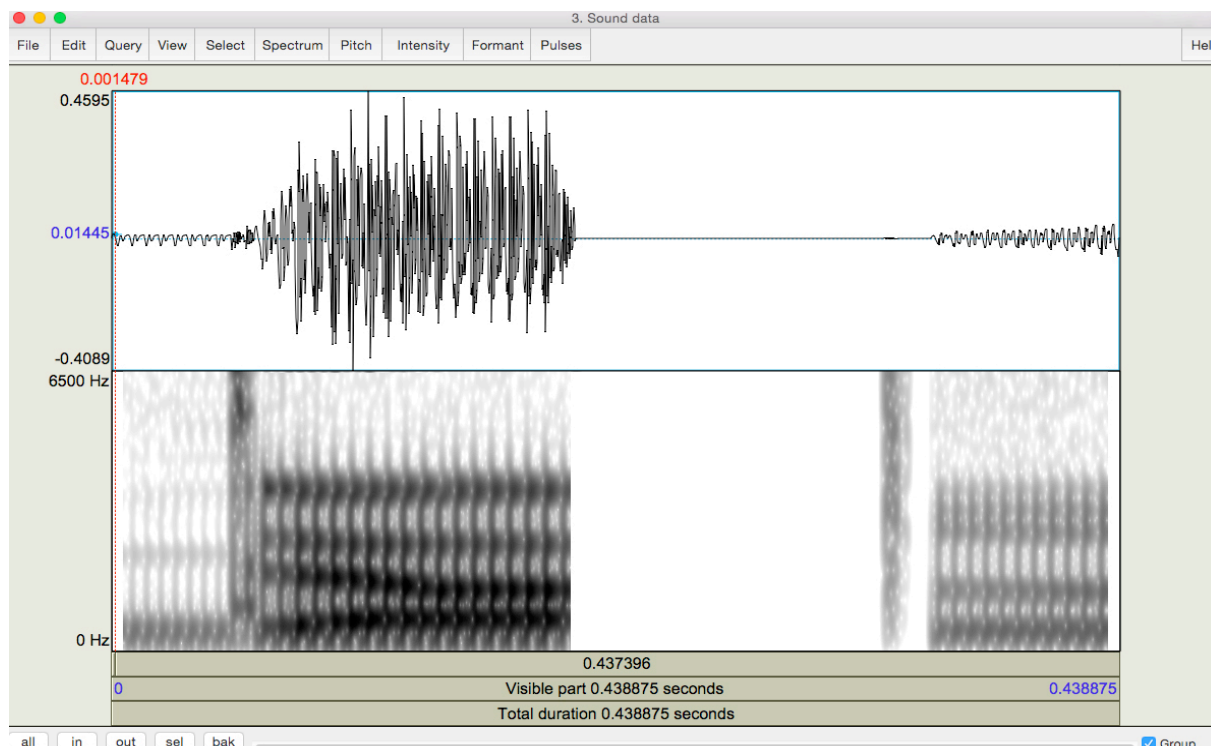


FIGURA 47 - Espectrograma do melhor exemplar da palavra /data/ sintetizada

Tanto os estímulos do continuum 1 quanto os do continuum 2 apresentaram duração total de 490ms e cada continuum continha nove estímulos.

5.1.4 Procedimento experimental

A apresentação dos estímulos aos participantes foi realizada mediante a utilização dos fones de ouvido *Philips*, acoplados ao *laptop HP Pavilion dv2000*, que continha o *software PercEval* (versão 3.0.5.0).

Com o uso deste *software*, desenvolveu-se uma tarefa de classificação dos estímulos, em que os participantes ouviam os estímulos por meio dos fones de ouvido e tinham que decidir qual a categoria mais coerente para os estímulos apresentados dentre as possibilidades dispostas na tela do *laptop* (por exemplo, /japa/-/sapa/). As figuras 48, 49 e 50 mostram a tela de exibição dos estímulos no *software PercEval* para os pares mínimos /japa/-/sapa/ e /bata/-/data/ e para o par de pseudopalavras /jaba/-/saba/. Abaixo das imagens à esquerda, que representavam /japa/, /jaba/ e /bata/, foi inserido um círculo vermelho e abaixo das imagens à direita, que representavam /sapa/, /saba/ e /data/, foi inserido um círculo azul. Os

mesmos símbolos coloridos também estavam dispostos nas teclas “ctrl” do lado direito e do lado esquerdo do *laptop*.

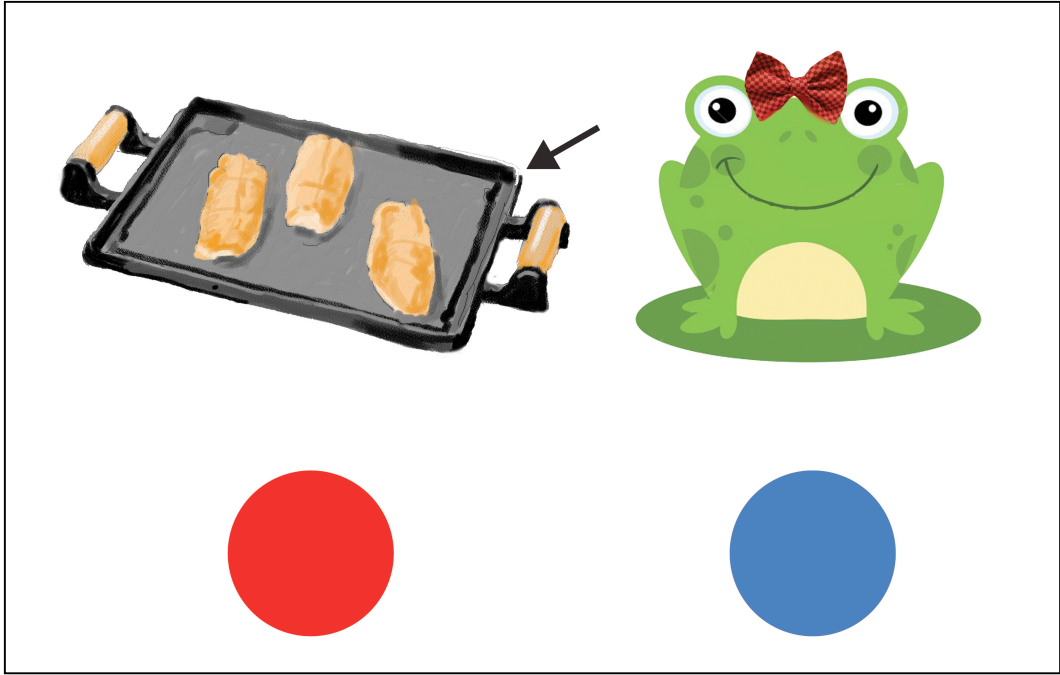


FIGURA 48 - Tela de exibição do par mínimo /japa/-/sapa/

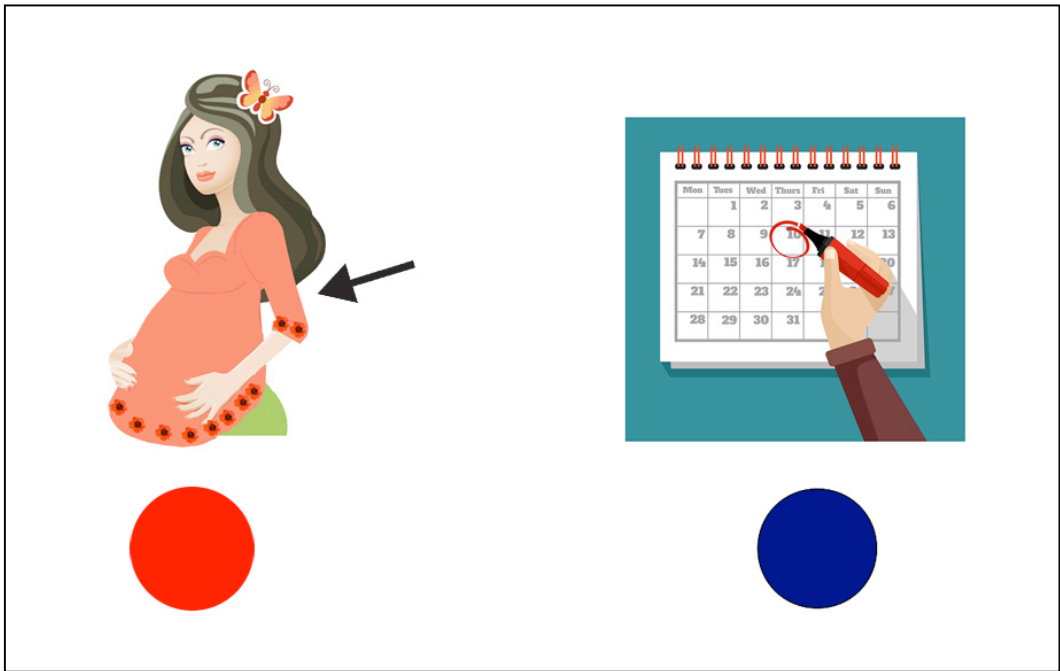


FIGURA 49 - Tela de exibição do par mínimo /bata/-/data/

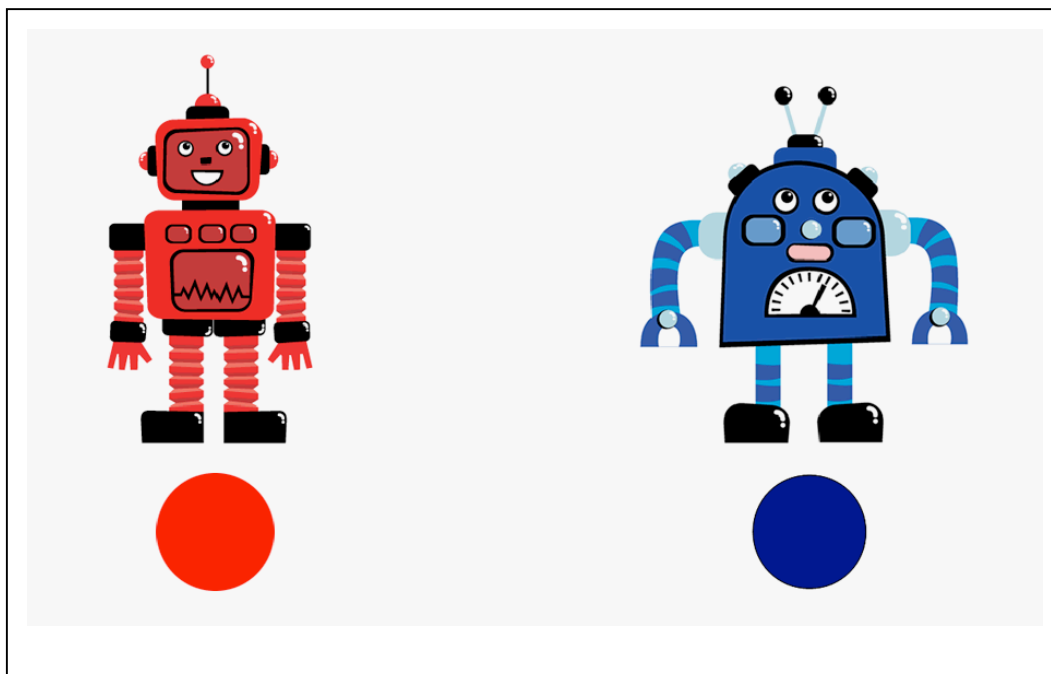


FIGURA 50 - Tela de exibição do par de pseudopalavras /faba-/saba/

No *software* PercEval foram produzidos quatro *scripts* para cada par de estímulos investigados. Os *scripts* foram denominados da seguinte forma: *script 1*, *script 2*, *script 3* e *script 4*. Para os pares de palavras compostos pelas fricativas, cada um destes *scripts* continha 50 julgamentos: 10 *estímulos treino* (5 apresentações para o melhor exemplar de cada categoria) e 40 *estímulos teste* (20 estímulos com transição de F2 e F3 apropriados para acompanhar [j] e 20 estímulos com transição de F2 e F3 apropriados para acompanhar [s]). No caso dos pares de palavras compostos pelas plosivas, cada um destes *scripts* era composto por 46 julgamentos: 10 *estímulos treino* (5 apresentações para o melhor exemplar de cada categoria) e 36 *estímulos teste* (18 estímulos com amplitudes do *burst* compatível com [b] e 18 estímulos com amplitudes do *burst* compatível com [d]).

Os quatro *scripts* dos pares mínimos /fapa-/sapa/ e /bata-/data/ e do par de pseudopalavras /faba-/saba/ diferiram em relação à ordem de apresentação dos *estímulos treino* e à ordem de apresentação dos *estímulos teste*.

Os estímulos dos pares mínimos /fapa-/sapa/ e /bata-/data/ e do par de pseudopalavras /faba-/saba/ foram apresentados em três sessões (uma sessão para cada par de estímulos).

Cada sessão foi composta por três blocos de teste para os pares mínimos /fapa/-/sapa/ e /bata/-/data/ e três blocos de teste para o par de pseudopalavras /faba/-/saba/. Em cada bloco de teste foi apresentado um *script*, sendo que 50% dos sujeitos foram submetidos aos *scripts* 1, 2 e 3, enquanto os outros 50% foram submetidos aos *scripts* 2, 3 e 4.

O intervalo de resposta (do inglês, *response delay*) foi de cinco segundos. Ou seja, depois que o participante respondia ou passados cinco segundos sem que o participante respondesse, iniciava-se uma nova prova, com a tela vazia exibida por um segundo, antes da apresentação do próximo estímulo.

5.1.5 Resultados e considerações do estudo

Esta seção apresenta os resultados obtidos por meio do experimento piloto de classificação dos pares mínimos /fapa/-/sapa/ e /bata/-/data/ e do par de pseudopalavras /faba/-/saba/, realizado com os adultos.

Os dados coletados foram tabulados no programa *Excel*, onde desenvolveu-se uma planilha com as seguintes variáveis independentes: idade, pistas auditivas manipuladas para as fricativas e plosivas e estatuto lexical (palavra e pseudopalavra). A variável dependente (respostas dos participantes da pesquisa) também foi inserida nesta planilha. As respostas brutas dos participantes foram analisadas e contabilizadas com o objetivo de gerar as curvas de classificação dos estímulos. As figuras apresentadas adiante relacionam-se à frequência de respostas /data/, /sapa/ e /saba/.

Os resultados foram apresentados separadamente, com a descrição das figuras. Primeiro, realizou-se as considerações referentes às respostas do par mínimo composto pelas plosivas /bata/-/data/. Depois, apresentou-se os resultados para o par mínimo /fapa/-/sapa/ e o par de pseudopalavras /faba/-/saba/, para os dois informantes selecionados e descritos na seção 5.1.2.1.2.

5.1.5.1 Resultados da tarefa de classificação com o par mínimo /bata/-/data/

A figura 51 permite observar visualmente que os adultos apresentaram curvas de classificação íngremes, o que indica respostas categóricas. Além disso, observa-se que as

curvas de classificação encontram-se mais íngremes quando se trata dos estímulos compostos pelas amplitudes do *burst* apropriadas para acompanhar [d].

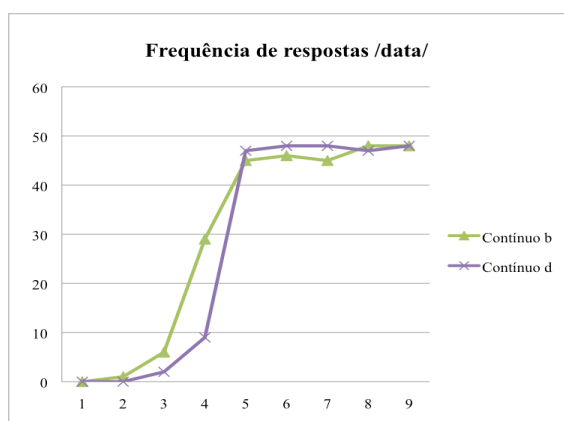
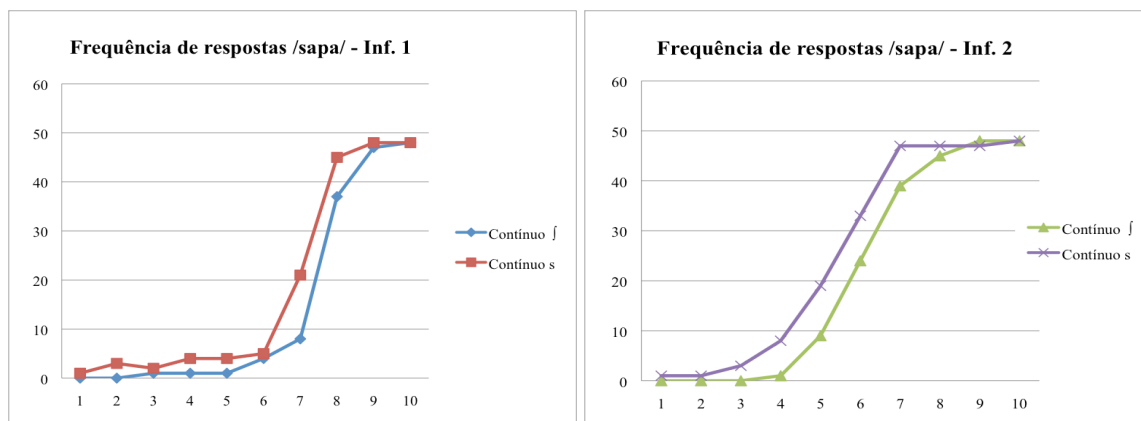


FIGURA 51 - Resultados da tarefa de classificação dos adultos para o par mínimo /bata/-/data/
 Legenda: eixo X: passos do continuum; eixo y: frequência de respostas /data/; continuum b: amplitudes do *burst* apropriadas para acompanhar [b]; continuum d: amplitudes do *burst* apropriadas para acompanhar [d]
 Fonte: Dados da pesquisa

A seguir os resultados com o par mínimo /ʃapa/-/sapa/.

5.1.5.2 Resultados da tarefa de classificação com o par mínimo /ʃapa/-/sapa/, dos informantes 1 e 2

A figura 52 relaciona-se ao resultado da tarefa de classificação utilizando os estímulos sintetizados com base nas propriedades acústicas dos sons produzidos pelo informante 1, enquanto a figura 53 refere-se ao resultado da tarefa de classificação utilizando os estímulos sintetizados com base nas propriedades acústicas dos sons produzidos pelo informante 2.



FIGURAS 52 e 53 - Resultados da tarefa de classificação dos adultos para o par mínimo /japa/-/sapa/, informante 1 e 2

Legenda: eixo X: passos do continuum; eixo y: frequência de respostas /japa/ (à esquerda) e frequência de respostas /sapa/ (à direita); contínuo [j]: transição do F2 e do F3 apropriada para acompanhar [j]; contínuo [s]: transição do F2 e do F3 apropriada para acompanhar [s]; figura do informante 1 à esquerda; figura do informante 2 à direita

Fonte: Dados da pesquisa

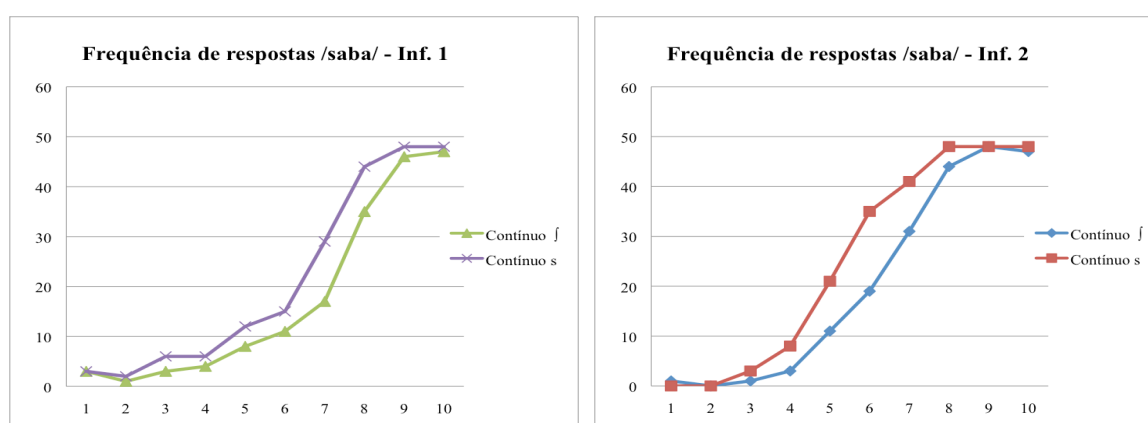
A partir da análise das figuras acima observa-se que as respostas dos adultos foram consistentes e categóricas diante dos estímulos produzidos com base nas características acústicas dos dois informantes selecionados. A região de fronteira fonêmica, isto é, a região onde observa-se a mudança da categoria /japa/ para /sapa/ é diferente quando compara-se as curvas de classificação dos participantes do estudo. Diante dos estímulos do informante 1, os participantes classificaram os estímulos como /japa/ até o sexto passo do continuum. Já com os estímulos do informante 2 observa-se que a partir do quarto passo do continuum inicia-se a mudança de categoria.

A diferença principal entre os estímulos dos dois informantes concentra-se no valor do F2. A vogal [a] criada com base nos valores do informante 1 apresentou um decréscimo de 126Hz da fase de transição para a fase estável da vogal [a] seguinte à fricativa [j]. Já para o informante 2, ocorreu um decréscimo de 210Hz. Ainda com esta diferença no valor do F2 entre os informantes, não observa-se visualmente uma diferença importante entre as curvas de classificação dos adultos.

Os resultados obtidos para o par de pseudopalavras /jaba/-/saba/ serão apresentados adiante.

5.1.5.3 Resultados da tarefa de classificação com o par de pseudopalavras /faba/-/saba/, dos informantes 1 e 2

As figuras abaixo mostram as respostas da tarefa de classificação dos adultos para o par de pseudopalavras /faba/-/saba/. Em comparação com o par mínimo /fapa/-/sapa/ observa-se que as curvas de classificação foram menos íngremes. As curvas mantiveram bem próximas diante dos estímulos produzidos com as propriedades acústicas do informante 1. A proximidade entre as curvas sinaliza que, provavelmente, os ouvintes não utilizaram a pista de transição dos formantes para classificar o par de palavras /faba/-/saba/. Tal fato será confirmado no experimento de ponderação de pistas auditivas.



FIGURAS 54 e 55 - Resultados da tarefa de classificação dos adultos para o par de pseudopalavras /faba/-/saba/, informante 1 e informante 2

Legenda: eixo X: passos do continuum; eixo y: frequência de respostas /fapa/ (à esquerda) e frequência de respostas /sapa/ (à direita); continuum f: transição do F2 e do F3 apropriada para acompanhar [f]; continuum s: transição do F2 e do F3 apropriada para acompanhar [s]; figura do informante 1 à esquerda; figura do informante 2 à direita

Fonte: Dados da pesquisa

A configuração mais plana das curvas acima indica que houve maior dúvida para as classificação das pseudopalavras /faba/-/saba/.

Após a descrição dos resultados obtidos com o estudo piloto de análise das propriedades acústicas dos estímulos e da tarefa de classificação selecionou-se os estímulos do informante 2, para execução do experimento final de ponderação de pistas auditivas. Os estímulos deste informante foram selecionados, uma vez que eles apresentam valores dos formantes vocálicos mais próximos dos valores médios do grupo.

6 CAPÍTULO 4 – A PERCEPÇÃO DAS FRICATIVAS CORONAIIS NÃO VOZEADAS [ʃ]-[s] E DAS PLOSIVAS VOZEADAS [b]-[d] POR CRIANÇAS E ADULTOS FALANTES NATIVOS DO PB

Este capítulo tem o objetivo de investigar a ponderação de pistas auditivas para a percepção das fricativas coronais não vozeadas [ʃ]-[s] e das plosivas vozeadas [b]-[d], por crianças e adultos falantes nativos do PB. Aqui, descreveu-se as informações relacionadas a: sujeitos, estímulos, desenho e procedimento experimental, plano de análise, resultados, discussões e conclusões do estudo.

Esta pesquisa foi realizada com a aprovação do Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, Parecer CAAE – 62617616.0.0000.5149.

A seguir, desenvolveu-se a metodologia da presente investigação.

6.1 Metodologia do estudo de ponderação de pistas auditivas

6.1.1 Sujeitos da pesquisa

Os participantes deste estudo foram: 12 crianças de quatro anos de idade, 12 crianças de sete anos e 12 adultos, todos falantes nativos do português brasileiro. As crianças de quatro anos foram recrutadas em duas instituições de ensino de Belo Horizonte/MG: Colégio Salesiano e Escola Infantil “O Pequeno Construtor”. As crianças de sete anos foram recrutados no Colégio Salesiano e os adultos foram recrutados na Superintendência Central de Perícia Médica e Saúde Ocupacional (SCPMSO), órgão público do Estado de Minas Gerais. A coleta de dados foi realizada no período de agosto a setembro de 2016, em uma sala silenciosa tanto nas escolas quanto na SCPMSO.

Os responsáveis pelas crianças de quatro e de sete anos, assim como as crianças de sete anos e os adultos assinaram o Termo de Consentimento Livre e Esclarecido, após serem informados sobre os aspectos gerais da pesquisa, tais como: objetivo do estudo, caráter voluntário de sua participação e ausência de riscos ou desconfortos (ANEXOS 5, 6 e 7).

Os critérios de seleção utilizados para a participação no experimento foram: desenvolvimento de fala conforme o esperado para idade e ausência de histórico significativo de otite média, de colocação de tubo de ventilação e de perda auditiva de qualquer grau. De acordo com o Consenso Brasileiro de Otites Médias de 1999, o histórico significativo de otites médias relaciona-se à ocorrência de três ou mais episódios de otite média aguda em seis meses ou quatro ou mais episódios em doze meses (CAMPOS, OLIVEIRA, ENDO *et al.*, 1999). Com o intuito de obter estas informações, os responsáveis pelas crianças de quatro e de sete anos responderam a um questionário, que se encontra em anexo (ANEXO 8).

Todos os sujeitos passaram previamente por uma triagem auditiva, em ambas as orelhas, que consistiu em tons puros, nas frequências de 0.5, 1, 2, 4 e 8kHz, apresentados em 25dB para as crianças de sete anos e para os adultos (DAVIS e SILVERMAN, 1970) e em 15 dB para as crianças quatro anos (NORTHERN e DOWNS, 1989). Para a realização da triagem utilizou-se o audiômetro AD-28.

As crianças de quatro anos também foram submetidas a uma avaliação fonológica (YAVAS, HERNANDORENA e LAMPRECHT, 2002), com o objetivo de verificar se os sons fricativos [s] e [ʃ] e os sons plosivos [b] e [d] faziam parte de seu inventário fonético e fonológico. Os sons fricativos seguem os plosivos e as nasais na ordem de aquisição segmental das línguas naturais. As plosivas e as nasais são os primeiros segmentos consonantais a serem adquiridos pelas crianças com desenvolvimento fonológico normal, sendo que ambos são adquiridos antes dos dois anos de idade. Direcionando para os sons investigados neste estudo, sugere-se que a aquisição da fricativa [s] antecede a aquisição da fricativa [ʃ], sendo que a primeira ocorre por volta dos dois anos e seis meses e a segunda por volta dos dois anos e dez meses (OLIVEIRA, 2002; SÁVIO, 2001). Assim, na aquisição normal, espera-se que as crianças de quatro anos sejam capazes não só de produzir estes sons adequadamente, mas também de empregá-los com função distintiva na língua.

Inicialmente, 14 crianças de quatro anos de idade foram selecionadas para participar do experimento de ponderação de pistas auditivas. Destas, 2 foram excluídas, uma vez que não compreenderam adequadamente a tarefa e apresentaram desempenho insatisfatório na classificação dos estímulos treino (melhores exemplares de cada categoria). Assim, 12 crianças (7 meninas e 5 meninos), com idade média de 4 anos e 5 meses participaram do

experimento de classificação das palavras /ʃapa-/sapa/, /bata-/data/ e das pseudopalavras /ʃaba-/saba/.

No grupo das crianças de sete anos de idade não houve a exclusão de nenhum participante. O grupo desta faixa etária ficou composto por 12 crianças (6 meninos e 6 meninas), com idade média de 7 anos e 5 meses. Do mesmo modo, os 12 adultos selecionados participaram do experimento. Os adultos apresentaram idades entre 20 e 40 anos, com idade média de 36 anos e 6 meses. O grupo foi composto por 12 mulheres.

6.1.2 Estímulos, continua e procedimento experimental

Os estímulos utilizados no experimento de ponderação de pistas auditivas (ruído fricativo, amplitudes do *burst* das plosivas [b] e [d] e porções vocálicas seguinte às fricativas e às plosivas), bem como os continua produzidos e o procedimento experimental foram descritos nas seções 5.1.2, 5.1.3 e 5.1.4, do capítulo anterior. A seguir descreveu-se as diferenças na coleta das respostas das crianças em relação aos adultos.

6.1.3 Coleta de dados com as crianças de quatro e de sete anos de idade e com os adultos

A coleta de dados com as crianças de quatro e de setes anos de idade apresentou uma diferença, uma vez que somente as crianças de quatro anos foram submetidas à avaliação fonológica. No primeiro encontro realizou-se a avaliação fonológica (com as crianças de quatro anos) e a triagem auditiva de todos os participantes. No segundo encontro, as crianças foram submetidas ao experimento de ponderação de pistas auditivas.

Conforme relatado, os estímulos dos pares mínimos /ʃapa-/sapa/ e /bata-/data/ e do par de pseudopalavras /ʃaba-/saba/ foram apresentados em três sessões (uma sessão para cada par de estímulos) e em dias diferentes.

A figura a seguir representa uma revisão do procedimento experimental utilizado com as crianças de quatro e de sete anos e com os adultos. Apenas as crianças de quatro e de sete anos realizaram o treino com nomeação de figuras e o treino com os melhores exemplares de

cada categoria separadamente. Os adultos iniciaram o experimento diretamente no bloco de teste 1.

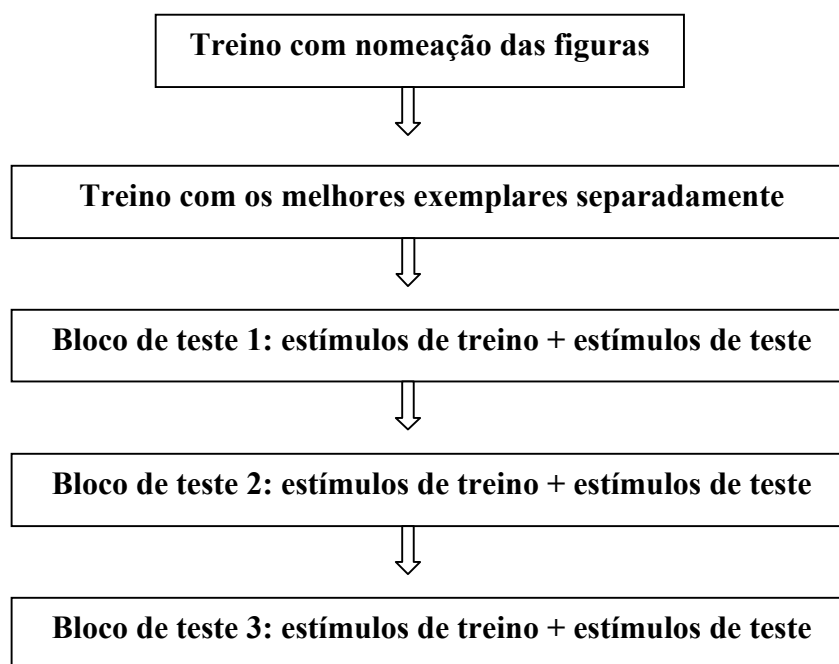


FIGURA 56 - Procedimento experimental utilizado com as crianças de quatro e de sete anos e com os adultos.

Fonte: (Adaptado de Gerrits, 2001, p.92)

Inicialmente, as crianças foram apresentadas às figuras que representavam os pares de palavras investigados. Narrou-se uma história criada para os pares mínimos /fapa/-/sapa/ e /bata/-/data/ e para o par de pseudopalavras /faba/-/saba/, com o objetivo de envolver a criança na tarefa e de contextualizar as imagens utilizadas durante todo o procedimento.

A fase de treino com a nomeação de figuras apresentou duas etapas. Na primeira etapa, a pesquisadora apontava para as figuras e as crianças tinham que nomeá-las. Após cinco julgamentos corretos, invertia-se o jogo e na segunda etapa, a pesquisadora falava o nome das figuras e as crianças tinham que localizá-las e nomeá-las.

Na segunda fase de treino, posicionava-se uma imagem na frente da criança e o melhor exemplar da categoria era apresentado pelos fones de ouvido. Após fazer isso para uma categoria de resposta, removia-se a imagem e o procedimento era repetido com a outra categoria de resposta.

Em seguida, iniciavam-se as fases de bloco de teste 1, 2 e 3, consecutivamente. Cada bloco de teste continha a apresentação de estímulos treino e de estímulos teste. Nestas fases, ambas as figuras eram posicionadas na frente da criança e esta era treinada a indicar e repetir a palavra ouvida. A pesquisadora registrou as respostas no computador tanto para as crianças de quatro quanto para as crianças de sete anos. Após cada bloco de teste realizava-se uma pausa curta, na qual a criança recebia um reforço positivo. Tal estratégia foi desenvolvida com o intuito de evitar o cansaço com a tarefa.

A sessão das crianças de quatro e de sete anos apresentou duração aproximada de 35 minutos e 25 minutos, respectivamente. Os adultos desempenharam a tarefa em aproximadamente 15 minutos e iniciaram no experimento na fase “Bloco de teste 1”, descrita na figura 56. O único grupo de idade que registrou as respostas no computador foi o dos adultos. Assim, este grupo tinha que indicar no teclado a imagem que representava o estímulo e repetir a palavra ouvida.

6.2 Plano de análise

As respostas obtidas com o experimento de ponderação de pistas auditivas para a percepção dos pares mínimos /ʃapa/-/sapa/ e /bata/-/data/ e do par de pseudopalavras /ʃaba/-/saba/, das crianças de quatro e de sete anos de idade e dos adultos, foram analisadas estatisticamente por meio do *software* R, versão 3.3.1 e do *software* SPSS, versão 10. Com o objetivo de permitir uma comparação direta dos resultados deste estudo com os estudos de Nittrouer e colaboradores (Nittrouer e Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer e Miller, 1997a) utilizou-se o modelo de análise probito (*probit analysis*; FARAWAY, 2006).

O modelo probito é um tipo de regressão utilizado para analisar variáveis de respostas binomiais. Ele transforma uma curva de resposta sigmóide em uma linha reta, que pode então ser avaliada por meio da análise de regressão. Desse modo, a transformação probito extrai informações da taxa de mudança dos dados de uma curva sigmóide e estima duas medidas para cada curva de respostas dos sujeitos: a fronteira fonêmica (do inglês, *phoneme boundary*) e a inclinação das curvas (do inglês, *slope*). No caso dos estímulos compostos pelas fricativas, a fronteira fonêmica relaciona-se ao ponto, ao longo do continuum fricativo, no qual 50% das respostas foram [ʃ] e 50% foram [s]. Já no caso dos estímulos compostos pelas plosivas, a fronteira fonêmica relaciona-se ao ponto, ao longo do continuum vocálico,

no qual 50% das respostas foram [b] e 50% foram [d]. O grau de separação das duas curvas de respostas foi calculado pela diferença das médias das duas curvas. Isso proporciona uma medida da extensão pela qual as fronteiras fonêmicas dos ouvintes foram influenciadas pela mudança na informação transicional dos dois continua. Assim, a separação das curvas de respostas representa o peso atribuído à pista auditiva de transição dos formantes para as fricativas e o peso atribuído à pista de amplitudes do *burst* para as plosivas.

A inclinação das curvas de classificação corresponde ao grau de categorização das respostas dos ouvintes e, além disso, corresponde ao peso atribuído à pista auditiva de frequência do ruído fricativo ou de transição dos formantes, diante do par mínimo /bata/-/data/.

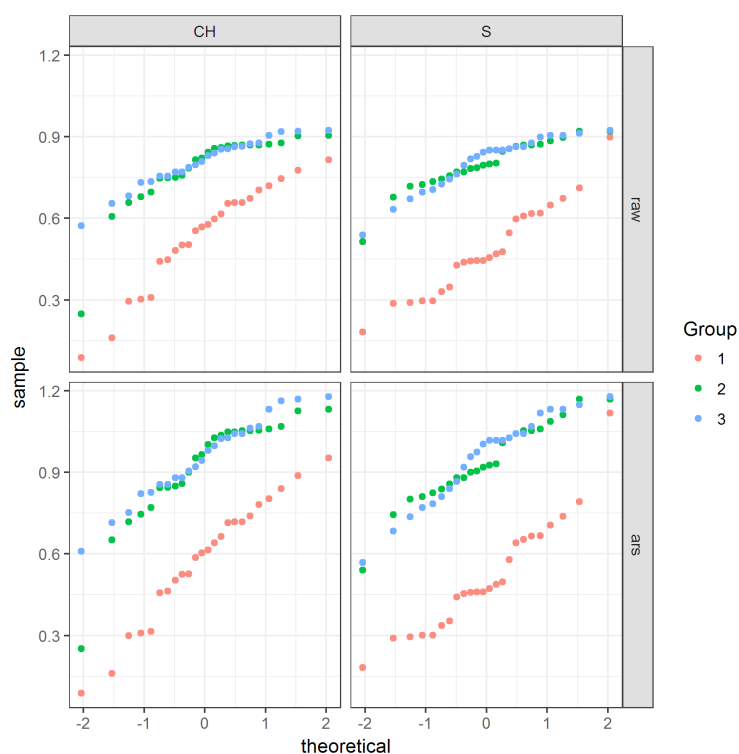
A separação das curvas foi retratada em Hertz (Hz), que corresponde a uma diferença na frequência de localização da fronteira fonêmica ao longo dos continua /ʃapa/-/sapa/, /ʃaba/-/saba/ e /bata/-/data/. Já a inclinação das curvas foi retratada em termos da mudança de unidades proibito por unidades do ruído de fricção ou unidades do continuum vocálico, para as plosivas.

A fronteira fonêmica e a inclinação das curvas foram utilizadas como medidas dependentes, e a influência do Grupo (idade) e do Continuum (/ʃ/, /s/ e continuum vocálico para as plosivas) sobre as variáveis dependentes foi testada realizando análise de variância (ANOVA) para as fricativas e o “teste t pareado” para as plosivas.

Antes de realizar a ANOVA para os dados coletados com as fricativas, verificou-se por meio do teste de Kolmogorov-Smirnov (com correção de significância de Lilliefors) que a distribuição dos dados quase sempre não se afastava da normalidade, exceto para (a) as “Fronteiras” fonêmicas do grupo de sete anos no continuum com transição dos formantes adequada para acompanhar [s], nas palavras [$D(12) = 0.244$, $p = 0.046$] e (b) as “Fronteiras” fonêmicas do grupo de adultos no continuum com transição dos formantes adequada para acompanhar [s], nas palavras [$D(12) = 0.256$, $p = 0.029$]. Com o objetivo de corrigir a anormalidade dos dados, todos os valores de “Inclinação” e “Fronteira” foram transformados em arcosenos. Antes da transformação, dividiram-se os valores de fronteira por 10, de modo que ficassem entre 0 e 1 necessários à transformação. O efeito da transformação pode ser vista no gráfico 3 (Q-Q plot).

GRÁFICO 3

Q-Q plot - Efeito da transformação dos dados em arcosenos



Legenda: raw = dados brutos; ars = dados transformados (arcosseno); group 1 = crianças de quatro anos; group 2 = crianças de sete anos e group 3 = adultos
 Fonte: Dados da pesquisa

Por fim, aplicou-se o teste de Levene baseado na mediana para verificar a homogeneidade de variância entre os grupos. O teste mostrou que todos os grupos tinham variâncias iguais.

Feita esta análise preliminar, realizou-se dois conjuntos de análise, um conjunto para as respostas relacionadas ao par mínimo /*sapa*-/*sapa*/, e outro, para as respostas relacionadas ao par de pseudopalavras /*faba*-/*saba*/ . O objetivo foi verificar como os fatores Grupo (Idade) e Continuum ([*f*] e [*s*]) afetaram as variáveis "Inclinação" e "Fronteira", respectivamente, considerando o estatuto lexical (palavra x pseudopalavra) como uma condição experimental. Desta forma, Grupo e Continuum foram as variáveis de interesse e não importou comparar os resultados dos diferentes grupos de idade ou dos dois continua em função do estatuto. Realizou-se ANOVA em dois níveis com medidas repetidas, com Grupo como fator entre-sujeitos e Continuum como fator inter-sujeitos. Como medida do tamanho do efeito para os

fatores utilizou-se o "eta² generalizado" (*generalized eta squared*; BAKEMAN, 2005)¹⁰. Para avaliar as diferenças entre os pares de grupos, utilizou-se como teste *post hoc* HSD de Tukey.

A análise estatística realizada para o par mínimo composto pelas plosivas foi desenvolvida de forma diferente das fricativas. Para tal, foram obtidas quatro medidas a partir das respostas dos sujeitos, “Inclinação” e “Fronteira” para o continuum com amplitudes do *burst* adequadas para acompanhar [b] e para o continuum com amplitudes do *burst* adequadas para acompanhar [d], respectivamente. O teste de Kolmogorov-Smirnov para normalidade revelou que tanto a “Inclinação” quanto a “Fronteira” mostraram-se adequados em todos os três grupos etários para ambos os continua. Entretanto, o teste de Levene para homogeneidade de variância entre os grupos revelou homoscedasticidade apenas para os dados do continuum [b]. No caso do continuum [d], em todos os três grupos, tanto para a “Inclinação” quanto para a “Fronteira”, várias transformações foram realizadas (logarítmica, exponencial, raiz quadrada, arcosseno), sem sucesso, para tentar obter dados adequados à realização de ANOVA com medidas repetidas. Com isso, optou-se por uma estratégia de análise que investiga apenas a diferença entre os grupos para a “Inclinação” e para a “Fronteira” por meio de testes “t pareados” e corrigidos pelo método de Bonferroni, tomando conjuntamente como variável dependente os valores dos continua [b] e [d]. Além disso, relatam-se adiante os valores de “r” como medidas do tamanho de efeito¹¹.

Após a realização da análise dos dados utilizando o modelo probito e a ANOVA realizou-se a análise de variância multivariada (MANOVA), com medidas repetidas, com o intuito de verificar se os participantes do estudo modificaram as suas estratégias de ponderação de pistas diante do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/, ou seja, se o estatuto lexical gerou mudanças no padrão de respostas dos participantes. Nesta análise, a “Inclinação” e a “Fronteira Fonêmica” foram as variáveis dependentes, o Grupo (Idade)

¹⁰ O eta² generalizado é uma medida do tamanho do efeito que pode ser utilizada na ANOVA. Cohen (1988 apud BAKEMAN, 2005, p. 383) define os seguintes valores do eta² generalizado: a) 0.02 (efeito pequeno); b) 0.13 (efeito médio); e c) 0.26 (efeito grande).

¹¹ O coeficiente de correlação “r” de Pearson também é uma medida do tamanho do efeito que pode ser utilizada no teste “t pareado”. Cohen (1988, 1992 apud FIELD, 2009, p. 57) estipulou os seguintes valores do “r” de Pearson: a) 0.1 (efeito pequeno); b) 0.3 (efeito médio); e c) 0.5 (efeito grande).

foi a variável entre grupos enquanto o Estatuto Lexical e o Continuum ([ʃ] e [s]) foram as variáveis dentro dos grupos.

A seguir serão descritos os resultados e a discussão do experimento de ponderação de pistas auditivas.

6.3 Resultados e discussão

Nesta seção apresentou-se os resultados alcançados com o experimento de ponderação de pistas auditivas com as crianças de quatro anos, as de sete anos e os adultos diante dos pares mínimos /ʃapa/-/sapa/ e /bata/-/data/ e do par de pseudopalavras /ʃaba/-/saba/.

Inicialmente, realizou-se uma descrição qualitativa das respostas dos participantes e logo após abordou-se de forma mais detalhada a análise quantitativa.

De um total de 4320 julgamentos para /ʃapa/-/sapa/ e 4320 /ʃaba/-/saba/ somente 4 julgamentos foram perdidos para /ʃapa/-/sapa/, devido a falta de resposta dos participantes no tempo apropriado, e apenas 1 julgamento foi perdido para /ʃaba/-/saba/. Diante do par mínimo /bata/-/data/, dos 3888 julgamentos, 7 foram perdidos.

A figura 57 mostra os resultados obtidos com a tarefa de classificação do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/, das crianças de quatro e das de sete anos de idade e dos adultos. O eixo X relaciona-se aos dez passos de frequência do continuum fricativo e o eixo Y refere-se a porcentagem de respostas [s]. Na medida em que ocorre o aumento da frequência do ruído fricativo, maior será a probabilidade do participante responder /sapa/ ou /saba/. Desta forma, o primeiro passo do continuum apresenta valores de frequência do ruído fricativo mais apropriados para [ʃa], enquanto o décimo passo do continuum apresenta valores de frequência do ruído fricativo mais apropriados para [sa]. Para cada grupo de idade observa-se quatro curvas de classificação, sendo que duas relacionam-se às respostas direcionadas ao par mínimo /ʃapa/-/sapa/ (uma curva com transição dos formantes apropriada para acompanhar [ʃ] e uma curva com transição dos formantes apropriada para acompanhar [s] - linhas cheias) e as outras duas referem-se às respostas direcionadas ao par de pseudopalavras /ʃaba/-/saba/ (uma curva com transição dos

formantes apropriada para acompanhar [ʃ] e uma curva com transição dos formantes apropriada para acompanhar [s] - linhas pontilhadas).

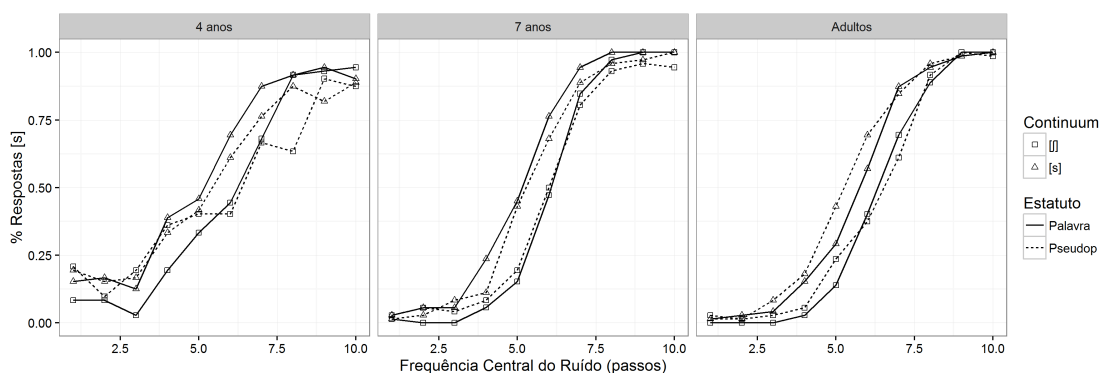


FIGURA 57 - Resultados da tarefa de classificação para o par mínimo /ʃapa/-/sapa/ e para o par de pseudopalavras /ʃaba/-/saba/

Legenda: eixo X: passos do contínuo fricativo; eixo Y: porcentagem de respostas [s]; Continuum [ʃ]: transição dos formantes apropriada para acompanhar [ʃ]; Continuum [s]: transição dos formantes apropriada para acompanhar [s]; Estatuto: Palavra (/ʃapa/-/sapa/) e Estatuto: Pseudo (/ʃaba/-/saba/)

Fonte: Dados da pesquisa

Se as crianças utilizaram mais a pista de transição dos formantes em relação à pista de altura da frequência do ruído fricativo, ao contrário dos adultos, espera-se observar uma diminuição no peso atribuído à pista de transição dos formantes, ou seja, espera-se verificar curvas menos separadas com o aumento da idade. Por outro lado, um aumento no peso atribuído à pista de altura da frequência do ruído fricativo será manifestado como um aumento gradual na inclinação das curvas.

Conforme esperado, observa-se um efeito da idade, onde as curvas de classificação dos adultos foram mais íngremes do que as curvas das crianças mais jovens, o que significa que os adultos apresentaram respostas mais categóricas, tanto diante do par mínimo /ʃapa/-/sapa/ quanto diante do par de pseudopalavras /ʃaba/-/saba/.

Outra informação importante fornecida pela figura 57 relaciona-se a separação das curvas de classificação dos participantes. Quanto mais separadas as curvas maior o peso atribuído à pista de transição dos formantes. A figura acima mostra que os adultos apresentaram curvas de classificação mais próximas, seguidos das crianças de sete e das de quatro anos de idade, para a classificação do par mínimo /ʃapa/-/sapa/. Com as pseudopalavras /ʃaba/-/saba/ observa-se que as crianças de sete anos e os adultos também apresentaram curvas de classificação mais próximas em relação as crianças de quatro anos. Desta forma, as crianças

mais jovens apresentaram uma separação maior entre curvas de classificação tanto diante do par mínimo /*ɟapa*-/*sapa*/ quanto diante do par de pseudopalavras /*ɟaba*-/*saba*/. Ao final da análise estatística será verificado se a separação entre as curvas de classificação foi significativa entre os grupos de idade.

A figura 58 mostra as curvas de classificação dos participantes para o par mínimo /*bata*-/*data*/. O eixo X relaciona-se aos valores de frequência do segundo formante do continuum vocálico (F2) e o eixo Y refere-se a porcentagem de respostas [d]. Do mesmo modo da figura 57, que retrata as curvas de classificação das fricativas, observa-se que na medida em que ocorre o aumento dos valores do F2, maior será a probabilidade do participante responder /*data*/. Assim, o primeiro passo do continuum apresenta valor de frequência do F2 mais apropriado para [ba], enquanto o nono passo do continuum apresenta valor de frequência do F2 mais apropriado para [da]. Para cada grupo de idade temos duas curvas de classificação: uma curva com amplitudes da fase de soltura (*burst*) apropriadas para acompanhar [b] (linhas compostas por quadrados) e outra curva com amplitudes da fase de soltura (*burst*) apropriadas para acompanhar [d] (linhas compostas por triângulos).

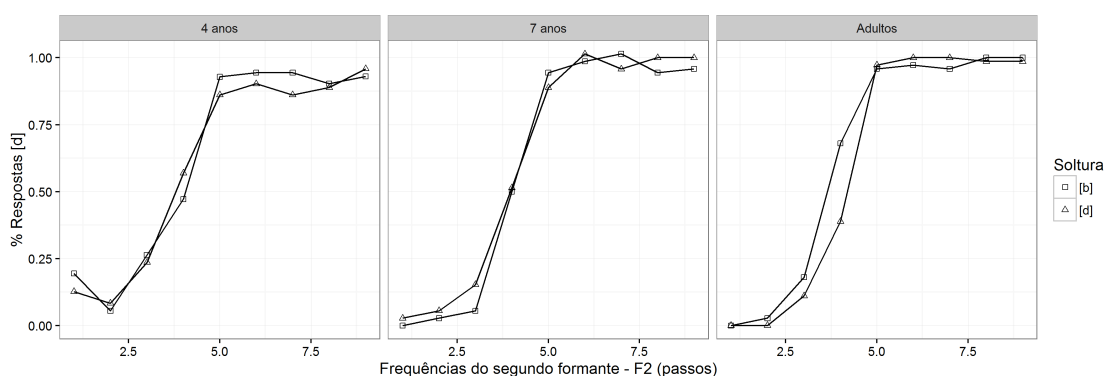


FIGURA 58 - Resultados da tarefa de classificação para o par mínimo /*bata*-/*data*/

Legenda: eixo X: passos do contínuo vocálico; eixo Y: porcentagem de respostas [d]; Soltura [b]: amplitudes da soltura (*burst*) apropriadas para acompanhar [b]; Soltura [d]: amplitudes da soltura (*burst*) apropriadas para acompanhar [d]

Fonte: Dados da pesquisa

Diante deste par observa-se que as curvas de respostas foram mais categóricas em relação aos pares /*ɟapa*-/*sapa*/ e /*ɟaba*-/*saba*/, até mesmo para o grupo de quatro anos de idade. No entanto, as curvas dos adultos e das crianças de sete anos ainda foram mais inclinadas do que as curvas das crianças de quatro anos.

Em relação à separação das curvas de classificação observa-se que as curvas de todos os grupos de idade estão bem próximas. Isto indica que, provavelmente, pouca importância foi dada a pista de amplitudes da fase de soltura para a classificação de /bata/-/data/. Cabe ressaltar que a análise individual das curvas de classificação dos participantes deste estudo para os pares mínimos /ʃapa/-/sapa/ e /bata/-/data/ e para o par de pseudopalavras /ʃaba/-/saba/ foi disponibilizada nos anexos 9, 10 e 11. As curvas de classificação que apresentaram valores de inclinação inferiores a 0.2 foram inseridas nas figuras em vermelho e indicaram as piores respostas do grupo. As curvas que apresentaram valores entre 0.2 e 0.3 foram inseridas em cinza e as curvas com os melhores valores de inclinação – isto é, superiores a 0.3 – foram inseridas em azul.

As impressões fornecidas pelas figuras 57 e 58 foram confirmadas pelos valores descritos na tabela 14, que mostra os valores médios da fronteira fonêmica e da inclinação das curvas de classificação estimados pelo modelo de análise probito, para os pares mínimos /ʃapa/-/sapa/ e /bata/-/data/ e para o par de pseudopalavras /ʃaba/-/saba/. Os valores médios da inclinação das curvas retratam as unidades probito por item (passos do continuum fricativo ou passos do continuum vocálico para as plosivas) e os valores médios da fronteira fonêmica relacionam-se aos passos do continuum fricativo ou do continuum vocálico em que se estimou que os participantes produziram 50% de respostas [ʃ] ou [s] e 50% de respostas [b] ou [d]. Como este é o nível do acaso, denota-se um ponto de incerteza decisória, que pode ser interpretado como a fronteira entre as duas categorias fônicas. Por exemplo, no par mínimo /ʃapa/-/sapa/, o valor médio da fronteira fonêmica para a curva de classificação com transição do F2 apropriada para acompanhar [ʃ] foi 5.78, para as crianças de quatro anos. Isso indica que a mudança da categoria [ʃ] para a categoria [s] ocorreu entre o quinto e o sexto passo do continuum fricativo.

TABELA 14

Valores médios e desvio padrão da fronteira fonêmica e da inclinação das curvas de classificação

	4 anos de idade		7 anos de idade		Adultos	
	M	SD	M	SD	M	SD
/japa-/sapa/						
Fronteira Fonêmica						
Cont [j]	5.78	0.850	5.77	0.673	6.06	0.784
Cont [s]	4.61	1.10	4.92	1.02	5.47	1.04
Diferença	1.17	-0.25	0.85	-0.347	0.59	-0.256
Inclinação						
Cont [j]	0.622	0.1384	0.846	0.0481	0.845	0.0644
Cont [s]	0.523	0.177	0.807	0.11	0.809	0.123
Diferença	0.099	-0.0386	0.039	-0.0619	0.036	-0.0586
/faba-/saba/						
Fronteira Fonêmica						
Cont [f]	5.71	1.434	5.84	1.51	5.85	1.102
Cont [s]	5.09	1.51	5.39	0.882	5.35	1.02
Diferença	0.62	-0.076	0.45	0.628	0.5	0.082
Inclinação						
Cont [f]	0.448	0.2029	0.718	0.1733	0.766	0.0954
Cont [s]	0.439	0.1532	0.791	0.0713	0.8	0.0753
Diferença	0.009	-0.5881	-0.073	0.102	-0.034	0.0201
/bata-/data/						
Fronteira Fonêmica						
Cont [b]	3.9	0.373	4.38	0.529	4.03	0.689
Cont [d]	4.14	1.104	4.04	0.474	4.29	0.458
Diferença	-0.24	-0.731	0.34	0.055	-0.26	0.231
Inclinação						
Cont [b]	0.65	0.136	0.88	0.084	0.882	0.084
Cont [d]	0.644	0.201	0.851	0.098	0.941	0.084
Diferença	0.006	-0.065	0.029	-0.014	-0.059	0

A tabela 14 revela que realmente os adultos e as crianças de sete anos apresentaram curvas de respostas mais íngremes do que as crianças de quatro anos, para o par mínimo /japa-/sapa/ (0.845, 0.809; 0.846, 0.807 e 0.622, 0.523, respectivamente) e para o par de pseudopalavras /faba-/saba/ (0.766, 0.8; 0.718, 0.791 e 0.448, 0.439, respectivamente). Tal fato, é consistente com os resultados obtidos por Nittrouer e Studdert-Kennedy (1987), Nittrouer e Miller (1996), Hanzan e Barret (2000), Gerrits (2001), Penido (2012), Mayo *et al.* (2003). Em todos estes estudos, desenvolvidos com o contraste [j]-[s], os pesquisadores observaram um efeito da idade, onde as curvas de classificação se tornaram mais íngremes com o aumento da idade, sugerindo que maior peso foi atribuído à pista de altura da frequência do ruído fricativo, na decisão do ponto de articulação das fricativas. No entanto,

Mayo e Turk (2004) não verificaram diferenças significativas nos valores de inclinação das duas curvas entre as crianças e os adultos, para a classificação do par /saI/-/ʃaI/ do inglês.

Da mesma forma, para o par mínimo /bata/-/data/ os adultos também apresentaram valores maiores da inclinação das curvas de classificação do que as crianças de quatro e de sete anos de idade (0.882, 0.941; 0.88, 0.851; 0.65, 0.644, respectivamente). Porém, diante deste par, tanto os adultos quanto as crianças apresentaram respostas mais categóricas quando comparado com os pares compostos pelas fricativas. No estudo de Hanzan e Barret (2000), as pesquisadoras também verificaram que os contrastes plosivos “Date-Gate” e “Goat-Coat”, do inglês britânico, foram classificados de forma mais categórica do que os contrastes fricativos “Sue-Shoe” e “Sue-Zoo”. Além disso, elas observaram um aumento na consistência da categorização fonêmica entre as idade de seis e doze anos, sendo que a idade de doze anos apresentou padrão de respostas semelhante aos adultos. Do mesmo modo, Gerrits (2001) verificou que as curvas de classificação dos adultos foram mais íngremes do que as crianças de quatro, de seis e de nove anos de idade. No entanto, a pesquisadora observou que as crianças de nove anos apresentaram valores de inclinação das curvas mais próximos dos adultos (1.19; 1.13 e 1.93, 1.29, respectivamente). Já no estudo de Mayo e Turk (2004), para os contrastes compostos pelas plosivas /de/-/be/, tanto os adultos quanto as crianças não produziram respostas categóricas. As pesquisadoras relatam que esta dificuldade pode ter sido verificada pelo fato das pistas manipuladas para classificação deste contraste terem sido mais fracas e menos informativas, o que acarretou dificuldades na classificação deste par. Para os contrastes /ta/-/da/ e /ti/-/di/, os adultos apresentaram curvas de repostas mais íngremes em comparação com as curvas das crianças de três, cinco e sete anos de idade.

Conforme descrito, além da inclinação das curvas de classificação, outro dado relevante que necessita ser analisado por meio da tabela 14 são os valores da fronteira fonêmica para cada grupo de idade. Para o par mínimo /ʃapa/-/sapa/, as crianças de quatro anos de idade apresentaram valores maiores de separação entre as duas curvas de classificação do que as crianças de sete anos e os adultos (1.17 (507Hz), 0.85 (166Hz) e 0.59 (119Hz), respectivamente). Nittrouer e Miller (1997a) observaram um padrão semelhante de diferença nos valores da fronteira fonêmica para as crianças de quatro anos e as de sete anos e para os adultos (598Hz, 544Hz e 498Hz, respectivamente). No entanto, neste estudo observa-se que as crianças de sete anos apresentaram valores de separação entre as curvas de respostas bem

próximos aos valores das curvas dos adultos. Esta separação maior entre as curvas de classificação das crianças em relação às curvas dos adultos também foi constatada nos estudos de Gerrits (2001), Mayo *et al.* (2003), Mayo e Turk (2004) e Penido (2012).

Para o par de pseudopalavras /faba-/saba/, as crianças mais jovens também apresentaram valores maiores de separação entre as duas curvas de classificação (0.099 (131Hz), 0.039 (95Hz) e 0.036 (106Hz), respectivamente) em comparação com as crianças de sete anos e com os adultos. No entanto, as curvas de respostas deste par apresentaram valores de separação menores em relação às curvas de classificação do par mínimo /fapa-/sapa/.

Conforme esperado, a tabela 14 também mostra que todos os grupos de idade apresentaram os valores da fronteira fonêmica maiores para o contínuo [ʃ] quando comparados com o continuum [s]. Isso ocorreu uma vez que o continuum composto pelo ruído fricativo e pela transição dos formantes apropriada para acompanhar [ʃ] ofereceu aos ouvintes pistas mais favoráveis à resposta /fapa/ no continuum /fapa-/sapa/ e à resposta /faba/ no continuum /faba-/saba/.

Para o par mínimo /bata-/data/, as crianças de quatro anos de idade apresentaram valores maiores de separação entre as curvas de classificação quando comparadas com as crianças de sete anos e com os adultos (1.17 (93Hz), 0.85 (30Hz), 0.59 (23Hz), respectivamente). Esta tendência também foi observada nos estudos de Gerrits (2001). No entanto, Mayo e Turk (2004) não verificaram uma diferença significativa entre as crianças e os adultos na separação das curvas de resposta para o contraste /ti-/di/ e para o contraste /ta-/da/ as crianças mostraram uma separação menor entre as curvas de resposta quando comparadas com os adultos.

Ao comparar os valores de separação entre as curvas de respostas das fricativas e das plosivas deste estudo, observa-se que diante das plosivas as curvas foram mais próximas nos três grupos de idade. Além disso, observa-se que apenas as crianças de sete anos de idade apresentaram os valores da fronteira fonêmica maiores para o contínuo [b] quando comparados com o continuum [d]. Era esperado verificar esta tendência em todos os grupos de idade, uma vez que isto ocorre porque o contínuo composto pelos valores de transição dos formantes vocálicos e pelas amplitudes do *burst* apropriadas para acompanhar [d] oferece aos ouvintes mais pistas favoráveis à resposta /data/ no contínuo /bata-/data/. Isso

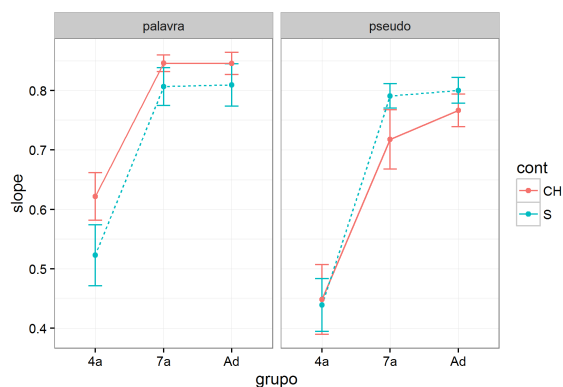
pode ter ocorrido pelo fato da pista de amplitudes do *burst* não ter sido considerada uma pista importante para a classificação deste par.

Com o intuito de ilustrar os dados referentes a fronteira fonêmica e a inclinação das curvas de classificação dispostos na tabela 14, construiu-se os gráficos abaixo, que permitem analisar visualmente a tendência dos três grupos de idade, para estes dois parâmetros, separadamente.

O gráfico 4 revela valores de inclinação inferiores das curvas de classificação das crianças de quatro anos quando comparadas com as crianças de sete anos e com os adultos, tanto para a classificação do par mínimo /*fapa*-/*sapa*/ quanto para a classificação do par de pseudopalavras /*faba*-/*saba*/. Como descrito, isto indica que as respostas das crianças mais jovens foram menos consistentes do que as respostas das crianças de sete anos e dos adultos. Também é possível observar que os valores de inclinação das curvas de respostas foram menores para os três grupos de idade diante da classificação das pseudopalavras. Tal fato, pode sugerir um efeito do estatuto lexical, onde todos os grupos de idade apresentaram respostas menos consistentes para classificar o par de pseudopalavras /*faba*-/*saba*/. No entanto, somente ao final da análise estatística será verificado se este efeito foi significativo entre os pares de palavras e pseudopalavras testados. Nos gráficos abaixo, as barras representam o erro padrão da média, ou seja, o desvio padrão entre as médias das amostras.

GRÁFICO 4

Inclinação (*slope*) das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Fricativas

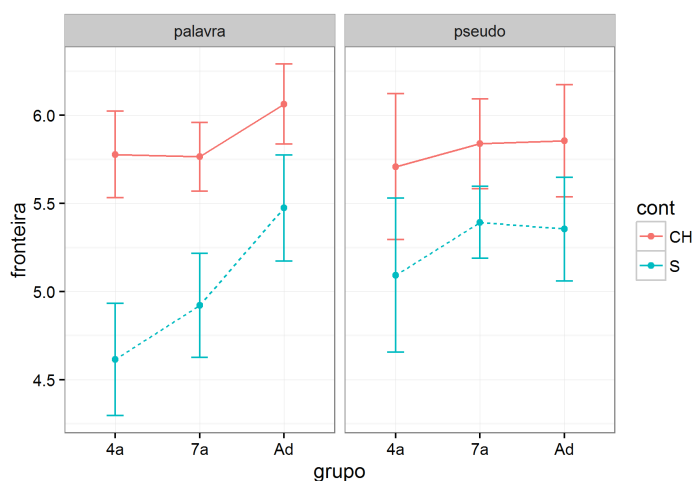


Legenda: eixo X: grupo das crianças de quatro anos (4a), grupo das crianças de sete anos (7a) e grupo de adultos (Ad); eixo y: inclinação das curvas de classificação das palavras /*fapa*-/*sapa*/ (à esquerda) e inclinação das curvas de classificação das pseudopalavras /*faba*-/*saba*/ (à direita); cont CH: curva com transição dos formantes apropriada para acompanhar [ʃ]; cont S: curva com transição dos formantes apropriada para acompanhar [s]. Fonte: Dados da pesquisa

Em relação à fronteira fonêmica, o gráfico abaixo revela que o grupo de idade que apresentou uma separação maior entre as curvas de classificação foi o de quatro anos. Esta separação maior pode ser verificada para os dois pares de palavras compostos pelas fricativas. Porém, uma separação menor entre as curvas de respostas, para os três grupos de idade, foi observada para a classificação das pseudopalavras. Além disso, o gráfico 5 mostra que os valores da fronteira fonêmica para o continuum CH (linhas vermelhas) foram maiores em relação ao continuum S (linhas azuis), diante da classificação das fricativas. A fronteira fonêmica da curva de classificação composta pela transição dos formantes apropriados para acompanhar [ʃ] ofereceu aos ouvintes pistas mais favoráveis à resposta /ʃapa/ ou /ʃapa/, conforme esperado.

GRÁFICO 5

Fronteira fonêmica das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Fricativas



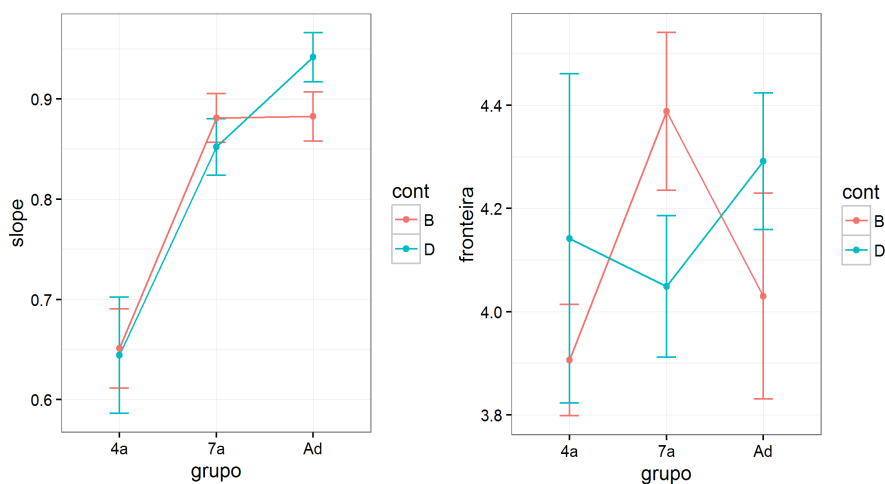
Legenda: eixo X: grupo das crianças de quatro anos (4a), grupo das crianças de sete anos (7a) e grupo de adultos (Ad); eixo y: fronteira fonêmica das curvas de classificação das palavras /ʃapa/-/sapa/ (à esquerda) e fronteira fonêmica das curvas de classificação das pseudopalavras /ʃaba/-/saba/ (à direita); cont CH: curva com transição dos formantes apropriada para acompanhar [ʃ]; cont S: curva com transição dos formantes apropriada para acompanhar [s]

Fonte: Dados da pesquisa

O gráfico 6 mostra que, da mesma forma que ocorreu com as fricativas, a inclinação das curvas de classificação das crianças de quatro anos de idade foi menor em relação às curvas dos adultos e das crianças de sete anos, para a classificação das plosivas. O gráfico adiante facilita a visualização do que foi relatado em relação aos valores da fronteira fonêmica: apenas o grupo de sete anos apresentou valor maior da curva de classificação com amplitudes do *burst* apropriadas para acompanhar [b], conforme esperado. Tal fato, sugere que a pista de amplitudes do *burst* foi pouco ponderada pelos participantes do estudo.

GRÁFICO 6

Fronteira fonêmica e inclinação (*slope*) das curvas de classificação das crianças de quatro e de sete anos e dos adultos – Plosivas



Legenda: eixo X: grupo das crianças de quatro anos (4a), grupo das crianças de sete anos (7a) e grupo de adultos (Ad); eixo y: slope das curvas de classificação das palavras /bata-/data/ (à esquerda), fronteira fonêmica das curvas de classificação das palavras /bata-/data/ (à direita); contB: curva com amplitudes da fase de soltura (*burst*) apropriadas para acompanhar [b] e curva com amplitudes da fase de soltura (*burst*) apropriadas para acompanhar [d]

Fonte: Dados da pesquisa

Por fim, procedeu-se a ANOVA e a MANOVA para as fricativas e o teste “t pareado” para as plosivas. Abaixo, os resultados obtidos para os pares mínimos /fapa-/sapa/ e /bata-/data/ e para o par de pseudopalavras /faba-/saba/ foram descritos separadamente.

Para o par mínimo /fapa-/sapa/ a ANOVA indicou um efeito principal altamente significativo de grupo sobre a “Inclinação” [$F(2.33) = 31.1, p < 0.0001, \eta^2_G = 0.53$]. Houve também um efeito principal significativo do Continuum sobre a “Inclinação” [$F(1.33) = 4.17, p = 0.049, \eta^2_G = 0.049$]. Além disso, não houve interação significativa entre Grupo e Continuum [$F(2.33) = 0.31, p = 0.74, \eta^2_G = 0.007$]. Tais resultados indicam que a inclinação das curvas de classificação diferiu, de forma significativa, entre os grupos de idade testados. O teste HSD de Tukey confirmou esta suposição e mostrou que a diferença entre os valores da “Inclinação” foi altamente significativa entre o grupo de quatro anos de idade e o de sete anos ($p < 0.0001$) e entre o grupo de quatro anos e o de adultos ($p < 0.0001$), mas a diferença foi praticamente inexistente entre o grupo de sete anos e o de adultos ($p = 0.989$), indicando uma mudança do padrão de respostas ao longo do desenvolvimento. Isso confirma a descrição anterior de que, para o par mínimo /fapa-/sapa/, os adultos e as crianças de sete anos apresentaram curvas de classificação mais íngremes

quando comparados com o grupo de quatro anos de idade. Como a “Inclinação” indica o índice de peso que os ouvintes atribuem ao nível de frequência de ruído fricativo, conclui-se que a frequência de ruído foi mais fortemente ponderada pelos adultos e pelas crianças de sete anos de idade do que pelas crianças mais jovens.

Ainda no par de palavras /fapa/-/sapa/, a análise sobre a “Fronteira” fonêmica mostrou que houve um efeito significativo do Continuum sobre a “Fronteira” [$F(1.33) = 39.11$, $p < 0.0001$, $\eta^2_G = 0.195$]. Não houve efeito principal de Grupo sobre “Fronteira” [$F(2.33) = 1.6$, $p = 0.22$, $\eta^2_G = 0.07$]. Também não houve interação significativa entre Grupo e Continuum [$F(2.33) = 1.31$, $p = 0.28$, $\eta^2_G = 0.016$]. O teste HSD de Tukey mostrou que não houve diferenças nem entre o grupo de quatro anos e o de sete anos ($p = 0.92$), nem entre o grupo de quatro anos e o de adultos ($p = 0.218$), nem tampouco entre o grupo de sete anos e o de adultos ($p = 0.396$). Como a fronteira fonêmica é considerada como índice do peso atribuído a pista de transição dos formantes, com estes achados conclui-se que a transição do F2 e do F3 não foi mais fortemente ponderada por nenhum dos grupos de idade. A fronteira fonêmica também foi analisada por meio da utilização de testes de diferença entre os continua (CH e S), dentro de cada grupo de idade, corrigido para medidas repetidas (método de Bonferroni). Os resultados revelaram que a diferença entre os continua CH e S foi significativa para as crianças de quatro ($p = 0.0089$) e para as crianças de sete anos ($p = 0.022$), mas não foi para os adultos ($p = 0.12$). Esta diferença entre os continua CH e S pode ser verificada no gráfico 4. Mas, ainda assim não houve diferença significativa entre os grupos para a fronteira fonêmica, conforme descrito.

Diante do par de pseudopalavras /faba/-/saba/, a ANOVA mostrou que, ao contrário do que ocorreu com o par mínimo /fapa/-/sapa/, não houve efeito do Continuum sobre a “Inclinação” [$F(1.33) = 1.7$, $p = 0.2$, $\eta^2_G = 0.017$]. Houve um efeito principal altamente significativo de Grupo sobre a “Inclinação” [$F(2.33) = 34$, $p < 0.0001$, $\eta^2_G = 0.58$]. Além disso, não verificou-se interação significativa entre Grupo e Continuum [$F(2.33) = 0.92$, $p = 0.41$, $\eta^2_G = 0.02$]. Assim como quando o contexto era o de palavras, o teste HSD de Tukey mostrou que a diferença foi altamente significativa entre o grupo de quatro anos de idade e o de sete anos ($p < 0.0001$) e entre o grupo de quatro anos e o de adultos ($p < 0.0001$), mas a diferença foi praticamente inexistente entre o grupo de sete anos e o de adultos ($p = 0.795$), o que também indica uma modificação do padrão de respostas ao longo do desenvolvimento. Da mesma forma da análise direcionada ao par mínimo /fapa/-/sapa/, com o par de

pseudopalavras, os resultados confirmam a descrição realizada na tabela 14, de que as crianças de quatro anos apresentaram curvas de classificação menos íngremes quando comparadas com o grupo de sete anos e o de adultos. Desta forma, conclui-se que, para a classificação do par de pseudopalavras, as crianças de quatro anos também atribuíram menor peso à pista de frequência do ruído fricativo quando comparadas com os demais grupos de idade.

Os resultados da ANOVA direcionados à “Fronteira” para as pseudopalavras revelaram que houve um efeito principal significativo do Continuum sobre a “Fronteira” [$F(1.33) = 13$, $p = 0.001$, $\eta^2_G = 0.056$], mas não houve efeito de Grupo sobre a “Fronteira” [$F(2.33) = 0.09$, $p = 0.92$, $\eta^2_G = 0.004$]. Além disso, observou-se que também não houve interação significativa entre Grupo e Continuum [$F(2.33) = 0.12$, $p = 0.88$, $\eta^2_G = 0.001$]. O teste HSD de Tukey não verificou diferenças significativas entre a “Fronteira” nem entre o grupo de quatro anos e o de sete anos ($p = 0.934$), nem entre o grupo de quatro anos e o de adultos ($p = 0.927$), nem tampouco entre o grupo de sete anos e o de adultos ($p = 1$). Por meio desta análise torna-se possível inferir que a pista de transição dos formantes não foi ponderada de forma diferente entre os grupos de idade. O teste de diferença entre os continua (CH e S), dentro de cada grupo de idade, corrigido para medidas repetidas (método de Bonferroni), revelou que não houve diferença significativa na fronteira fonêmica no grupo de quatro anos de idade ($p = 0.31$), nem no grupo de sete anos ($p = 0.17$) e nem no grupo de adultos ($p = 0.25$).

Para o par mínimo /bata-/data/, os resultados do teste “t pareado” direcionado à “Inclinação” mostraram um efeito de Grupo, no mesmo padrão descrito para as fricativas. Não houve diferença entre os grupos de adultos e de crianças de sete anos [$t = 1.30$, $p = 0.4$, $r = 0.22$], mas houve uma diferença altamente significativa já entre quatro e sete anos [$t = 6.25$, $p < 0.0001$, $r = 0.74$]. Consequentemente, houve também uma diferença altamente significativa entre os adultos e as crianças de quatro anos de idade [$t = 7.55$, $p < 0.0001$, $r = 0.8$]. Este achado confirma o que está exposto na tabela 14, que mostra que as crianças de quatro anos apresentaram curvas de classificação menos íngremes do que as crianças de sete anos e os adultos. No caso das plosivas, como a “Inclinação” indica o peso atribuído à pista de transição dos formantes vocálicos, conclui-se que esta pista foi mais ponderada pelos adultos e pelas crianças de sete anos de idade quando comparados com as crianças mais jovens.

Os resultados do teste “t pareado” voltados para a “Fronteira” revelaram que, assim como as fricativas, não houve qualquer diferença significativa entre os Grupos em nenhum Continuum. A diferença entre crianças de quatro e de sete anos é maior [$t = 1.03$, $p = 0.56$, $r = 0.18$] do que a diferença entre adultos e as crianças de quatro anos [$t = 0.73$, $p = 0.75$, $r = 0.13$] ou do que a diferença entre os adultos e as crianças de sete anos [$t = -0.30$, $p = 0.95$, $r = 0.05$]. Os achados para a “Fronteira” fonêmica indicam que os três grupos de idade não foram influenciados pela pista estática de amplitudes do *burst* para classificação do par mínimo composto pelas plosivas, apesar de ter-se observado um padrão de desenvolvimento entre os grupos etários inspecionados. Assim como nas pseudopalavras, o teste de diferença entre os continua (B e D), dentro de cada grupo de idade, corrigido para medidas repetidas (método de Bonferroni) revelou que não houve diferença significativa na fronteira fonêmica no grupo de quatro anos de idade ($p = 0.61$), nem no grupo de sete anos ($p = 0.19$) e nem no grupo de adultos ($p = 0.38$).

Por fim, os resultados da análise MANOVA, realizada para as fricativas, mostrou um efeito significativo da idade [$F(1.57) = 52.1$, $p < 0.0001$] e do estatuto lexical [$F(0.94) = 10.1$, $p = 0.0032$] sobre a “Inclinação”. Este achado aponta para duas conclusões: conforme descrito, os adultos apresentaram curvas de classificação mais íngremes do que as crianças para a classificação das fricativas e os três grupos de idade modificaram o padrão de respostas diante do par mínimo /*ʃapa*-/*sapa*/ e do par de pseudopalavras /*ʃaba*-/*saba*/, ou seja, o estatuto lexical gerou diferenças no padrão de respostas dos participantes. Além disso, os participantes apresentaram um melhor desempenho para classificar o par mínimo /*ʃapa*-/*sapa*/. No entanto, os três grupos de idade não mostraram o mesmo padrão de respostas para a classificação do par de pseudopalavras, uma vez que as crianças mais jovens demonstraram respostas menos categóricas do que os adultos. Isto indica que as crianças mais jovens tiveram maior dificuldade para classificar o par de pseudopalavras quando comparadas com os adultos. Os resultados para fronteira fonêmica não revelaram nenhum efeito significativo da idade [$F(1.24) = 0.61$, $p = 0.55$] ou do estatuto lexical [$F(0.45) = 0.75$, $p = 0.39$]. Tal fato indica que não houve uma diferença significativa nos resultados de fronteira fonêmica entre o par mínimo /*ʃapa*-/*sapa*/ e o par de pseudopalavras /*ʃaba*-/*saba*/. Cabe ressaltar, que tanto as crianças de quatro anos de idade quanto as crianças de sete anos receberam treinamento intensivo, criterioso e prévio à tarefa de ponderação de pistas (treino com nomeação das figuras e treino com os melhores exemplares separadamente). Desta

forma, a influência maior do estatuto lexical sobre as crianças mais jovens não pode ter sido decorrente do declínio da memória fonológica e da memória visual.

Os dados descritos são suficientes para responder as seguintes questões do estudo de ponderação de pistas: (1) As crianças mais jovens, que possuem pouca experiência na língua, pesaram as pistas auditivas disponíveis no sinal acústico de fala de forma diferente dos adultos? (2) As crianças e os adultos modificaram o padrão de respostas diante do par mínimo /ʃapa-/sapa/ e do par de pseudopalavras /ʃaba-/saba/ devido a frequência de uso desses pares e a experiência linguística? (3) As crianças de quatro anos de idade ponderaram mais a pista dinâmica do que os adultos para classificar os pares de palavras compostos pelas plosivas e pelas fricativas? (4) As crianças do grupo de quatro anos de idade apresentaram um padrão de respostas mais consistente diante do par de palavras /ʃapa-/sapa/ do que do par de pseudopalavras /ʃaba-/saba/?

Os resultados alcançados tanto para o par mínimo /ʃapa-/sapa/ quanto para o par de pseudopalavras /ʃaba-/saba/ corroboram parcialmente com os achados encontrados na literatura. Diante do par mínimo e do par de pseudopalavras compostos pelas fricativas observou-se que os adultos e as crianças de sete anos ponderaram mais fortemente a pista de frequência do ruído fricativo do que as crianças de quatro anos de idade. Esse padrão de resposta também foi verificado nos estudos de Nittrouer e Studdert-Kennedy (1987), Nittrouer e Miller (1996), Hanzan e Barret (2000), Gerrits (2001) e Penido (2012). No estudo de Mayo e Turk (2004), as pesquisadoras não verificaram diferenças significativas entre os adultos e as crianças para a “Inclinação” das curvas de classificação, o que indicou que nenhum grupo pesou mais fortemente a pista de frequência do ruído fricativo para classificar o par de palavras do inglês britânico /ʃaI-/saI/. Para este par de palavras, Mayo e Turk (2004) observaram uma grande separação entre as curvas de classificação das crianças em relação aos adultos, o que mostrou que este grupo foi mais influenciado pela pista de transição dos formantes quando comparado com os adultos. Tal fato não foi observado neste estudo, uma vez que para a classificação dos pares de palavras /ʃapa-/sapa/ e /ʃaba-/saba/, os três grupos de idade utilizaram a pista de transição dos formantes da mesma maneira. Desta forma, este achado não corrobora com os resultados descritos nos estudos supramencionados.

Ainda direcionando a análise para as fricativas, observou-se que as crianças de quatro anos apresentaram curvas de classificação menos íngremes em relação às crianças de sete anos e aos adultos o que indica uma menor consistência no padrão de respostas das crianças mais jovens. Este resultado está de acordo com os resultados descritos nos estudos de Nittrouer e Studdert-Kennedy (1987), Nittrouer e Miller (1996), Hanzan e Barret (2000), Gerrits (2001) e Mayo e Turk (2004). Verificou-se também que as crianças de quatro anos de idade não possuem as categorias fonológicas tão bem definidas e foram mais inconsistentes em decidir o ponto de articulação das fricativas [ʃ] e [s] do que os demais grupos. Já as crianças de sete anos apresentaram um padrão de respostas próximos aos dos adultos, com curvas de classificação íngremes e, conseqüentemente, respostas mais categóricas. Ou seja, estes grupos dividiram o contínuo fricativo de maneira mais abrupta do que as crianças de quatro anos. Este achado também foi observado no estudo de Gerrits (2001) com os contrastes fricativos do holandês. A pesquisadora verificou que aos seis anos de idade o esquema de ponderação de pistas das crianças foi similar ao dos adultos.

Outro achado relevante obtido com este estudo relaciona-se a influência do estatuto lexical no padrão de respostas dos participantes. Os resultados revelaram que o estatuto lexical gerou modificação no padrão de respostas dos participantes e os três grupos de idade apresentaram respostas menos categóricas para a classificação do par de pseudopalavras /ʃaba/-/saba/. No entanto, o grupo de quatro anos de idade foi o mais influenciado pelo estatuto lexical quando comparado com os demais grupos. Ora, se a criança adquire e os adultos armazenam segmentos, conforme postulam os modelos fonológicos tradicionais, os três grupos de idade deveriam ter apresentado o mesmo padrão de respostas diante do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/, uma vez que estas palavras possuem os mesmos segmentos em posição inicial. Em contrapartida, os resultados descritos vão ao encontro dos pressupostos das Teorias Fonológicas Multirrepresentacionais, mais especificamente com a Teoria de Exemplos, que consideram que o falante armazena em seu léxico mental os exemplos atestados em sua experiência e a frequência de um exemplar desempenha um papel importante na organização das representações fonológicas. Conforme a Teoria, o léxico desempenha o papel central na aquisição da língua e as representações mentais dos componentes fonológicos é composto de múltiplos exemplos. Assim, o falante armazena em seu léxico mental exemplos atestados em sua experiência e os exemplos mais produzidos tem um espaço maior nas representações mentais. De acordo com Pierrehumbert (2003), a criança aprende os detalhes fonéticos específicos e finos de

uma língua particular e a aquisição da língua envolve uma atualização constante das distribuições de probabilidade. As generalizações baseadas nestas distribuições são dinâmicas e submetidas a mudanças de acordo com o uso dos exemplares e tornam-se mais robustas à medida que as amostras desses exemplares aumentam. Além disso, a palavra e não o segmento, é uma unidade de análise básica e as categorias emergem a partir da memória da palavra inteira (ou de enunciados maiores) que são armazenados no léxico mental (COLE, 2009). A ativação de uma palavra também ativa outras foneticamente semelhantes.

Ainda que os três grupos de idade não possuam como exemplares as pseudopalavras /faba/-/saba/ e ainda que estas pseudopalavras não estejam armazenadas no léxico mental dos participantes testados, os resultados revelaram que as crianças de quatro anos demonstraram uma diferença mais evidente entre os resultados de classificação das palavras e das pseudopalavras, com respostas mais categóricas nas palavras. Este achado mostra um efeito maior do estatuto lexical na classificação das crianças mais jovens, que possuem menos experiência com a língua e possuem um léxico mais limitado. Em contrapartida, as crianças de sete anos mostraram um padrão de classificação semelhante aos adultos e os dois grupos, que possuem uma experiência maior com a língua e um léxico mais amplo, foram menos influenciados pelo estatuto lexical. Se as categorias emergem a partir da memória da palavra e no PB não temos a representação e o armazenamento das pseudopalavras /faba/-/saba/ no léxico, o efeito do estatuto lexical verificado pode ser decorrente da experiência com o uso da língua nativa e da ativação de palavras foneticamente semelhantes em um léxico mais amplo (crianças de sete anos e adultos). Ou seja, na medida em que se ganha experiência com o uso da língua e diante da ampliação dos exemplares dispostos no léxico, os falantes se tornam mais flexíveis para lidar com a informação acústica disponível no sinal de fala, que não fazem parte do léxico mental.

Em relação às plosivas, observou-se que as crianças de quatro anos de idade apresentaram curvas de classificação menos íngremes do que as crianças de sete anos e os adultos. Assim, as crianças mais jovens utilizaram menos intensamente a pista dinâmica de transição dos formantes em relação aos demais grupos de idade. Ohde e Haley (1997) também verificaram que nem sempre as crianças mais jovens utilizaram a pista dinâmica para classificar os contrastes do inglês americano compostos pelas plosivas [b, d, g] e pelas vogais [i, u, a], conforme é postulado pela hipótese DWS. Os achados deste estudo e do estudo de Ohde e

Halley (1997) não corroboram com o que Gerrits (2001) verificou ao investigar a ponderação de pistas auditivas utilizando o par de palavras composto pelas plosivas do holandês “pop-kop”. A pesquisadora observou que apenas as crianças mais jovens atribuíram mais peso à pista de transição dos formantes. Os demais grupos de idade utilizaram tanto a pista de transição dos formantes quanto a pista de amplitudes do *burst*.

Neste estudo não verificou-se diferenças significativas entre os Grupos na “Fronteira” fonêmica das curvas de classificação do par mínimo /bata/-/data/. Conforme descrito, isto indica que os grupos de idade testados não foram influenciados pela pista de amplitudes do *burst* e, conseqüentemente, fizeram pouco uso desta pista para classificar o par mínimo composto pelas plosivas. Gerrits (2001) também observou que apesar das crianças mais jovens e os adultos utilizarem tanto a pista de transição dos formantes quanto a pista de amplitudes do *burst* para classificação das plosivas, os adultos mostraram um amplo efeito da pista de transição dos formantes, o que contradiz os pressupostos da hipótese DWS. Tal fato gerou diferenças nas estratégias de ponderação de pistas dos participantes do estudo de Gerrits (2001) diante dos pares de palavras compostos pelas fricativas e pelas plosivas. Assim, a pesquisadora relata que a pista de amplitudes do *burst* possui uma duração menor (30ms) quando comparada com a pista de frequência do ruído fricativo (190ms), o que poderia proporcionar ao ouvinte uma informação fonética mais limitada. Esta questão da informatividade das pistas poderia explicar o motivo pelo qual os sujeitos deste estudo tenham utilizado pouco a pista de amplitudes do *burst*, que apresentou uma duração curta de 15ms.

O modelo DWS (*Developmental Weighting Shift*), proposto por Nittrouer, Manning e Meyer (1993) supõe que as crianças mais jovens processam a fala de forma diferente dos adultos, de uma maneira mais global e tal fato gera diferenças nas estratégias de ponderação de pistas auditivas entre estes dois grupos. Ainda de acordo com o modelo, as crianças fornecem uma preferência maior pela pista dinâmica do que os adultos. Os resultados deste estudo contradizem parcialmente o modelo DWS, uma vez que as crianças mais jovens não ponderaram mais fortemente a pista dinâmica em relação às crianças de sete anos e aos adultos. Este achado também foi observado no estudo de Penido (2012), em que tanto as crianças quanto os adultos foram influenciados pela pista de frequência do ruído fricativo para classificar o par mínimo /aʃaR/-/asaR/. A pesquisadora sugeriu que nesta língua existem indícios de que esta pista fornece informações fonéticas mais favoráveis para a classificação

do contraste /ʃ-vogal/-/s-vogal/ desde muito cedo no desenvolvimento da criança. Pode-se atribuir esta característica à língua, o que aponta para o fato de que as pistas são dependentes da língua. Assim, para que a mudança no desenvolvimento da ponderação de pistas entre crianças e adultos seja considerada um fenômeno universal entre as línguas, conforme propõem Nittrouer e colaboradores no modelo DWS, ela deveria ocorrer em qualquer par de palavra testado em qualquer língua.

Analisando os resultados dos estudos de ponderação de pistas auditivas descritos em conjunto com os achados obtidos nesta tese supõe-se que a ponderação de pistas de crianças e adultos, ou seja, a importância que os ouvintes atribuem a uma determinada pista, parece modificar de acordo com (1) a exposição da criança em sua língua nativa, (2) o contraste fonológico testado (fricativas, plosivas, nasais, dentre outros) e (3) as características acústicas da língua em que a criança é exposta. De fato, em grande parte dos estudos, observa-se que as crianças mais jovens possuem respostas menos consistentes e curvas de classificação menos íngremes. Com o aumento da idade e com a exposição em uma determinada língua, espera-se que os ouvintes apresentem um padrão de respostas mais consistente, com categorias fonológicas bem definidas, conforme foi verificado neste estudo. No entanto, esperar um padrão de ponderação de pistas, onde as crianças mais jovens sempre pesam mais fortemente as pistas dinâmicas do que os adultos para classificar os contrastes é inapropriado porque as pistas auditivas variam entre as línguas. Ou seja, uma pista pode ser mais informativa em um determinada língua quando comparada com outra. Por exemplo, na seção 3.5.3, na qual analisou-se os valores de transição dos formantes da vogal [a] seguinte às fricativas [ʃ] e [s], foi possível observar uma variação maior dos valores do F2 (inicial e estável) da vogal [a] seguinte a fricativa [ʃ] quando comparado com a fricativa [s]. Esta tendência também é observada em outras línguas, como no inglês. Porém, no presente estudo, verificou-se uma menor variação dos valores de F2 da vogal [a] seguinte a fricativa [s] quando comparado com o inglês. Isto pode apontar para o fato de que a transição dos formantes da vogal [a] seguinte as fricativas [ʃ] e [s] do PB pode ser menos informativa quando comparada com o inglês. Aqui se estabelece uma relação entre a fonética e a fonologia e entre a percepção e a produção de fala, em que diferenças na produção dos sons pode gerar diferenças na informatividade das pistas auditivas e conseqüentemente na percepção. Se as diferenças articulatórias dos falantes de uma língua geram diferenças em relação à informatividade das pistas provavelmente elas não serão ponderadas da mesma maneira em todas as línguas. Desta forma, não quer dizer que as

crianças mais jovens sempre pesarão mais fortemente as pistas dinâmicas em relação aos adultos para classificar os contrastes, uma vez que a pista estática pode estar mais disponível.

Conforme descrito na seção 2.4.2, não foi objetivo deste estudo propor uma hipótese que consiga explicar todas as diferenças na ponderação de pistas auditivas entre crianças e adultos. Porém, verificou-se que os achados desta investigação não sustentaram a hipótese DWS em sua totalidade. Os resultados alcançados poderiam sugerir que na medida em que as crianças são expostas à sua língua nativa elas aprendem a lidar com as pistas auditivas disponíveis no sinal de fala. Ou seja, as crianças aprendem a rastrear as distribuições das pistas e a eleger quais pistas são mais informativas quando elas necessitam discriminar determinados contrastes.

Em suma, o padrão de respostas das crianças de quatro anos diante das fricativas e das plosivas neste estudo foi diferente do padrão de respostas das crianças de sete anos e dos adultos, faixas etárias que apresentaram estratégias de ponderação de pistas semelhantes. Desta forma, este resultado poderá ir ao encontro com a proposta do aprendizado estatístico¹², que sugere o cérebro humano mantém o controle de muitas características dos sons que ouvimos ao longo do tempo (pistas auditivas) e determina as probabilidades ou as frequências de certas características. No caso do desenvolvimento da percepção da fala, as crianças aprendem a determinar as estatísticas de distribuição de um conjunto de pistas auditivas e, além disso, aprendem a determinar se estas pistas auditivas são relevantes para distinguir os contrastes fonológicos de sua língua nativa. Esta linha de pensamento, poderia explicar o motivo pelo qual as crianças mais jovens apresentaram um padrão de resposta mais inconsistente e diferenciado em relação aos demais sujeitos. Ou seja, verificou-se uma mudança nas estratégias de ponderação de pistas, das crianças de quatro anos para as de sete anos de idade e para os adultos, provavelmente pelo ganho de experiência com o PB.

A realização de mais estudos que investiguem como as múltiplas pistas auditivas são aprendidas ao longo do desenvolvimento linguístico e se as pistas mais informativas, disponíveis no sinal acústico de fala, são mais fortemente ponderadas pelos falantes nativos

¹² Ver Maye, Weiss e Aslin (2008); Romberg e Saffran (2010); e Lany e Saffran (2013) para um conhecimento mais aprofundado em relação ao aprendizado estatístico na infância.

do PB torna-se necessário. Como direcionamento para uma investigação futura, levanto uma variável adicional que poderia contribuir para as diferenças na categorização dos dados entre as crianças jovens em comparação com as crianças mais velhas e com os adultos: a escrita, uma vez que esta também impõe categorias. Desta forma, os grupos alfabetizados teriam maior facilidade de categorizar os dados do que as crianças de quatro anos de idade? A forma de testar esta hipótese seria submeter sujeitos iletrados ao mesmo experimento de ponderação de pistas desenvolvido aqui. Além disso, como seria a ponderação de pistas auditivas em crianças que apresentam desvio fonológico? Estes estudos podem fornecer contribuições importantes para o melhor entendimento da relação entre o desenvolvimento da criança e sua exposição na língua nativa e a modificação das suas estratégias de ponderação de pistas.

7 CONSIDERAÇÕES FINAIS

Esta tese apresentou como objetivo principal investigar o desenvolvimento da ponderação de pistas auditivas para a percepção das fricativas coronais não vozeadas ([ʃ] e [s]) e das plosivas bilabial e alveolar ([b] e [d]) por crianças e adultos falantes nativos do português brasileiro. Assim, realizou-se dois estudos prévios para posterior execução do experimento final de ponderação de pistas. O primeiro estudo foi desenvolvido com o intuito de analisar as propriedades acústicas dos pares de palavras /ʃapa-/sapa/, /ʃaba-/saba/ e /bata-/data/, que foram sintetizados e utilizados no experimento final. Já o segundo estudo envolveu a análise, por meio de uma tarefa de classificação, da qualidade dos estímulos sintetizados.

Com a realização do primeiro estudo verificou-se que grande parte dos resultados obtidos foi coerente com os achados descritos na literatura, tais como, os valores do centro de gravidade da fricativa [s] foram maiores do que os valores da fricativa [ʃ]; os picos de amplitude da fricativa [s] foram mais proeminentes em regiões mais altas de frequência quando comparada com a fricativa [ʃ]; a plosiva alveolar [d] apresentou picos de energia mais intensos em regiões mais altas de frequência quando comparada com a plosiva bilabial [b]; a variação nos valores inicial e estável do F2 foi maior para a vogal [a] seguinte a fricativa [ʃ] quando comparada com a fricativa [s] e a variação nos valores inicial e estável do F2 e do F3 foi maior para a vogal [a] seguinte a plosiva [d] quando comparada com a plosiva [b]. Porém, observou-se que esta variação nos valores do F2 (inicial e estável) da vogal [a] seguinte a fricativa [s], verificada neste estudo (1311 Hz-1324Hz), foi menor quando comparada com os estudos desenvolvidos no inglês americano (1250Hz-1130Hz).

Após a realização do primeiro estudo, os estímulos foram sintetizados com base nas propriedades acústicas dos sons produzidos por um dos informantes. Em seguida, realizou-se o segundo estudo, que objetivou analisar, por meio de uma tarefa de classificação, a qualidade dos estímulos sintetizados compostos pelas palavras /ʃapa-/sapa/, /ʃaba-/saba/ e /bata-/data/. Os resultados desta investigação mostraram que os estímulos sintetizados apresentaram boa qualidade, uma vez que os participantes obtiveram curvas de classificação íngremes e, conseqüentemente, produziram respostas categóricas.

Por fim, desenvolveu-se o experimento final com o intuito de investigar o desenvolvimento da ponderação de pistas auditivas, para a percepção das fricativas coronais não vozeadas e

das plosivas vozeadas [b] e [d] por crianças e adultos, falantes nativos do português brasileiro e as hipóteses inicialmente levantadas foram, de um modo geral, comprovadas: as crianças mais jovens, que possuem pouca experiência linguística, pesaram as pistas auditivas disponíveis no sinal acústico de fala de forma diferente dos adultos, foram menos estáveis e produziram respostas mais inconsistentes. Além disso, observou-se que as crianças da faixa etária de quatro anos de idade, assim como os demais grupos de idade testados, apresentaram modificação do padrão de resposta diante do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/. No entanto, o grupo de quatro anos de idade foi mais influenciado pelo estatuto lexical quando comparado com os demais grupos.

Os resultados do experimento de ponderação de pistas permitiram concluir que:

- 1) As crianças de quatro anos de idade apresentaram curvas de classificação menos íngremes do que as crianças de sete anos e os adultos diante dos pares de palavras compostos pelas fricativas e pelas plosivas. Este achado sugere que as crianças mais jovens apresentaram um padrão de respostas mais inconsistente e não dividiram o continuum de forma abrupta conforme verificou-se nas respostas dos adultos e das crianças de sete anos.
- 2) As crianças de quatro anos de idade nem sempre ponderaram mais fortemente a pista dinâmica em relação aos demais grupos de idade. Diante das palavras compostas pelas fricativas, a pista dinâmica de transição dos formantes foi ponderada de forma semelhante nos três grupos de idade. Diante das palavras compostas pelas plosivas, esta pista foi mais ponderada pelas crianças de sete anos e pelos adultos para a classificação do par mínimo /bata/-/data/. Tais achados não estão de acordo com as premissas do modelo DWS.
- 3) Houve uma modificação significativa do padrão de respostas dos três grupos de idade entre a classificação do par mínimo /ʃapa/-/sapa/ e do par de pseudopalavras /ʃaba/-/saba/, indicando uma influência do estatuto lexical nas respostas dos participantes. No entanto, as crianças mais jovens foram as mais influenciadas pelo estatuto lexical.
- 4) Nos três pares de palavras testados, as crianças de sete anos de idade apresentaram estratégias de ponderação de pistas auditivas diferentes em relação as crianças de quatro anos e semelhantes em relação aos adultos, o que indica uma modificação do padrão de respostas ao longo do desenvolvimento.

Diante dos resultados obtidos, concluiu-se que nem sempre as crianças mais jovens pesam mais fortemente a pista dinâmica e global do que os adultos quando elas necessitam classificar os pares de palavras, conforme o modelo DWS postula. Concluiu-se também que as estratégias de ponderação de pistas parecem modificar de acordo com: a) a experiência da criança com sua língua nativa; b) o contraste fonológico testado e c) a informatividade que as pistas auditivas fornecem aos ouvintes para a classificação dos contrastes.

8 REFERÊNCIAS BIBLIOGRÁFICAS

- ALVES, M.A. (2015). *Estudo dos parâmetros acústicos relacionados à produção das plosivas do Português Brasileiro na fala adulta: análise acústico-quantitativa*. (Doutorado) – Universidade Federal de Santa Catarina.
- ASLIN, R.N.; WERKER, J.F.; MORGAN, J.L. (2002). Innate phonetic boundaries revisited. *Journal of the Acoustical Society of America*, 112, pp. 1257-60.
- BAKEMAN, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, pp. 379-84.
- BARBOSA, P.A. (1999). Revelar a estrutura rítmica de uma língua construindo máquina falantes: pela integração da ciência e tecnologia de fala. In: SCARPA, E. (Org.). *Estudos de prosódia*. Campinas: Unicamp, pp. 21-52.
- BARBOSA, P.A.; MADUREIRA, S. (2015). Manual de fonética acústica experimental: aplicações a dados do português. São Paulo: Cortez.
- BARTON, D. (1980). Phonemic perception in children. In: YENI-KOMSHIAN, G.H.; KAVANAGH, J.F.; FERGUSON, C.A. (Orgs.). *Child Phonology: Perception*. New York: Academic Press.
- BEHLAU, M.S.; PONTES, P.A.L.; GANANÇA, M.M.; TOSI, O. (1988). Análise espectrográfica de formantes das vogais do português brasileiro. *Acta Awho*, pp. 67-73.
- BEST, C.T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In: GOODMAN, J.C.; NUSBAUM, H.C. *The developmental of speech perception: the transition from speech sounds to spoken words*. Cambridge, MA: MIT Press, pp. 167-224.
- BOERSMA, P. (2009). Cue constraints and their interactions in phonological perception and production. In: BOERSMA, P.; HAMANN, S. *Phonology in perception*. Berlin: Mouton de Gruyter, pp. 55-110.
- BOERSMA, P., WEENINK, D. (2011). *Praat: doing phonetics by computer*. Versão 5.2.25. Disponível em: <www.praat.org>.
- BONILHA, G.F.G. (2003). Teoria da otimidade. In: MATZENAUER, C.L.B; BONILHA, G.F.G. *Aquisição da fonologia e teoria da otimidade*. Pelotas: EDUCAT, pp. 14-24.
- BORZONE DE MANRIQUE, A.M.; MASSONE, M.I. (1981). Acoustic analysis and perception of speech fricative consonants. *Journal of the Acoustical Society of America*, 69(4), pp. 1145-53.
- BURNHAM, D.K.; EARNSHAW, L.J.; CLARK, J.E. (1991). Development of native and non- native bilabial stops: infants, children and adults. *Journal of Child Language*, 18, pp. 231-60.
- BYBEE, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

- CÂMARA, J.M. (1970). *Estrutura da língua portuguesa*. Petrópolis: Vozes.
- CAMPOS, C.A., OLIVEIRA, J.A., ENDO, L., BENTO, R., PIGNATARI, S., WECKX, L.L.M. (1999). Consenso sobre otite média. *Revista Brasileira de Otorrinolaringologia*, 65 Supl 8: 14-7.
- CARTERETTE, E.C.; JONES, M.V. (1974). *Informal speech*. Berkeley, University of California Press, p.11-48.
- COLE, J. (2009). Emergent feature structures: harmony systems in exemplar models of phonology. *Language Sciences*, 31, pp. 144-60.
- CORREA, L.M.S. (1999). Aquisição da linguagem: uma retrospectiva dos últimos trinta anos. *D.E.L.T.A.*, 15, pp. 339-83.
- CHOMSKY, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- CROWTHER, C.S.; MANN, V. (1992). Native language factors affecting use of vocalic cues to final consonant voicing. *Journal of the Acoustical Society of America*, 92, pp. 711-22.
- DAVIS, H.; SILVERMAN, S.R. (1970). Auditory Test Hearing Aids. In: DAVIS, H.; SILVERMAN, S.R. *Hearing and Deafness*. Holt: Rinehart and Winston.
- DELATTRE, P. (1951). The physiological interpretation of sound spectrogram. *The modern language association of America*, 66, pp. 864-74.
- DIEHL, R.; LOTTO, A.; HOLT, L. (2004). Speech Perception. *Annual Reviews of Psychology*, 55, pp. 149-79.
- DORMAN, M.F.; STUDDERT-KENNEDY, M.; RAPHAEL, L.J. (1977). Stop-consonant recognition: release bursts and formant transitions as a functionally equivalent, context-dependent cues. *Perception and Psychophysics*, 22, pp. 109-22.
- EIMAS, P.D. (1985). The equivalence of cues in the perception of speech by infants. *Infant Behavior and Development*, 8, pp. 125-38.
- EIMAS, P.D.; MILLER, J.L. (1992). Organization in the perception of speech by young infants. *Psychological Science*, 3, pp. 340-45.
- EIMAS, P.D.; SIQUELAND, F.R.; JUSCZYK, P.; VIGORITO, J. (1971). Speech perception in infants. *Science*, 171, pp. 303-6.
- ESCUDERO, P.; BOERSMA, P. (2001). Modelling the perceptual developmental of phonological contrasts with optimality theory and the gradual learning algorithm. *Proceedings of the 25th Penn Linguistics Colloquium*, pp. 1-14.
- EISENBERG, L.S.; SHANNON, R.V.; SCHAEFER, A.; WYGONSKI, J.; BOOTHROYD, A. (2000). Speech recognition with reduced spectral cues as a function of age. *Journal of the Acoustical Society of America*, 107, pp. 2704-10.

- FANT, G. (1960). *Acoustic theory of speech production: with calculations based on X-Ray studies of Russian articulations*. Amsterdam: Mouton e Co.
- FERNALD, A.; SIMON, T. (1984). Expanded imitation contours in mother's speech to newborns. *Developmental Psychology*, 20, pp. 104-13.
- FIELD, A. (2009). *Descobrendo a estatística usando o SPSS*. Porto Alegre: Artmed.
- FLAVELL, J.H.; MILLER, P.H.; MILLER, S.A. (1999). *Desenvolvimento cognitivo*. São Paulo: Artmed. Cap. 7: *A Linguagem*.
- GERRITS, E. (2001). *The categorization of speech sounds by adults and children. A study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues*. Tese (Doutorado) – Utrecht University, The Netherlands, LOT series.
- GORDON, M.; BARTHMAIER, P.; SANDS, K. (2002). A cross linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32, pp. 141-74.
- GOUDBEEK, M. (2006). *The acquisition of auditory categories*. Tese (Doutorado) – Nijmegen University.
- GRIESER, D.L.; KUHL, P.K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24, pp. 14-20.
- HALLE, M.; HUGHES, G.W.; RADLEY, J.P.A. (1957). Acoustic properties of stop consonants. *Journal of the Acoustical Society of America*, 29, pp. 107-16.
- HANZAN, V.; BARRET, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics*, 28, pp. 377-96.
- HARNAD, S. (1987). *Categorical Perception*. New York: Cambridge University Press.
- HAUPT, C. (2007). As fricativas [s], [z], [ʃ] e [ʒ] do português brasileiro. In: *Estudos Linguísticos XXXVI(1)*. Florianópolis: UFSC.
- HEINZ, J.M.; STEVENS, K.N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America*, 33, pp. 589-96.
- HIXON, T. (1966). Turbulent noise sources for speech. *Folia phoniat.*, 18, pp. 168-82.
- HOLT, L.L.; LOTTO, A.J. (2008). Speech perception within an auditory cognitive science framework. *Current Directions in Psychological Science*, 17, pp. 42-46.
- HUANG, X.; GERHARDT, K.J.; ABRAMS, R.M.; ANTONELLI, P. (1997). Temporary threshold shifts induced by low-pass and high-pass filtered noises in fetal sheep in utero. *Hearing Research*, 113, pp. 173-81.
- ISTRE, G. (1983). *A Fonética Acústica*. In _____. *Fonologia transformacional e natural: uma introdução crítica*. Florianópolis: pp. 37-72.

- JENKINS, J.J. (1992). The organization and reorganization of categories: the case of speech perception. In: PICK, H.L.; BROOK, P.V.D.; KNILL, D.C. (Orgs.). *Cognition, conceptual and methodological issues*. Washington: American Psychological Association.
- JOHNSON, K. (2011). *Acoustic and auditory phonetics*. Oxford: Blackwell.
- JONGMAN, A.; WAYLAND, R. WONGS, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108, pp. 1252-63.
- JUSCZYK, P.W.; PISONI, D.B.; MULLENIX, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, 43, pp. 253-91.
- JUSCZYK, P. W.; HOUSTON, D. M.; NEWSOME, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, pp. 159-207.
- KLATT, D.H. (1979). Software for cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, pp. 971-95.
- KLATT, D.H. (1989). Review of selected models of speech perception. In W.D., MARSLEN-WILSON (Ed.). *Lexical representation and process*. Cambridge, MA: MIT Press.
- KLATT, D.H.; KLATT, L.C. (1990). "Analysis, synthesis and perception of voice quality variations among male and female talkers." *Journal of the Acoustical Society of America*, 87, pp. 820-56.
- KRAUSE, S.E. (1982). Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults. *Journal of the Acoustical Society of America*, 71, pp. 990-95.
- KUHL, P.K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, 6, pp. 263-85.
- KUHL, P.K. (1987). Perception of speech and sound in early infancy. In: SALAPATEK, P.; COHEN, L. (Orgs.). *Handbook of Infant Perception*. Orlando: Academic.
- KUHL, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5, pp. 831-45.
- KUIJPERS, C.T.L. (1996). Perception of voicing contrast by Dutch children and adults. *Journal of Phonetics*, 24, pp. 367-82.
- LACERDA, F.P. (1982). Acoustic perceptual study of the Portuguese voiceless fricative. *J. Phonet.*, 10, pp. 11-22.
- LADEFOGED, P.; MADDIESON, I. (1996). *The sound's of the world's languages*. Massachusetts: Blackwell, pp. 47-101.
- LANY, J.; SAFFRAN, J.R. (2013). Statistical learning mechanism in infancy. In: RUBENSTEIN, J.L.R; RAKIC, P. (Org). *Neural circuit development and function in the healthy and diseased brain*. Amsterdam: Elsevier, pp. 231-48.

- LECANUET, J. (1998). Fetal responses to auditory speech stimuli. In: SLATER, A. (Org.). *Perceptual development – visual, auditory and speech perception in infancy*. Reino Unido: Hove.
- LIBERMAN, A.; HARRIS, K.; HOFFMAN, H.; GRIFFITH, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, 58, pp. 358-68.
- LIBERMAN, A.M., HARRIS, K.S., KINNEY, J.A.; LANE, H. (1961). The discrimination of relative onset time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology*, 61, pp. 379-88.
- LIEBERMAN, P.; BLUMSTEIN, S.E. (1988). *Speech physiology, speech perception and acoustic phonetics*. Cambridge: Cambridge, University Press.
- LINDBLOM, B.; MACNEILAGE, P.; STUDDERT-KENNEDY, M. (1983). Self-organizing processes and the explanation of phonological universals. *Linguistics*, pp. 181-204.
- LISKER, L.; ABRAMSOM, A.S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word*, 20, pp. 384-422.
- LOUSADA, M.; JESUS, L.M.T.; HALL. (2010). Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association*, 40, pp. 261-75.
- MANN, V.A.; REPP, B.H. (1980). Influence of vocalic context on perception of the [S]-[s] distinction. *Perception e Psychophysics*, 28, pp. 213-28.
- MANN, V.A.; SHARLIN, H.M.; DORMAN, M. (1985). Children's perception of sibilants: The relation between articulation and perceptual development. *Journal of Experimental Child Psychology*, 39, pp. 252-64.
- MASSARO, D.W.; COHEN, M.M. (1983). Categorical or continuous speech perception: A new test. *Speech Communication*, 2, pp. 15-35.
- MAYE, J.; WEISS, D.; ASLIN, R. (2008). Statistical phonetic learning in infants: facilitation and features generalization. *Developmental science*, 11, pp. 122-34.
- MAYO, C.; TURK, A. (2004). Adult-child differences in acoustic cue weighting are influenced by segmental context: children are not always perceptually biased toward transitions. *Journal of the Acoustical Society of America*, 115, pp. 3184-94.
- MAYO, C.; TURK, A. (2005a). The influence of spectral distinctiveness on acoustic cue weighting in children's and adult's speech perception. *Journal of the Acoustical Society of America*, 118, pp. 1730-41.
- MAYO, C.; TURK, A. (2005b). No available theories currently explain all adult-child cue weighting differences. *Centre for Speech Technology Research e Theoretical and Applied Linguistics*. University of Edinburgh: UK.

- McCARTHY, J.J.; PRINCE, A.S. (1993). *Prosodic morphology I: Constraint interaction and satisfaction*. New Brunswick: Rutgers University Center for Cognitive Science.
- MELO, R.M.; MOTA, H.B.; MEZZOMO, C.L.; BRASIL, C.B.; LOVATTO, L.; ARZENO, L. (2011). Caracterização acústica da sonoridade dos fonemas plosivos do português brasileiro. *Revista CEFAC*, 14, pp. 487-99.
- MENN, L. (1971). Phonotactic rules in beginning speech. *Lingua*, 26, pp. 225-51.
- MILLER, J.L., EIMAS, P.D. (1995). Speech Perception: From Signal to Word. *Annual Review of Psychology*, 46, pp. 467-92.
- MONTGOMERY, C.R.; CLARKSON, M.G. (1997). Infant's speech perception: masking by low and high-frequency noises. *Journal of the Acoustical Society of America*, 102, pp. 3665-72.
- MORRONGIELLO, B. A.; ROBSON, R. C.; BEST, C. T.; CLIFTON, R. K. (1984). Trading relations in the perception of speech by 5-year-old children. *Journal of Experimental Child Psychology*, 37(2), 231-250.
- NITTROUER, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, pp. 351-82.
- NITTROUER, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech, Language and Hearing Research*, 39, pp. 278-97.
- NITTROUER, S. (2002). From ear to cortex: a perspective on what clinicians need to understand about speech perception and language processing. *Language, Speech, and Hearing Services in Schools*, 33, pp. 237-52.
- NITTROUER, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America*, 112, pp. 711-19.
- NITTROUER, S. (2004). The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults. *Journal of the Acoustical Society of America*, 115, pp. 1777-90.
- NITTROUER, S.; CROWTHER, C.S. (1998). Examining the role of auditory sensitivity in the developmental weighting shift. *Journal of Speech, Language and Hearing Research*, 41, pp. 809-18.
- NITTROUER, S.; LOWENSTEIN, J.H. (2007). Children's weighting strategies for word-final stop voicing are not explained by auditory sensitivities. *Journal of Speech, Language, and Hearing Research*, 50, pp. 58-73.
- NITTROUER, S.; MANNING, C.; MEYER, G. (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *Journal of the Acoustical Society of America*, 94, pp. 1865.

- NITTROUER, S.; MILLER, M.E. (1997a). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, 101, pp. 2253-66.
- NITTROUER, S.; MILLER, M.E. (1997b). Developmental weighting shifts for noise components of fricative syllables. *Journal of the Acoustical Society of America*, 102, pp. 572-80.
- NITTROUER, S.; MILLER, M.E.; CROWTHER, C.S.; MANHART, M.J. (2000). The effect of segmental order on fricative labeling by children and adults. *Perception e Psychophysics*, 62, pp. 266-84.
- NITTROUER, S.; STUDDERT-KENNEDY, M. (1987). The role of coarticulatory effects on the perception of fricatives by children and adults. *Journal of Speech, Language and Hearing Research*, 30, pp. 319-29.
- NITTROUER, S.; STUDDERT-KENNEDY, M.; MCGOWAN, P. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research*, 32, pp. 120-32.
- NORTHERN, J.L.; DOWNS, M.P. (1989). *Audição em crianças*. São Paulo: Manole.
- OHDE, R.N.; HALEY, K.L. (1997). Stop-consonant and vowel perception in 3- and 4-year-old children. *Journal of the Acoustical Society of America*, 102, pp. 3311-22.
- OHDE, R.N.; HALEY, K.L.; McMAHON, C.W. (1996). A developmental study of vowel perception from brief synthetic consonant-vowel syllables. *Journal of the Acoustical Society of America*, 100, pp. 3813-24.
- PENIDO, F.A. (2012). *A percepção das fricativas coronais não vozeadas por crianças e adultos falantes nativos do português brasileiro: em busca de evidências para uma mudança evolutiva nos esquemas perceptuais*. Dissertação (Mestrado em Estudos Linguísticos). Belo Horizonte: FALE/UFMG.
- PETERSON, G.R.; SHOUP, J.E. (1966). Glossary of terms from the physiological and acoustic phonetic theories. *J. Speech Res.*, 9, pp. 100-20.
- PIERREHUMBERT, J. (2003). *Probabilistic Phonology: discrimination and robustness*. In: R. Bod, J. Hay, S. Jannedy (Eds). pp. 177-228.
- POTTER, R.K.; KOPP, G.A.; GREEN KOPP, H. (1947) – apud – *Reading the fricatives sounds*. In: _____ - Visible speech. New York, Dover, 1966. P. 110-65.
- PRINCE, A.; SMOLENSKY, P. (1993). *Optimality theory: Constraint interaction and generative grammar*. Report n. RuCCS-TR-2. New Brunswick: Rutgers University Center for Cognitive Science.
- PUJOL, R.; LAVIGNE-REBILLARD, M.; UZIEL, A. (1991). Development of the human cochlea. *Acta Oto-laryngologica*, Oslo, 111, pp. 7-123.

- QUERLEU, D.; RENARD, X.; VERSYP, F.; PARIS-DELRUE, L.; CRÉPIN, G. (1988). Fetal hearing. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, Amsterdam, 29, pp. 191-212.
- REPP, B.H. (1984). Categorical Perception: Issues, methods, findings. In: Lass, N.J. (Ed.). *Speech and language: Advances in Basic Research and Practice*, 10, pp. 244-322.
- RIKI, T.S.; MINKA, H.; LIAT, K.R. (2007). The effect of age on acoustic cue weighting in the perception of initial stop voicing contrast in Hebrew. In: *ICPhS XVI*, Saarbrücken, pp. 705-8.
- ROMBERG, A.R.; SAFFRAN, J.F. (2010). Statistical learning and language acquisition. *Cognitive Science*, 1, pp. 906-14.
- ROTHE-NEVES, R.; PENIDO, F. A. (2013). F2 transition as a cue to place of articulation in Brazilian Portuguese coronal fricatives. In: *VIIth GSCP International Conference Speech and Corpora*, Belo Horizonte. Proceedings of the VIIth GSCP International Conference Speech and Corpora, pp. 278-280.
- ROTHE-NEVES, R., VITOR, R.M. (2008). Desenvolvimento da linguagem e procedimentos de avaliação In: *Aspectos biopsicossociais da saúde na infância e na adolescência*. Belo Horizonte: Coopmed, pp. 313-35.
- RUSSO, I.C.P.; BEHLAU, M.S. (1993). Percepção da fala: análise acústica do português brasileiro. São Paulo: Lovise.
- SAMCZUK, I.B.; GAMA-ROSSI, A. (2004). *Descrição fonético-acústica das fricativas do português brasileiro*. PUC, São Paulo.
- SAMUEL, A.G. (2011). Speech Perception. *Annual Review of Psychology*, 62, pp. 49-72.
- SANTOS, M.T.M. (1987). *Uma análise espectrográfica dos sons fricativos surdos e sonoros do Português Brasileiro*. Tese (Especialização em Fonoaudiologia) – Escola Paulista de Medicina, São Paulo.
- SANTOS, R. S. (2008). Adquirindo a fonologia de uma língua: produção, percepção e representação fonológica. *Alfa*, 52, pp. 465-81.
- SCHINDLER, O. (1974). *Material fonemico*. In: _____ - *Manuale di audiofono – logopedia*. Torino, Editrice Omega, p.552.
- SCHOUTEN, B.; GERRITS, E.; VAN HESSEN, A. (2003). The end of categorical perception as we know it. *Speech Communication*, 41, pp. 71-80.
- SENSIMETRICS ORG. (n.d). SenSyn: Speech Synthesizer Package, Cambridge, MA
- SILVA, T.C. (2001). *Fonética e fonologia do português. Roteiros de estudos e guia de exercícios*. São Paulo: Contexto, 2001. P. 38-9.

SILVA, D.M.R.; ROTHE-NEVES, R. (2009). Um estudo experimental sobre a percepção do contraste entre as vogais médias posteriores do português brasileiro. *D.E.L.T.A.*, 25, pp. 319-45.

STEVENS, K.N.; BLUMSTEIN, S.E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, pp. 1358-68.

STEVENS, K.N.; KLATT, D.H. (1968). *Studies of acoustic properties of speech sound*. Massachusetts, Air Force Cambridge research Laboratories, pp. 92.

STOEP, J.; VERHOEVEN, L. (2001). Family and classroom predictors of children's early language and literacy development. In: SHANAHAN, T.; RODRIGUEZ-BROWN, F.V. (Eds.), *National Reading Conference Yearbook*, 49, pp. 209-21.

SUSSMAN, J.E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *Journal of the Acoustical Society of America*, 109, pp. 1173-80.

SVACHKIN, N.K. (1973). The development of phonemic speech perception in early childhood. In: FERGUSON, C.A.; SLOBIN, D.I. (Orgs.). *Studies of child language development*. New York: Holt, Rinehart and Winston, (Trabalho original publicado em 1948), pp. 91-127.

TOSI, O. (1979). *Vibratory motion*. In: _____ - Textbook of physics for audiologists and speech pathologists. Lansing: Linko Copies, pp. 278.

WERKER, J.F. (1993). Developmental changes in cross-language speech perception: implications for cognitive models of speech processing. In: ALTMANN, G.; SHILLCOCK, R. *Cognitive models of speech processing: the second sperlonga workshop*. Essex, England: Lawrence Erlbaum Associates, pp. 57-78.

WERKER, J.F.; POLKA, L. (1993). Developmental changes in speech perception: new challenges and new directions. *Journal of Phonetics*, 21, pp. 83-101.

WERKER, J. F.; POLKA, L.; PEGG, J. E. (1997). The conditioned head turn procedure as a method for testing infant speech perception. *Early Development and Parenting*, 6, pp. 171-178.

WERKER, J.F.; TESS, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, pp. 49-63.

WALLEY, A.C.; CARRELL, T.D. (1983). Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 73, pp. 1011-22.

WRIGHT, R. (2004). A review of perceptual cues and cue robustness. In: BRUCE, H.; ROBERT, K.; DONCA, S. *Phonetically Based Phonology*. Cambridge: Cambridge University Press, pp. 34-57.

YAVAS, M.; HERNANDORENA, C.L.M.; LAMPRECHT, R.R. (2002). Avaliação fonológica da criança. Porto Alegre: Artmed.

ANEXOS

ANEXO 1. Termo de consentimento livre e esclarecido para os participantes adultos do estudo piloto de análise das propriedades acústicas dos estímulos

Nº Registro COEP: CAAE – 62617616.0.0000.5149

Título do Projeto: Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro.

1. Introdução

Este projeto de pesquisa possui como tema principal a investigação do desenvolvimento da percepção da fala. No entanto, antes de pesquisar este tema tornou-se necessário desenvolver um estudo piloto, que tem como proposta analisar as características acústicas dos sons presentes nas palavras “sapa”, “chapa”, “saba”, “chaba”, “bata” e “data”.

Você está sendo convidado(a) a participar da coleta de dados deste estudo prévio. Você foi selecionado, uma vez que apresenta idade entre 20 e 45 anos e sua participação não é obrigatória.

2. Procedimento do estudo

O procedimento deste estudo envolve a participação nas gravações das palavras “sapa”, “chapa”, “saba”, “chaba”, “bata” e “data”, no Laboratório de Fonética da Universidade Federal de Minas Gerais.

3. Riscos e desconfortos

Consideramos que a metodologia utilizada para coleta de dados oferece muito pouco ou, eventualmente, nenhum prejuízo para você. Talvez possa haver cansaço durante a execução da gravação. Desta forma, se tal fato ocorrer, será realizada uma pausa na coleta para o seu descanso.

4. Benefícios

A sua participação é voluntária e não trará qualquer benefício direto, mas proporcionará à pesquisadora um melhor conhecimento sobre as características acústicas dos sons gravados, o que contribuirá efetivamente para o desenvolvimento da pesquisa.

5. Custos/Reembolso

Não existirão despesas e/ou compensações financeiras relacionadas à sua participação no estudo.

6. Caráter Confidencial dos Registros

A sua identidade será mantida em sigilo. Os resultados da pesquisa serão sempre apresentados como o retrato de um grupo e não de uma pessoa. Dessa forma, você não será identificado quando o material de seu registro for utilizado ou quando os seus resultados forem publicados em jornais, revistas e eventos científicos ou educativos. O material coletado ficará sob a guarda da coordenação do projeto e não será divulgado publicamente.

7. Participação

A sua participação neste estudo é muito importante e voluntária. Você tem o direito de não querer participar ou de sair deste estudo a qualquer momento. Em caso de você decidir retirar-se do estudo, favor informar o pesquisador.

8. Informações

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, que poderá ser contatado para esclarecimentos pelo telefone 3499-4592, por email coep@prpq.ufmg.br ou no seguinte endereço: Av. Antonio Carlos, 6627 – Unidade Administrativa II, sala 2005. CEP 31270-901 – Belo Horizonte, MG.

Os pesquisadores responsáveis poderão fornecer qualquer esclarecimento sobre essa pesquisa, assim como tirar dúvidas, bastando contato no seguinte endereço e/ou telefone:

Nome do pesquisador: Fabiana Andrade Penido

Endereço: Rua Stella Hanriot, 211/801

Telefone: (31) 99641-4171

Email: fabianapenido@gmail.com

9. Declaração de Consentimento

Li ou alguém leu para mim as informações contidas neste documento antes de assinar este termo de consentimento. Declaro que toda a linguagem técnica utilizada na descrição deste estudo de pesquisa foi satisfatoriamente explicada e que recebi respostas para todas as minhas dúvidas. Confirmando também que recebi uma cópia deste Termo de Consentimento Livre e Esclarecido. Compreendo que sou livre para me retirar do estudo em qualquer momento, sem perda de benefícios ou qualquer outra penalidade.

Dou meu consentimento de livre e espontânea vontade para participar deste estudo.

Nome do participante

Assinatura do participante

Obrigada pela sua colaboração e por merecer sua confiança

Fabiana Andrade Penido
(Pesquisadora responsável)

_____/_____/_____
Data

ANEXO 2. Parâmetros de controle do *software* Sensyn para a síntese dos estímulos

PARÂMETROS DE CONTROLE CONSTANTES

PARAM	SYM	MIN	VAL	MAX	DESCRIPTION
1.	DU	30	500	5000	Duration of synthesis utterance, in ms
2.	UI	1	5	20	Update interval for parameters reset, in ms
3.	SR	5000	10000	20000	Output sampling rate, in samples/sec
4.	NF	1	5	6	Number of formants in cascade branch
5.	SS	1	2	3	Source switch (1=impulse, 2=natural, 3=LF model)
6.	RS	1	8	8191	Random seed(initial value of random number generator)
7.	SB	0	1	1	Same noise burst, reset RS if SF =0 and AH =0 (0=no, 1=yes)
8.	CP	0	0	1	0 implies cascade, 1 implies parallel tract excitation by AV
9.	OS	0	0	20	Output selector (0=normal voice, see below for other values)
10.	GV	0	60	80	Overall gain scale factor for AV , in dB
11.	GH	0	60	80	Overall gain scale factor for AH , in dB
12.	GF	0	0	80	Overall gain scale factor for AF , in dB

PARÂMETROS DE CONTROLE VARIÁVEIS

PARAM	SYM	MIN	VAL	MAX	DESCRIPTION
13.	F0	0	1000	5000	Fundamental frequency, in tenths of a Hz
14.	AV	0	60	80	Amplitude of voicing, in dB
15.	OQ	10	50	99	Open quotient (voicing open-time/period), in %
16.	SQ	100	200	500	Speech quotient (rise/fall time of open period, LF model), in %
17.	TL	0	0	41	Extra tilt of voicing spectrum, dB down @ 3 kHz
18.	FL	0	0	100	Flutter (random fluctuation in F0), in % of maximum
19.	DI	0	0	100	Diplophonia (pairs of periods migrate together), in 5 of maximum
20.	AH	0	0	80	Amplitude of aspiration, in dB
21.	AF	0	0	80	Amplitude of frication, in dB
22.	F1	180	500	1300	Frequency of the 1 st formant, in Hz
23.	B1	30	60	1000	Bandwidth of the 1 st formant, in Hz
24.	DF1	0	0	100	Change in F1 during open portion of a period, in Hz
25.	DB1	0	0	400	Change in B1 during open portion of a period, in Hz
26.	F2	550	1500	3000	Frequency of the 2 nd formant, in Hz
27.	B2	40	90	1000	Bandwidth of the 2 nd formant, in Hz
28.	F3	1200	2500	4800	Frequency of the 3 rd formant, in Hz
29.	B3	60	150	1000	Bandwidth of the 3 rd formant, in Hz
30.	F4	2400	3250	4990	Frequency of the 4 th formant, in Hz
31.	B4	100	200	1000	Bandwidth of the 4 th formant, in Hz
32.	F5	3000	3700	4990	Frequency of the 5 th formant, in Hz
33.	B5	100	200	1500	Bandwidth of the 5 th formant, in Hz
34.	F6	3000	4990	4990	Frequency of the 6 th formant, in Hz
35.	B6	100	500	4000	Bandwidth of the 6 th formant, in Hz
36.	FNP	180	500	2000	Frequency of the nasal pole, in Hz
37.	BNP	40	90	1000	Bandwidth of the nasal pole, in Hz
38.	FNZ	180	500	2000	Frequency of the nasal zero, in Hz
39.	BNZ	40	90	1000	Bandwidth of the nasal zero, in Hz
40.	FTP	300	2150	3000	Frequency of the tracheal pole, in Hz
41.	BTP	40	180	1000	Bandwidth of the tracheal pole, in Hz
42.	FTZ	300	2150	3000	Frequency of the tracheal zero, in Hz
43.	BTZ	40	180	1000	Bandwidth of the tracheal zero, in Hz
44.	A2F	0	0	80	Amplitude of frication-excited parallel in 2 nd formant, in dB
45.	A3F	0	0	80	Amplitude of frication-excited parallel in 3 rd formant, in dB
46.	A4F	0	0	80	Amplitude of frication-excited parallel in 6 th formant, in dB
47.	AB	0	0	80	Amplitude of frication-excited parallel bypass path, in dB
48.	B2F	40	250	1000	Bandwidth of frication-excited parallel in 2 nd formant, in dB
49.	B3F	60	320	1000	Bandwidth of frication-excited parallel in 3 rd formant, in dB
50.	B4F	100	350	1000	Bandwidth of frication-excited parallel in 4 th formant, in dB
51.	B5F	100	500	1500	Bandwidth of frication-excited parallel in 5 th formant, in dB
52.	B6F	100	1500	4000	Bandwidth of frication-excited parallel in 6 th formant, in dB
53.	ANV	0	0	80	Amplitude of voicing-excited parallel nasal formant, in dB
54.	A1V	0	60	80	Amplitude of voicing-excited parallel 1 st formant, in dB
55.	A2V	0	60	80	Amplitude of voicing-excited parallel 2 nd formant, in dB
56.	A3V	0	60	80	Amplitude of voicing-excited parallel 3 rd formant, in dB
57.	A4V	0	60	80	Amplitude of voicing-excited parallel 4 th formant, in dB
58.	ATV	0	0	80	Amplitude of voicing-excited parallel tracheal formant, in dB

ANEXO 3. Termo de consentimento livre e esclarecido para os participantes adultos do estudo piloto de classificação do estímulos sintetizados

N.º Registro COEP: CAAE – 62617616.0.0000.5149

Título do Projeto: Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro.

1. Introdução

Os sons da fala são constituídos de consoantes e vogais. Cada som do português brasileiro apresenta diversas características acústicas. Assim, para perceber e discriminar estes sons, os ouvintes prestam atenção nestas características, que são consideradas pistas acústicas.

Estudos têm mostrado que para discriminar os pares (sa)-(cha) e (su)-(shu), crianças e adultos ponderam as pistas acústicas diferentemente (ponderação de pistas acústicas), uma vez que as crianças prestam mais atenção em uma pista e os adultos em outra.

Você está sendo convidado(a) a participar de uma pesquisa, que busca investigar o desenvolvimento da percepção da fala. Você foi selecionado, uma vez que apresenta idade entre 20 e 40 anos e sua participação não é obrigatória.

2. Procedimento do estudo

O procedimento deste estudo consiste em ouvir os sons por meio de um fone de ouvido e responder qual som foi escutado dentre as possibilidades, (chapa)-(sapa), (chaba)-(saba) e (bata)-(data). Assim, uma imagem representando estas palavras serão colocadas em cada lado da tela do computador e, após percebê-los, você deverá indicá-las de acordo com as respostas presentes na tela do monitor.

3. Riscos e desconfortos

Consideramos que a metodologia utilizada para coleta de dados oferece muito pouco ou, eventualmente, nenhum prejuízo para você. Talvez possa haver cansaço durante a execução da tarefa. Desta forma, a coleta de dados será realizada com as pausas necessárias.

4. Benefícios

A sua participação é voluntária e não trará qualquer benefício direto, mas proporcionará à pesquisadora um melhor conhecimento sobre o desenvolvimento da percepção da fala.

5. Custos/Reembolso

Não existirão despesas e/ou compensações financeiras relacionadas à sua participação no estudo.

6. Caráter Confidencial dos Registros

A sua identidade será mantida em sigilo. Os resultados da pesquisa serão sempre apresentados como o retrato de um grupo e não de uma pessoa. Dessa forma, você não será identificado quando o material de seu registro for utilizado ou quando os seus resultados forem publicados em jornais, revistas e eventos científicos ou educativos. O material coletado ficará sob a guarda da coordenação do projeto e não será divulgado publicamente.

7. Participação

A sua participação neste estudo é muito importante e voluntária. Você tem o direito de não querer participar ou de sair deste estudo a qualquer momento. Em caso de você decidir retirar-se do estudo, favor informar o pesquisador.

8. Informações

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, que poderá ser contatado para esclarecimentos pelo telefone 3499-4592, por email coep@prpq.ufmg.br ou no seguinte endereço: Av. Antonio Carlos, 6627 – Unidade Administrativa II, sala 2005. CEP 31270-901 – Belo Horizonte, MG.

Os pesquisadores responsáveis poderão fornecer qualquer esclarecimento sobre essa pesquisa, assim como tirar dúvidas, bastando contato no seguinte endereço e/ou telefone:

Nome do pesquisador: Fabiana Andrade Penido

Endereço: Rua Stella Hanriot, 211/801

Telefone: (31) 99641-4171

Email: fabianapenido@gmail.com

9. Declaração de Consentimento

Li ou alguém leu para mim as informações contidas neste documento antes de assinar este termo de consentimento. Declaro que toda a linguagem técnica utilizada na descrição deste estudo de pesquisa foi satisfatoriamente explicada e que recebi respostas para todas as minhas dúvidas. Confirmando também que recebi uma cópia deste Termo de Consentimento Livre e Esclarecido. Compreendo que sou livre para me retirar do estudo em qualquer momento, sem perda de benefícios ou qualquer outra penalidade.

Dou meu consentimento de livre e espontânea vontade para participar deste estudo.

Nome do participante

Assinatura do participante

Obrigada pela sua colaboração e por merecer sua confiança

Fabiana Andrade Penido
(Pesquisadora responsável)

_____/_____/_____
Data

ANEXO 4. *Script* utilizado para a criação do continuum fricativo

```
if numberOfSelected("Sound") <>2
  exit select two sounds
endif

s1 = selected ("Sound", 1)
s2$ = selected$("Sound",2)

numberOfSteps = 9

numberOfStep = numberOfSteps - 1

for i from 0 to 'numberOfSteps'
  f = 'i'/numberOfSteps'
  i_f = 1 - 'f'
  select 's1'
  Copy... interpolation'i'
  Formula... 'i_f' * self[col] + 'f' *Sound_'s2$'[col]
endfor
```

ANEXO 5. Termo de consentimento livre e esclarecido para os participantes de quatro anos de idade – experimento de ponderação de pistas auditivas

Título do Projeto: Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro.

N.º Registro COEP: CAAE – 62617616.0.0000.5149

Prezados pais ou responsável,

Este Termo de Consentimento pode conter palavras que você não entenda. Caso isso ocorra, a pesquisadora encontra-se à disposição para explicar eventuais dúvidas.

1. Introdução

Os sons da fala são constituídos de consoantes e vogais. Cada som do português brasileiro apresenta diversas características acústicas. Assim, para perceber e discriminar estes sons, os ouvintes prestam atenção nestas características, que são consideradas pistas auditivas.

Estudos têm mostrado que para discriminar os pares (sa)-(cha) e (su)-(chu), crianças e adultos ponderam as pistas auditivas diferentemente, uma vez que as crianças prestam mais atenção em uma pista e os adultos em outra. Este fenômeno chama-se ponderação de pistas auditivas e a investigação deste tema é considerada importante para compreendermos melhor como o desenvolvimento da percepção da fala ocorre.

Seu(sua) filho(a) está sendo convidado(a) a participar de uma pesquisa de percepção da fala, que tem como objetivo verificar se as crianças atribuem os mesmos pesos às pistas auditivas disponíveis nas palavras “sapa”, “chapa”, “bata”, “data”, “saba” e “chaba” do que os adultos para a identificação destas palavras do português brasileiro. Seu(sua) filho(a) foi selecionado, uma vez que apresenta idade de 4 anos. Vale ressaltar, que a participação de seu(sua) filho(a) não é obrigatória.

2. Procedimento do estudo

O procedimento deste estudo consiste em ouvir as palavras por meio de um fone de ouvido e responder qual palavra foi escutada dentre as possibilidades, (sapa)-(chapa), (saba)-(chaba) e (bata)-(data). Assim, figuras representando estas palavras serão colocadas em cada lado da tela do computador e, após percebê-las, seu(sua) filho(a) deverá indicá-las de acordo com as respostas presentes na tela do monitor.

3. Riscos e desconfortos

Consideramos que a metodologia utilizada para coleta de dados oferece muito pouco ou, eventualmente, nenhum prejuízo ao seu(sua) filho(a). Devido à idade de quatro anos, talvez possa haver cansaço do seu(sua) filho(a) durante a execução da tarefa. Desta forma, a coleta de dados será realizada com as pausas necessárias.

4. Benefícios

Consideramos que essa pesquisa não trará benefícios diretos para seu(sua) filho(a). Porém, vale ressaltar, que para participação na mesma, será realizada uma triagem auditiva com o objetivo de descartar qualquer grau de perda auditiva. Diante da necessidade de realizar uma avaliação audiológica completa em seu filho(a) você será comunicado(a) e receberá as devidas orientações. Além disso, será realizada uma avaliação do desenvolvimento da fala

de seu(sua) filho(a). Se for constatado qualquer alteração você também será informado(a) e orientado(a).

5. Custos/Reembolso

Você não terá nenhum gasto com a participação de seu(sua) filho(a) no estudo, uma vez que este será realizado na própria escola. Além disso, não receberá pagamento pela participação de seu(sua) filho(a) no mesmo.

6. Caráter Confidencial dos Registros

A identidade de seu(sua) filho(a) será mantida em sigilo. Os resultados da pesquisa serão sempre apresentados como o retrato de um grupo e não de uma pessoa. Dessa forma, seu(sua) filho(a) não será identificado(a) quando o material for utilizado ou quando os seus resultados forem publicados em jornais, revistas e eventos científicos ou educativos. O material coletado ficará sob a guarda da coordenação do projeto e não será divulgado publicamente.

7. Participação

A participação de seu(sua) filho(a) neste estudo é voluntária e poderá ser interrompida a qualquer momento, caso julgue necessário. Você tem o direito de não autorizar a participação de seu(sua) filho(a). Caso você decida retirá-lo do estudo, favor informar a pesquisadora.

8. Informações

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, que poderá ser contatado para esclarecimentos pelo telefone 3499-4592, por email coep@prpq.ufmg.br ou no seguinte endereço: Av. Antônio Carlos, 6627 – Unidade Administrativa II, sala 2005. CEP 31270-901 – Belo Horizonte, MG.

A pesquisadora responsável poderá fornecer qualquer esclarecimento sobre essa pesquisa, assim como tirar dúvidas, bastando contato no seguinte endereço e/ou telefone:

Nome do pesquisador: Fabiana Andrade Penido

Endereço: Rua Stella Hanriot, 211/801

Telefone: (31) 99641-4171

Email: fabianapenido@gmail.com

9. Declaração de Consentimento

Li ou alguém leu para mim as informações contidas neste documento antes de assinar este termo de consentimento. Declaro que toda a linguagem técnica utilizada na descrição deste estudo de pesquisa foi satisfatoriamente explicada e que recebi respostas para todas as minhas dúvidas. Confirmando também que recebi uma cópia deste Termo de Consentimento Livre e Esclarecido. Compreendo que sou livre para me retirar meu filho(a) do estudo em qualquer momento, sem perda de benefícios ou qualquer outra penalidade.

Dou meu consentimento de livre e espontânea vontade para meu filho(a) participar deste estudo.

Nome do participante

Assinatura do representante legal

Obrigado pela sua colaboração e por merecer sua confiança.

Fabiana Andrade Penido (pesquisadora)

_____/_____/_____
Data

ANEXO 6. Termo de consentimento livre e esclarecido para os participantes de sete anos de idade – experimento de ponderação de pistas auditivas

Título do Projeto: Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro.

N.º Registro COEP: CAAE – 62617616.0.0000.5149

Prezados pais ou responsável,

Este Termo de Consentimento pode conter palavras que você não entenda. Caso isso ocorra, a pesquisadora encontra-se à disposição para explicar eventuais dúvidas.

1. Introdução

Os sons da fala são constituídos de consoantes e vogais. Cada som do português brasileiro apresenta diversas características acústicas. Assim, para perceber e discriminar estes sons, os ouvintes prestam atenção nestas características, que são consideradas pistas auditivas.

Estudos têm mostrado que para discriminar os pares (sa)-(cha) e (su)-(chu), crianças e adultos ponderam as pistas auditivas diferentemente, uma vez que as crianças prestam mais atenção em uma pista e os adultos em outra. Este fenômeno chama-se ponderação de pistas auditivas e a investigação deste tema é considerada importante para compreendermos melhor como o desenvolvimento da percepção da fala ocorre.

Seu(sua) filho(a) está sendo convidado(a) a participar de uma pesquisa de percepção da fala, que tem como objetivo verificar se as crianças atribuem os mesmos pesos às pistas auditivas disponíveis nas palavras “sapa”, “chapa”, “bata”, “data”, “saba” e “chaba” do que os adultos para a identificação destas palavras do português brasileiro. Seu(sua) filho(a) foi selecionado, uma vez que apresenta idade de 7 anos. Vale ressaltar, que a participação de seu(sua) filho(a) não é obrigatória.

2. Procedimento do estudo

O procedimento deste estudo consiste em ouvir as palavras por meio de um fone de ouvido e responder qual palavra foi escutada dentre as possibilidades, (sapa)-(chapa), (saba)-(chaba) e (bata)-(data). Assim, figuras representando estas palavras serão colocadas em cada lado da tela do computador e, após percebê-las, seu(sua) filho(a) deverá indicá-las de acordo com as respostas presentes na tela do monitor.

3. Riscos e desconfortos

Consideramos que a metodologia utilizada para coleta de dados oferece muito pouco ou, eventualmente, nenhum prejuízo ao seu(sua) filho(a). Devido à idade de sete anos, talvez possa haver cansaço do seu(sua) filho(a) durante a execução da tarefa. Desta forma, a coleta de dados será realizada com as pausas necessárias.

4. Benefícios

Consideramos que essa pesquisa não trará benefícios diretos para seu(sua) filho(a). Porém, vale ressaltar, que para participação na mesma, será realizada uma triagem auditiva com o objetivo de descartar qualquer grau de perda auditiva. Diante da necessidade de realizar uma avaliação audiológica completa em seu filho(a) você será comunicado(a) e receberá as devidas orientações.

5. Custos/Reembolso

Você não terá nenhum gasto com a participação de seu(sua) filho(a) no estudo, uma vez que este será realizado na própria escola. Além disso, não receberá pagamento pela participação de seu(sua) filho(a) no mesmo.

6. Caráter Confidencial dos Registros

A identidade de seu(sua) filho(a) será mantida em sigilo. Os resultados da pesquisa serão sempre apresentados como o retrato de um grupo e não de uma pessoa. Dessa forma, seu(sua) filho(a) não será identificado(a) quando o material for utilizado ou quando os seus resultados forem publicados em jornais, revistas e eventos científicos ou educativos. O material coletado ficará sob a guarda da coordenação do projeto e não será divulgado publicamente.

7. Participação

A participação de seu(sua) filho(a) neste estudo é voluntária e poderá ser interrompida a qualquer momento, caso julgue necessário. Você tem o direito de não autorizar a participação de seu(sua) filho(a). Caso você decida retirá-lo do estudo, favor informar a pesquisadora.

8. Informações

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, que poderá ser contatado para esclarecimentos pelo telefone 3499-4592, por email coep@prpq.ufmg.br ou no seguinte endereço: Av. Antônio Carlos, 6627 – Unidade Administrativa II, sala 2005. CEP 31270-901 – Belo Horizonte, MG.

A pesquisadora responsável poderá fornecer qualquer esclarecimento sobre essa pesquisa, assim como tirar dúvidas, bastando contato no seguinte endereço e/ou telefone:

Nome do pesquisador: Fabiana Andrade Penido

Endereço: Rua Stella Hanriot, 211/801

Telefone: (31) 99641-4171

Email: fabianapenido@gmail.com

9. Declaração de Consentimento

Li ou alguém leu para mim as informações contidas neste documento antes de assinar este termo de consentimento. Declaro que toda a linguagem técnica utilizada na descrição deste estudo de pesquisa foi satisfatoriamente explicada e que recebi respostas para todas as minhas dúvidas. Confirmando também que recebi uma cópia deste Termo de Consentimento Livre e Esclarecido. Compreendo que sou livre para me retirar meu filho(a) do estudo em qualquer momento, sem perda de benefícios ou qualquer outra penalidade.

Dou meu consentimento de livre e espontânea vontade para meu filho(a) participar deste estudo.

Nome do participante

Assinatura do representante legal

Obrigado pela sua colaboração e por merecer sua confiança.

Fabiana Andrade Penido (pesquisadora)

_____/_____/_____
Data

ANEXO 7. Termo de consentimento livre e esclarecido para os participantes adultos – experimento de ponderação de pistas auditivas

Título do Projeto: Ponderação de pistas auditivas por crianças e adultos falantes nativos do português brasileiro.

N.º Registro COEP: CAAE – 62617616.0.0000.5149

Prezado (a) Senhor (a),

Este Termo de Consentimento pode conter palavras que você não entenda. Caso isso ocorra, a pesquisadora encontra-se à disposição para explicar eventuais dúvidas.

1. Introdução

Os sons da fala são constituídos de consoantes e vogais. Cada som do português brasileiro apresenta diversas características acústicas. Assim, para perceber e discriminar estes sons, os ouvintes prestam atenção nestas características, que são consideradas pistas auditivas.

Estudos têm mostrado que para discriminar os pares (sa)-(cha) e (su)-(chu), crianças e adultos ponderam as pistas auditivas diferentemente, uma vez que as crianças prestam mais atenção em uma pista e os adultos em outra. Este fenômeno chama-se ponderação de pistas auditivas e a investigação deste tema é considerada importante para compreendermos melhor como o desenvolvimento da percepção da fala ocorre.

Você está sendo convidado(a) a participar de uma pesquisa de percepção da fala, que tem como objetivo verificar se as crianças atribuem os mesmos pesos às pistas auditivas disponíveis nas palavras “sapa”, “chapa”, “bata”, “data”, “saba” e “chaba” do que os adultos para a identificação destas palavras do português brasileiro. Você foi selecionado, uma vez que apresenta idade entre 20 e 40 anos e sua participação não é obrigatória.

2. Procedimento do estudo

O procedimento deste estudo consiste em ouvir as palavras por meio de um fone de ouvido e responder qual palavra foi escutada dentre as possibilidades, (sapa)-(chapa), (saba)-(chaba) e (bata)-(data). Assim, figuras representando estas palavras serão colocadas em cada lado da tela do computador e, após percebê-las, você deverá indicá-las de acordo com as respostas presentes na tela do monitor.

3. Riscos e desconfortos

Consideramos que a metodologia utilizada para coleta de dados oferece muito pouco ou, eventualmente, nenhum prejuízo para você. Talvez possa haver cansaço durante a execução da tarefa. Desta forma, a coleta de dados será realizada com as pausas necessárias.

4. Benefícios

Consideramos que essa pesquisa não trará benefícios diretos para você. Porém, vale ressaltar, que para participação na mesma, será realizada uma triagem auditiva com o objetivo de descartar qualquer grau de perda auditiva. Diante da necessidade de realizar uma avaliação audiológica completa você será comunicado(a) e receberá as devidas orientações.

5. Custos/Reembolso

Você não terá nenhum gasto com a sua participação no estudo, uma vez que este será realizado na própria escola, ou seja, em seu local de trabalho. Além disso, não receberá pagamento pela participação no mesmo.

6. Caráter Confidencial dos Registros

A sua identidade será mantida em sigilo. Os resultados da pesquisa serão sempre apresentados como o retrato de um grupo e não de uma pessoa. Dessa forma, você não será identificado quando o material de seu registro for utilizado ou quando os seus resultados forem publicados em jornais, revistas e eventos científicos ou educativos. O material coletado ficará sob a guarda da coordenação do projeto e não será divulgado publicamente.

7. Participação

Sua participação neste estudo é muito importante e voluntária. Você tem o direito de não querer participar ou de sair deste estudo a qualquer momento. Se você decidir retirar-se do estudo, favor informar o pesquisador.

8. Informações

Este estudo foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais, que poderá ser contatado para esclarecimentos pelo telefone 3499-4592, por email coep@prpq.ufmg.br ou no seguinte endereço: Av. Antônio Carlos, 6627 – Unidade Administrativa II, sala 2005. CEP 31270-901 – Belo Horizonte, MG.

A pesquisadora responsável poderá fornecer qualquer esclarecimento sobre essa pesquisa, assim como tirar dúvidas, bastando contato no seguinte endereço e/ou telefone:

Nome do pesquisador: Fabiana Andrade Penido

Endereço: Rua Stella Hanriot, 211/801

Telefone: (31) 99641-4171

Email: fabianapenido@gmail.com

9. Declaração de Consentimento

Li ou alguém leu para mim as informações contidas neste documento antes de assinar este termo de consentimento. Declaro que toda a linguagem técnica utilizada na descrição deste estudo de pesquisa foi satisfatoriamente explicada e que recebi respostas para todas as minhas dúvidas. Confirmando também que recebi uma cópia deste Termo de Consentimento Livre e Esclarecido. Compreendo que sou livre para me retirar do estudo em qualquer momento, sem perda de benefícios ou qualquer outra penalidade.

Dou meu consentimento de livre e espontânea vontade para participar deste estudo.

Nome do participante

Assinatura do participante

Obrigado pela sua colaboração e por merecer sua confiança.

Fabiana Andrade Penido (pesquisadora)

_____/_____/_____
Data

ANEXO 8. Questionário direcionado aos responsáveis pelas crianças de quatro e sete anos de idade

Prezados pais ou responsável,

Meu nome é Fabiana Penido, sou fonoaudióloga formada desde 2004 pela PUCMinas, Mestre e Doutoranda em Estudos Linguísticos pela UFMG. Há 6 anos iniciei a investigação do desenvolvimento percepção da fala da criança.

Convido o seu(sua) filho(a) para participar deste estudo que realizo desde 2010. Informações mais detalhadas sobre o procedimento estão disponíveis no Termo de Consentimento Livre e Esclarecido, que será entregue a você.

Para a participação do seu(sua) filho(a) no estudo será necessário que você responda o questionário abaixo e assine a autorização. Vale ressaltar, que este questionário é confidencial e as informações obtidas serão mantidas em sigilo.

Conto com a sua colaboração e desde já agradeço! Coloco-me à disposição para esclarecer eventuais dúvidas.

Atenciosamente,

Fabiana Penido – Fonoaudióloga CRFa6/MG 4270. Contato: (31) 99641-4171.

Nome da criança: _____

Data de Nascimento: ____ / ____ / ____

Escolaridade: _____

1. Desenvolvimento da linguagem

1.1 Em que idade a criança:

1.1.1 Pronunciou as primeiras palavras: ____ anos ____ meses Não sabe ou não lembra

1.1.2 Disse as primeiras frases: ____ anos ____ meses Não sabe ou não lembra

1.2 A criança comunica-se através de: Gestos Palavras Palavras e gestos

1.3 As pessoas compreendem o que o seu(sua) filho(a) fala? Não Sim

1.4 Seu(sua) filho(a) apresenta dificuldade de produzir algum som da fala? Não Sim

Qual(is)? _____

2. Histórico auditivo

2.1 Seu(sua) filho(a) tem algum familiar que apresente perda da audição? Não Sim. Se sim, qual é o grau de parentesco? _____

2.2 Apresenta otites (infecções de ouvido) frequentes? Não Sim. Se sim, qual é a frequência? _____

Quantos episódios por ano? _____

2.3 Realizou alguma cirurgia no ouvido? Não Sim. Se sim, qual cirurgia e quando? _____

2.4 Faz uso de tubo de ventilação (carretel) no ouvido? Não Sim

2.5 Fala alto? Não Sim

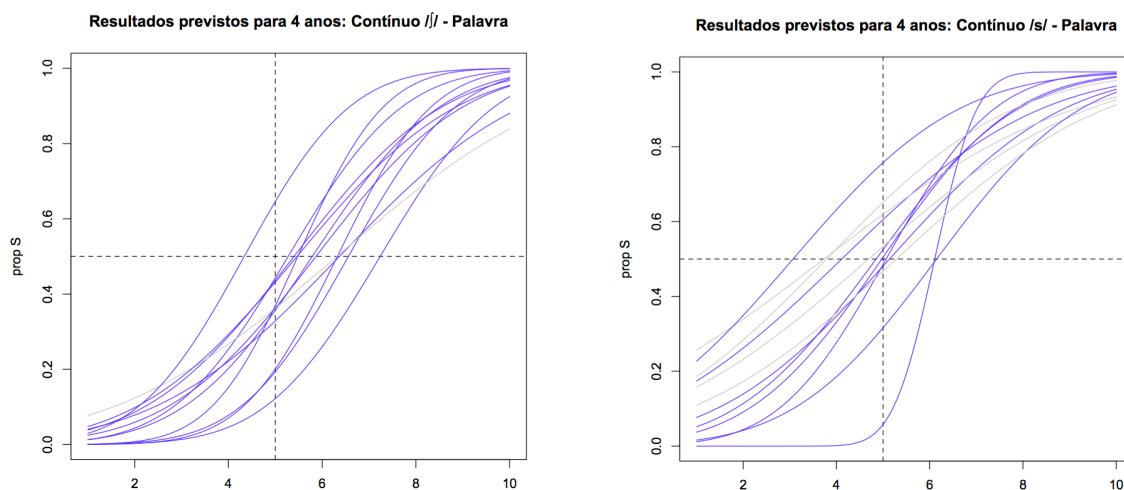
2.6 Pede para repetir o que lhe é falado? Não Sim Sempre Às vezes

2.7 Pede para aumentar o som da televisão, rádio? Não Sim Sempre Às vezes

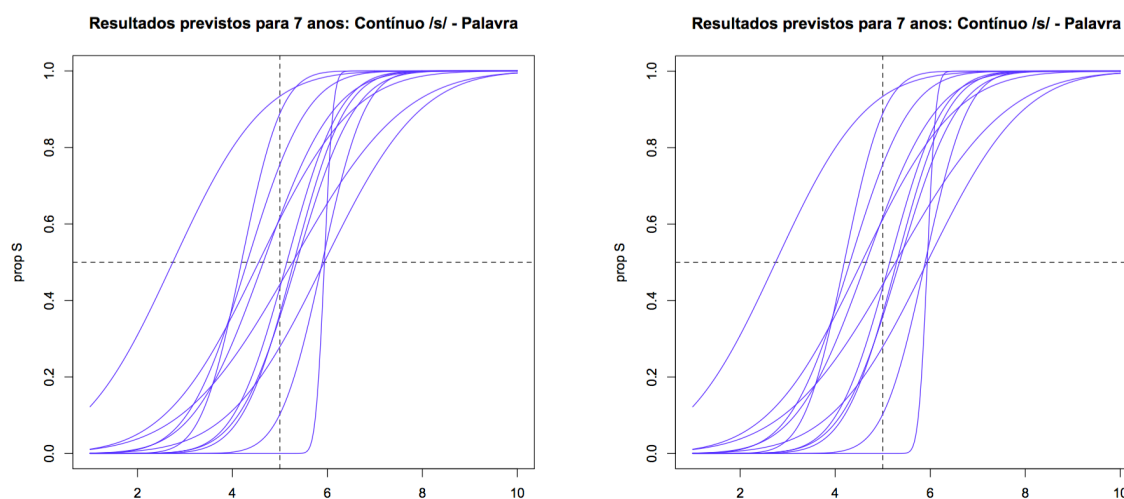
2.8 Já realizou exames auditivos? Não Sim. Se sim, quais? _____

Resultados: _____

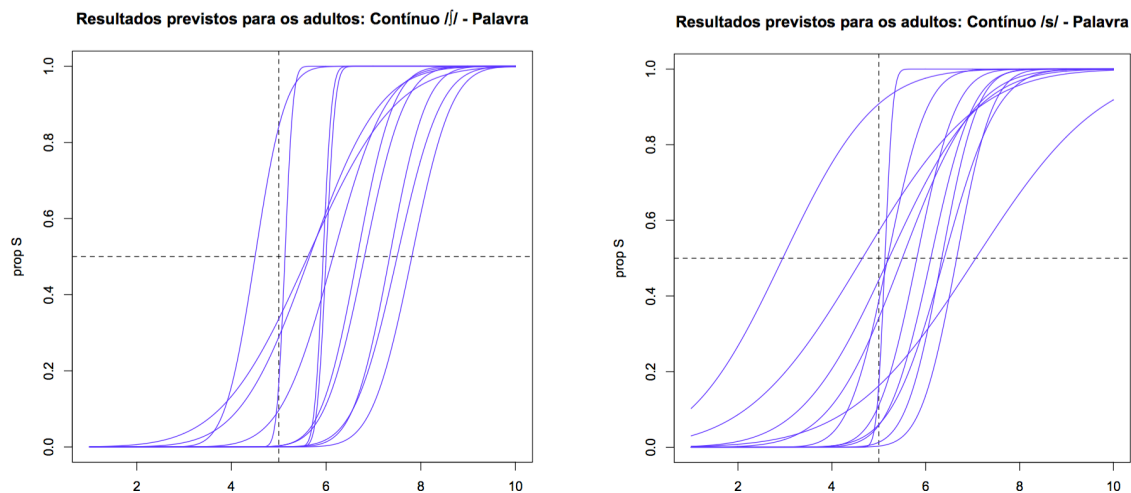
ANEXO 9. Curvas de classificação das crianças de quatro e sete anos e dos adultos para o par mínimo /ʃapa/-/sapa/. Os dez passos de frequência (Hz) do contínuo fricativo estão dispostos na abscissa e a probabilidade de respostas /sapa/ estão dispostas na ordenada. À esquerda observa-se as curvas de classificação do contínuo fricativo composto pelos valores de transição dos formantes apropriados para acompanhar [ʃ] e à direita observa-se as curvas de classificação do contínuo fricativo composto pelos valores de transição dos formantes apropriados para acompanhar [s]



Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

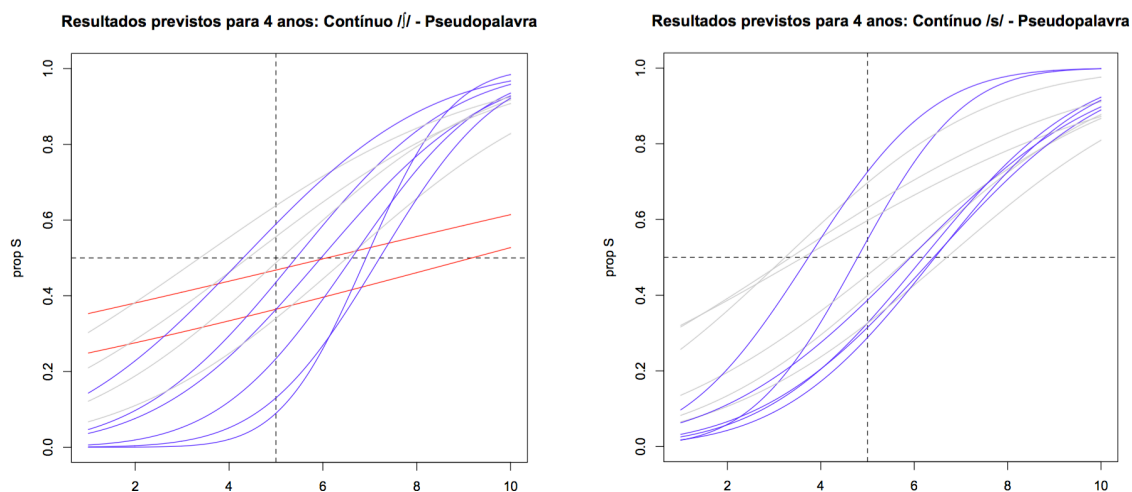


Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

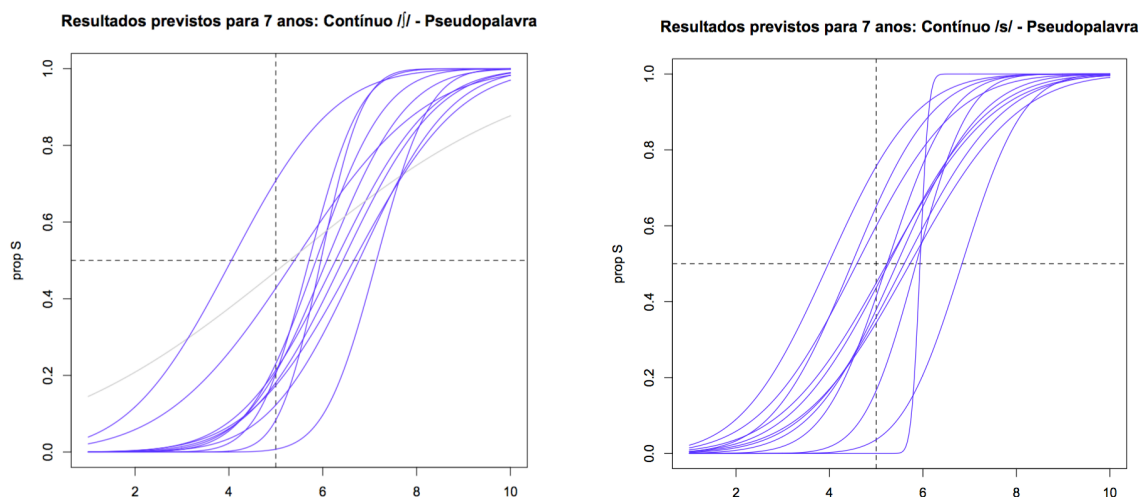


Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

ANEXO 10. Curvas de classificação das crianças de quatro e sete anos e dos adultos para o par de pseudopalavras /jaba/-/saba/. Os dez passos de frequência (Hz) do contínuo fricativo estão dispostos na abscissa e a probabilidade de respostas /sapa/ estão dispostas na ordenada. À esquerda observa-se as curvas de classificação do contínuo fricativo composto pelos valores de transição dos formantes apropriados para acompanhar [ʃ] e à direita observa-se as curvas de classificação do contínuo fricativo composto pelos valores de transição dos formantes apropriados para acompanhar [s]

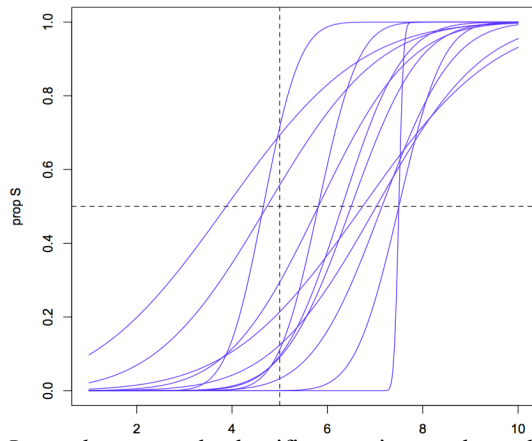


Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

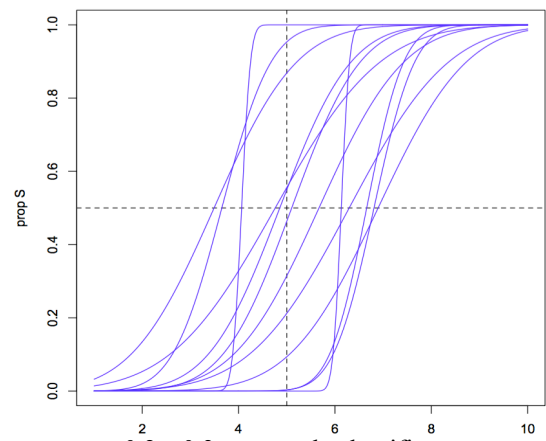


Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

Resultados previstos para os adultos: Contínuo /l/ - Pseudopalavra

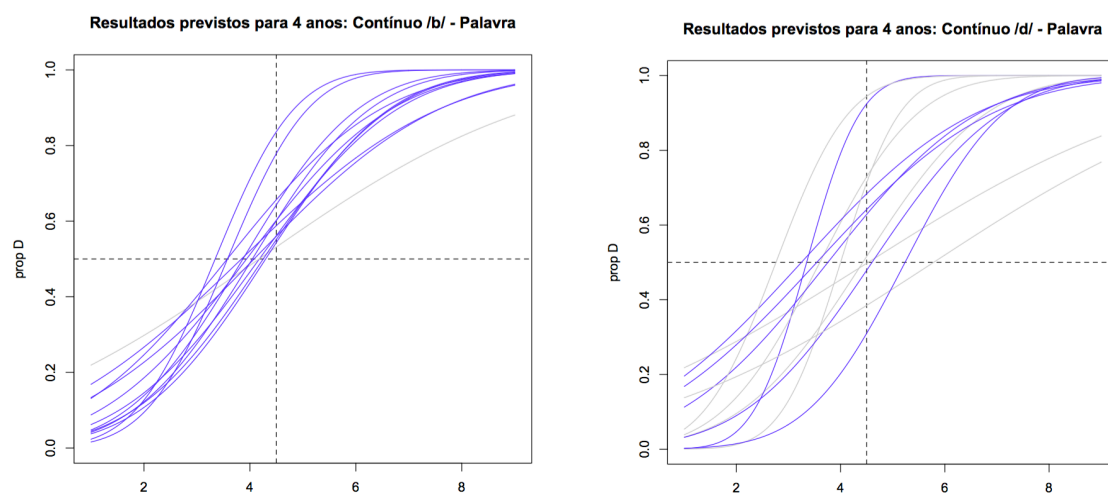


Resultados previstos para os adultos: Contínuo /s/ - Pseudopalavra

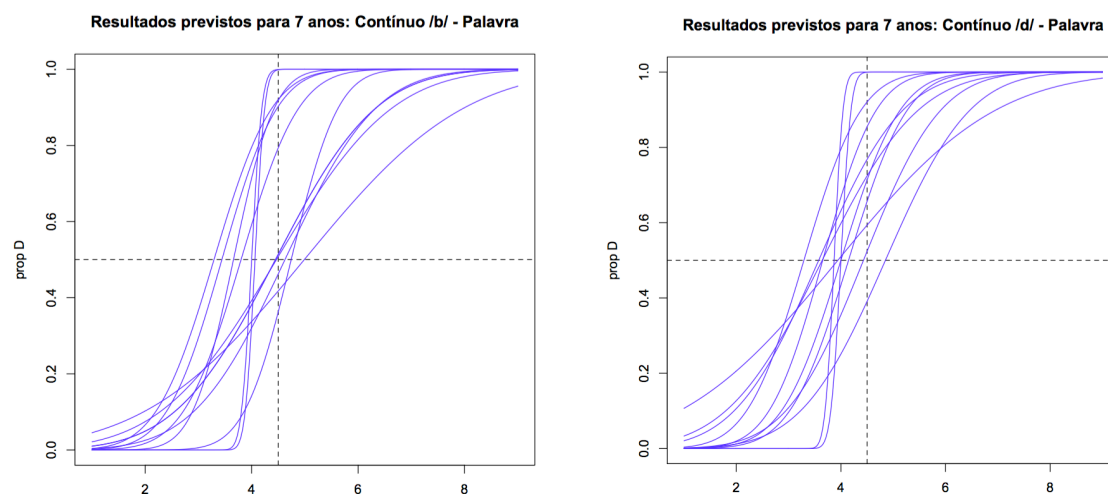


Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.

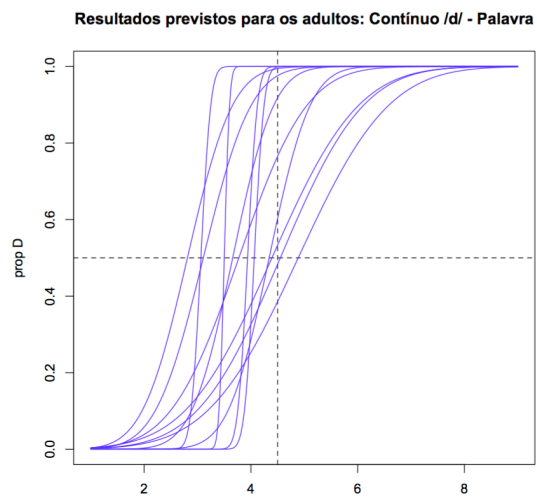
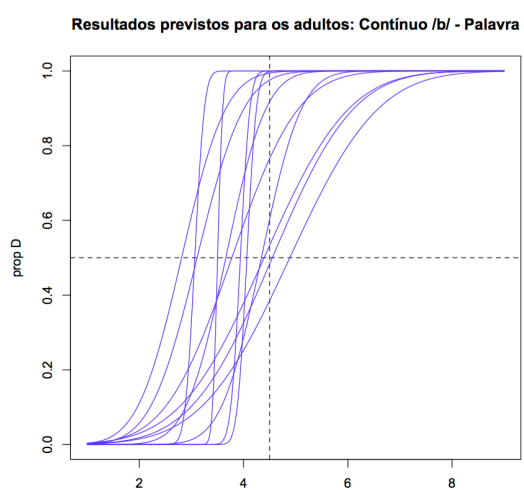
ANEXO 11. Curvas de classificação das crianças de quatro e sete anos e dos adultos para o par mínimo /bata-/data/. Os nove passos de frequência do segundo formante (F2) do contínuo vocálico estão dispostos na abscissa e a probabilidade de respostas /data/ estão dispostas na ordenada. À esquerda observa-se as curvas de classificação do contínuo vocálico composto pelos valores das amplitudes do *burst* apropriadas para acompanhar [b] e à direita observa-se as curvas de classificação do contínuo vocálico composto pelos valores das amplitudes do *burst* apropriadas para acompanhar [d]



Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.



Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.



Legenda: curvas de classificação cinza: valores de inclinação entre 0.2 e 0.3; curvas de classificação vermelha: valores de inclinação inferiores a 0.2; curvas de classificação azul: valores de inclinação superiores a 0.3.