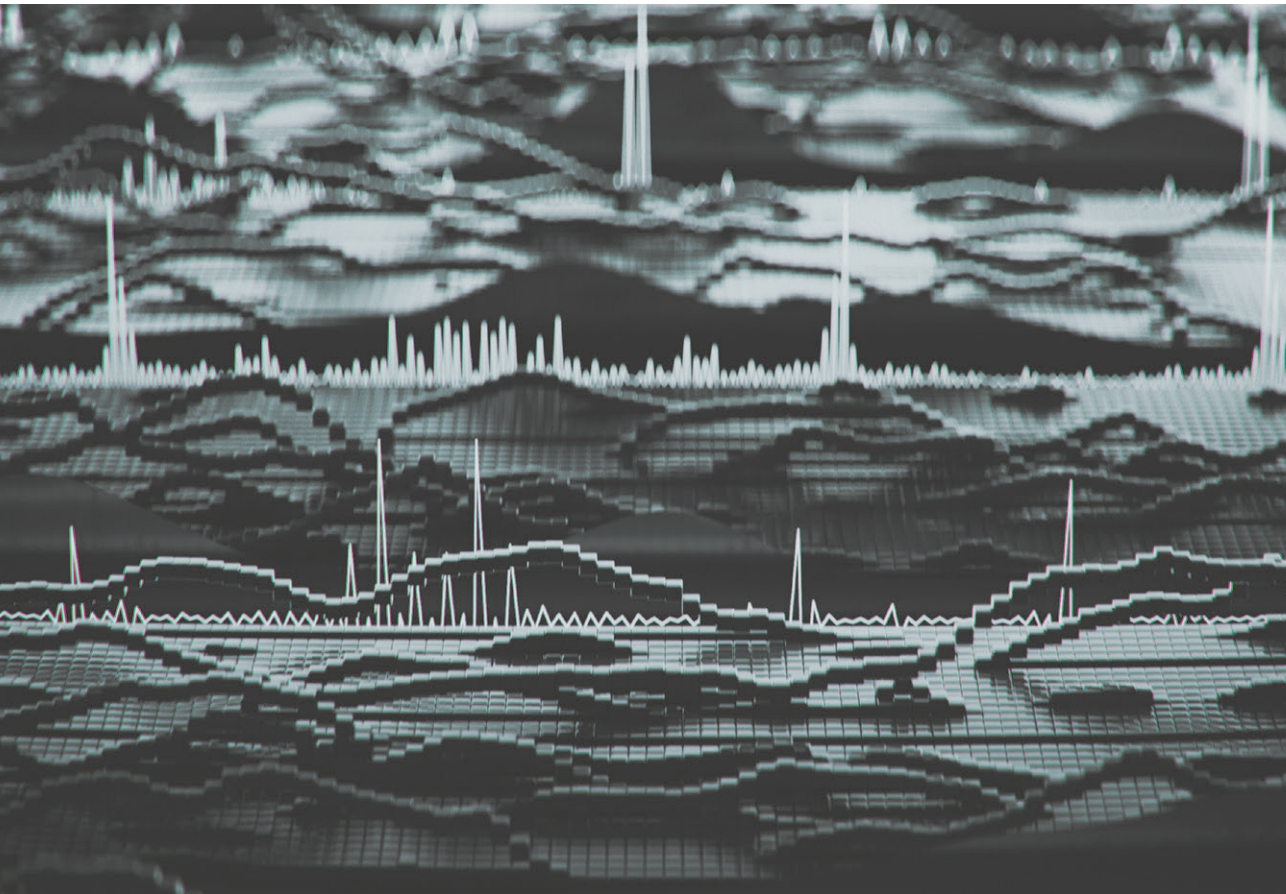


Januacele da Costa  
Miguel Oliveira Jr  
organizadores

# ESTUDOS EM FONÉTICA E FONOLOGIA



ESTUDOS EM FONÉTICA  
E FONOLOGIA

*CONSELHO EDITORIAL*

André Costa e Silva

Cecilia Consolo

Dijon de Moraes

Jarbas Vargas Nascimento

Luis Barbosa Cortez

Marco Aurélio Cremasco

Rogério Lerner

JANUACELE DA COSTA  
MIGUEL OLIVEIRA JR  
(organizadores)

ESTUDOS EM FONÉTICA  
E FONOLOGIA

2021

*Estudos em Fonética e Fonologia*

© 2021 Januacele da Costa, Miguel Oliveira Jr  
Editora Edgard Blücher Ltda.

*Publisher* Edgard Blücher

*Editor* Eduardo Blücher

*Coordenação editorial* Jonatas Eliakim

*Produção editorial* Aline Fernandes

*Diagramação e capa* Laércio Flenic

*Revisão de texto* Samira Panini

*Imagem da capa* iStockphoto

---

# Blucher

Rua Pedroso Alvarenga, 1245, 4º andar  
04531-934 – São Paulo – SP – Brasil  
Tel 55 11 3078-5366  
contato@blucher.com.br  
www.blucher.com.br

Segundo Novo Acordo Ortográfico, conforme 5. ed.  
do Vocabulário Ortográfico da Língua Portuguesa,  
Academia Brasileira de Letras, março de 2009.

É proibida a reprodução total ou parcial por quaisquer  
meios, sem autorização escrita da Editora.

---

Todos os direitos reservados pela Editora  
Edgard Blücher Ltda.

Dados Internacionais de Catalogação na Publicação (CIP)  
Angélica Ilacqua CRB-8/7057

---

Estudos em Fonética e Fonologia / Januacele da  
Costa, Miguel Oliveira Jr -- São Paulo: Blucher, 2021  
326p.

Bibliografia  
ISBN 978-65-5550-093-6 (impresso)  
ISBN 978-65-5550-092-9 (eletrônico)

Open Access

1. Fonética 2. Fonologia I. Título II. Costa, Januacele  
da III. Oliveira Júnior, Miguel

---

21-2456

CDD 414

---

Índices para catálogo sistemático:  
1. Fonética

---

# SUMÁRIO

<b>1. ANÁLISE ACÚSTICA DAS VOGAIS PRODUZIDAS POR FALANTES COM DOENÇA DE PARKINSON .....</b>	<b>11</b>
1. INTRODUÇÃO.....	11
2. METODOLOGIA.....	14
2.1 MEDIDAS ACÚSTICAS.....	16
2.2 TRATAMENTO ESTATÍSTICO.....	17
3. RESULTADOS E DISCUSSÃO.....	17
3.1. FREQUÊNCIA FUNDAMENTAL.....	17
3.2. INTENSIDADE.....	19
3.3 FORMANTES VOCÁLICOS, TRIÂNGULO E ÁREA DO ESPAÇO VOCÁLICO .....	21
3.4 PRISMA ACÚSTICO-ARTICULATÓRIO .....	24
3.5 ÍNDICES.....	26
4. CONSIDERAÇÕES FINAIS.....	28
REFERÊNCIAS .....	29
<b>2. A MENTIRA TEM PERNA CURTA? PISTAS PROSÓDICAS NA IDENTIFICAÇÃO DE DISCURSO ENGANOSO.....</b>	<b>35</b>
1. INTRODUÇÃO.....	35
2. ESTADO DA ARTE.....	37
2.1. A MENTIRA .....	37
2.2. A PROSÓDIA.....	39
3. METODOLOGIA.....	42
3.1. PERFIL DOS PARTICIPANTES .....	42
3.2. <i>CORPUS</i> .....	42
3.3. PROCEDIMENTOS DE COLETA DE DADOS.....	43
3.4. ANÁLISES.....	44
4. RESULTADOS.....	45

4.1. LATÊNCIA.....	45
4.2. PAUSA.....	46
5. DISCUSSÕES.....	50
6. CONCLUSÃO.....	52
REFERÊNCIAS.....	53
<b>3. CORRELATOS ELETROFISIOLÓGICOS DO PROCESSAMENTO DE FRONTEIRA PROSÓDICA NO PORTUGUÊS BRASILEIRO.....</b>	<b>59</b>
1. INTRODUÇÃO.....	59
2. MÉTODO.....	63
2.1 MATERIAL.....	63
2.2 ESTÍMULOS: DESCRIÇÃO GERAL E ANÁLISE ACÚSTICA.....	65
2.3 PARTICIPANTES.....	68
2.4 PARADIGMA EXPERIMENTAL DE EEG.....	69
2.5 AQUISIÇÃO DOS SINAIS DIGITAIS DE EEG.....	70
2.6 EXTRAÇÃO DE ERPS (POTENCIAIS RELACIONADOS A EVENTO).....	70
2.7 A ANÁLISE DE ERPS.....	72
3. RESULTADOS.....	73
4. DISCUSSÃO.....	78
5. CONCLUSÃO.....	82
REFERÊNCIAS.....	82
<b>4. ASPECTOS PROSÓDICOS DA LIBRAS EM NÚMEROS DE TELEFONE MÓVEL...91</b>	<b>91</b>
1. INTRODUÇÃO.....	92
2. PROCEDIMENTOS METODOLÓGICOS.....	94
3. RESULTADOS E DISCUSSÃO.....	97
3.1 DISTRIBUIÇÃO DE FREQUÊNCIA NUMÉRICA E DECIMAL.....	97
3.2 PARÂMETROS PROSÓDICOS DURACIONAIS.....	100
4 CONCLUSÕES.....	106
REFERÊNCIAS.....	107
<b>5. PARADIGMA DO MUNDO VISUAL: MÉTODO DE RASTREAMENTO OCULAR.....</b>	<b>111</b>
1. INTRODUÇÃO.....	112

<b>2. CUIDADOS METODOLÓGICOS PARA EXPERIMENTOS COM O VWP.....</b>	<b>115</b>
2.1. ELABORAÇÃO E SELEÇÃO DE SENTENÇAS.....	115
2.2. PREPARAÇÃO DE ESTÍMULOS AUDITIVOS.....	120
2.3. PREPARAÇÃO DE ESTÍMULOS VISUAIS.....	122
<b>3. RECOMENDAÇÕES FINAIS .....</b>	<b>130</b>
<b>REFERÊNCIAS .....</b>	<b>131</b>
<b>6. SEGMENTAÇÃO AUTOMÁTICA E MANUAL DE NARRATIVAS ESPONTÂNEAS: METODOLOGIA DE UM ESTUDO DE PERCEPÇÃO DA PROSÓDIA.....</b>	<b>135</b>
1. INTRODUÇÃO.....	135
2. NARRATIVAS ESPONTÂNEAS.....	136
3. MONTAGEM DO EXPERIMENTO DE PERCEPÇÃO.....	137
4. SEGMENTAÇÃO MANUAL .....	139
5. SEGMENTAÇÃO AUTOMÁTICA.....	144
5.1 GRUPOS ACENTUAIS .....	145
5.2 UNIDADES VV .....	147
6. CONSIDERAÇÕES FINAIS.....	148
REFERÊNCIAS .....	149
<b>7. UMA BREVE INTRODUÇÃO AOS ASPECTOS TEÓRICO-METODOLÓGICOS RELACIONADOS À COMPARAÇÃO FORENSE DE LOCUTOR.....</b>	<b>151</b>
1. INTRODUÇÃO.....	151
2. A FONÉTICA FORENSE.....	151
3. QUESTÕES TERMINOLÓGICAS.....	154
4. FATORES E VARIÁVEIS.....	155
4.1 TRANSMISSÃO E QUALIDADE DA GRAVAÇÃO .....	156
4.2 ESTILO DE FALA.....	160
4.3 SOBRE A ESPONTANEIDADE NA FALA .....	161
4.4 SOBRE A REPRESENTATIVIDADE DOS DADOS.....	164
5. CRITÉRIOS PARA A SELEÇÃO DE PARÂMETROS FONÉTICO-ACÚSTICOS ....	166
REFERÊNCIAS .....	169
<b>8. RELAÇÕES DE SENTIDO EM CONSTRUÇÕES GRAMATICAIS: HOMOMÍMIA, POLISSEMIA E SINONÍMIA EM TIRAS DA MAFALDA .....</b>	<b>173</b>
RESUMO .....	173



1. INTRODUÇÃO .....	174
2. AS TIRAS DE MAFALDA COMO OBJETO DE ESTUDO E APRENDIZAGEM... 176	
3. OBSERVANDO ALGUMAS TIRAS DE MAFALDA..... 177	
4. ALGUMAS CONSIDERAÇÕES FINAIS..... 185	
REFERÊNCIAS .....	186
<b>9. A RELAÇÃO DAS MUDANÇAS LINGÜÍSTICAS HISTÓRICAS COM OS ERROS FONÉTICO-FONOLÓGICOS DO SISTEMA VOCÁLICO NO ENSINO-APRENDIZAGEM DE ESPANHOL/LE POR BRASILEIROS..... 189</b>	
1. INTRODUÇÃO .....	189
2. ESTUDOS REALIZADOS .....	191
3. INTERFERÊNCIAS VOCÁLICAS..... 193	
3.1. ABERTURA DAS VOGAIS MÉDIAS..... 193	
3.2. HARMONIA VOCÁLICA..... 196	
3.3. NEUTRALIZAÇÃO DAS VOGAIS ÁTONAS EM POSIÇÃO FINAL .....	198
3.4. NASALIZAÇÃO DAS VOGAIS..... 200	
3.5. ALONGAMENTO VOCÁLICO .....	203
4. CONCLUSÕES..... 205	
REFERÊNCIAS .....	206
<b>10. EXPLICIT INSTRUCTION OF POST-LEXICAL PROCESSES TO AID LISTENING COMPREHENSION IN THE TEACHING OF ENGLISH AS AN ADDITIONAL LANGUAGE .....</b>	<b>211</b>
1. INTRODUCTION .....	211
2. CONNECTED SPEECH .....	213
2.1 RHYTHM AND INTONATION..... 214	
2.2 WORD AND SENTENCE STRESS .....	215
2.3 POST-LEXICAL PROCESSES..... 216	
3. LISTENING COMPREHENSION .....	217
4. METHODOLOGY .....	220
5. RESEARCH RESULTS..... 224	
5.1 RESEARCH CONCLUSIONS..... 228	
REFERENCES .....	232

<b>11. ANÁLISE VARIACIONISTA DE /R/ NA CODA SILÁBICA MEDIAL NO PORTUGUÊS FALADO EM ALAGOAS.....</b>	<b>235</b>
1. INTRODUÇÃO.....	235
2. A COMPLEXIDADE DOS RÓTICOS.....	237
3. PROCEDIMENTOS METODOLÓGICOS.....	241
3.1 CONSTITUIÇÃO DA AMOSTRA.....	242
3.2 COLETA E TRANSCRIÇÃO DOS DADOS.....	243
3.3 TRATAMENTO DOS DADOS.....	245
3.4 A VARIÁVEL DEPENDENTE E SUAS VARIANTES.....	245
3.5 VARIÁVEIS INDEPENDENTES.....	246
3.6 ANÁLISE ESTATÍSTICA DOS DADOS.....	247
4. RESULTADOS E DISCUSSÃO.....	247
4.1 ANÁLISE DA VARIAÇÃO FRICATIVA GLOTA [h] <i>VERSUS</i> APROXIMANTE [j] E FENÔMENO RELACIONADO.....	251
5. CONCLUSÕES.....	262
REFERÊNCIAS.....	264
<b>12. ORGANIZAÇÃO DAS UNIDADES FONOLÓGICAS NO INÍCIO DA AQUISIÇÃO DE LINGUAGEM SOB A PERSPECTIVA DA FONOLOGIA PROSÓDICA E A AUTOSSEGMENTAL.....</b>	<b>267</b>
1. INTRODUÇÃO.....	267
2. O ACENTO ENTONACIONAL E AS ANÁLISES CENTRÍPETA E CENTRÍFUGA.....	270
3. A GEOMETRIA DE TRAÇOS E OS PREENCHIMENTOS SEGMENTAIS DA SÍLABA NA AQUISIÇÃO.....	275
4. DISCUSSÃO.....	278
5. CONCLUSÕES.....	280
REFERÊNCIAS.....	280
<b>13. A PALAVRA NOMINAL EM YAATHE (FULNI-Ô), LÍNGUA INDÍGENA BRASILEIRA.....</b>	<b>283</b>
1. INTRODUÇÃO.....	283
1.1 A LÍNGUA E SEUS FALANTES.....	283
1.2 ESTUDOS SOBRE A LÍNGUA: O ESTADO DA ARTE.....	285
1.3 OBJETIVOS DO TRABALHO.....	285

2. APORTE TEÓRICO.....	286
3. PROCEDIMENTOS METODOLÓGICOS.....	287
4. A PALAVRA FONOLÓGICA.....	290
4.1 A PALAVRA COMO UNIDADE LINGUÍSTICA.....	290
4.2 A PALAVRA NOMINAL FONOLÓGICA EM YAATHE.....	293
5. CONCLUSÃO.....	305
REFERÊNCIAS.....	306
<b>14. E AGORA, JOSÉ? VOGAIS NASAIS FONOLÓGICAS EM YAATHE EXISTEM? ....</b>	<b>309</b>
1. INTRODUÇÃO.....	309
2. PROCEDIMENTOS METODOLÓGICOS.....	311
3. O COMPORTAMENTO DO TRAÇO NASAL EM YAATHE.....	313
3.1 INFORMAÇÕES PRELIMINARES.....	313
3.2 VOGAIS NASALIZADAS.....	315
3.3 VOGAIS NASAIS.....	320
4. CONCLUSÃO.....	323
REFERÊNCIAS.....	324

## ANÁLISE ACÚSTICA DAS VOGAIS PRODUZIDAS POR FALANTES COM DOENÇA DE PARKINSON

*Susana de Carvalho<sup>1</sup>  
Miguel Oliveira Jr.<sup>2</sup>*

### 1. INTRODUÇÃO

A doença de Parkinson (DP) é uma doença degenerativa e progressiva, caracterizada pela perda de neurônios na parte compacta da substância negra e consequente prejuízo na produção da dopamina, responsável pelo controle dos movimentos (BRASOLOTTO et al., 2014).

Os sinais e sintomas da DP são, principalmente, motores: rigidez, bradicinesia, tremor e instabilidade postural. As dificuldades motoras prejudicam também a produção da fala interferindo negativamente na comunicação verbal dos indivíduos. A prevalência de alterações de fala é alta na DP, atingindo cerca de 90% dos casos em algum estágio da doença e, infelizmente, as intervenções farmacológicas produzem efeitos mínimos sobre essas dificuldades na fala (BRASOLOTTO et al., 2014; MARTÍNEZ-SANCHÉZ, 2010; SAPIR; RAMIG; FOX, 2008; RUSZ et al., 2017).

---

<sup>1</sup> Universidade Federal de Sergipe.

<sup>2</sup> Universidade Federal de Alagoas.

Mudanças na voz e na fala podem ser notadas por familiares e amigos próximos das pessoas com DP, vários anos antes que o diagnóstico seja estabelecido. Recente investigação sobre os sinais prodrômicos da DP revelou que a acinesia de face e as alterações vocais foram os primeiros indicadores da doença, com um intervalo estimado de 9,8 anos antes do diagnóstico (POSTUMA et al., 2012; RUSZ et al., 2013; SAPIR; SPRECHER; SKODDA, 2013).

No início, essas alterações na fala podem ser bastante sutis e quase imperceptíveis aos ouvidos, mas medidas acústicas objetivas têm contribuído para o registro de anormalidades, mesmo em pacientes sem disartria visível (FORREST; WEISMER, 2009; LANSFORD; LISS, 2014a; RUSZ et al., 2013).

Um estudo longitudinal, que acompanhou pacientes com DP durante 36 meses, demonstrou que a maioria dos sintomas motores puderam ser controlados e mantiveram-se estáveis com medicação, enquanto a articulação das vogais teve uma deterioração significativa com o passar do tempo, concluindo que as medidas acústicas das vogais teriam potencial para servir como marcadores do avanço da doença (SKODDA; GRÖNHEIT; SCHLEGEL, 2012).

Durante a emissão das vogais, a posição dos articuladores resulta em cavidades de ressonância que amplificam determinadas frequências. Tais frequências, típicas para cada vogal, são denominadas formantes (ROY et al., 2009; SKODDA; GRÖNHEIT; SCHLEGEL, 2012).

Os valores das frequências do primeiro e segundo formantes (F1 e F2), expressas em Hertz (Hz), são as medidas acústicas mais relevantes na percepção e produção das vogais. A relação acústico-articulatória pode ser explicada por uma regra mais geral na qual os valores de F1 variam em função da altura da língua (no eixo vertical) e os valores de F2 variam em função da anterioridade ou posterioridade da língua (no eixo horizontal) (KENT; READ, 2015; SAPIR et al., 2010).

Na DP, a produção dos sons da fala é prejudicada pela rigidez e/ou bradicinesia, caracterizando-se por uma diminuição na amplitude e velocidade dos movimentos de mandíbula, língua e lábios e resultando em uma restrição na geração dos formantes (SKODDA; GRÖNHEIT; SCHLEGEL, 2012).

Hipotetiza-se que, em uma fala disártrica como ocorre na DP, formantes que normalmente apresentam frequências altas tendem a apresentar valores baixos, enquanto frequências baixas tendem a apresentar valores mais altos (SAPIR et al., 2007).

A fim de identificar a fala disártrica, característica da DP, os valores absolutos dos formantes F1 e F2 têm sido utilizados em cálculos e configurações diversas, como o triângulo acústico vocálico, a área do espaço vocálico (do original *vowel space area* – VSA), o índice de centralização formântica (do original *formant centralization ratio* – FCR), dentre outros:

- a) **Triângulo acústico vocálico e Área do espaço vocálico:** O triângulo vocálico é uma representação visual da relação acústico-articulatória. Configura-se por meio de um gráfico *xy*, com os valores de F1 na ordenada e F2 na abscissa. Assim, as coordenadas (F2a, F1a), (F2i, F1i) e (F2u, F1u) definem os vértices do triângulo. A área do espaço vocálico, por sua vez, consiste na distância euclidiana entre esses vértices e pode ser obtida por meio da expressão:  $VSA = \text{ABS}((F1i*(F2a-F2u)+F1a*(F2u-F2i)+F1u*(F2i-F2a))/2)$ , como proposto por Liu, Tsao e Kuhl (2005).

A limitação de movimentos observada na DP pode levar a constrição do triângulo e uma redução da área do espaço vocálico. Os estudos que utilizaram essas medidas, entretanto, nem sempre conseguiram revelar diferenças significativas entre falantes saudáveis e falantes com DP (BANG et al., 2013; OKADA; MURATA; TODA, 2015; OROZCO-ARROYAVE et al., 2014; SAPIR et al., 2010; SKODDA; VISSER; SCHLEGEL, 2011; TJADEN et al., 2013).

- b) **Índice de centralização formântica e Índice de articulação vocálica:** são medidas que pretendem superar as inconsistências encontradas com o uso da área do espaço vocálico – aparentemente muito sensível a variações interfalantes – e representar a retração do espaço acústico por meio de um único valor numérico. O índice de centralização formântica (tradução livre do original *formant centralization ratio*) pode ser calculado com a expressão:  $(F2u+F2a+F1i+F1u)/(F2i+F1a)$ . O índice de articulação vocálica (tradução do original *vowel articulation index*) pode ser calculado por meio da expressão:  $(F2i+F1a)/(F2u+F2a+F1i+F1u)$ , que é o inverso da anterior. Tais medidas foram propostas por Sapir et al. (2010) e Skodda, Visser e Schlegel (2011), respectivamente.

Para Sapir et al. (2010) e Skodda, Visser e Schlegel (2011), o índice de centralização formântica e o índice de articulação vocálica mostraram-se mais sensíveis que o triângulo acústico vocálico, na detecção de uma fala disártrica.

- c) **Prisma vocálico (*vPrism*):** Em 2014, Orozco-Arroyave et al. propuseram uma representação tridimensional da relação acústico-articulatória, com a

adição do eixo  $z$  ao plano  $xy$  da área do espaço vocálico. Em seu modelo, a base do prisma é o triângulo acústico vocálico e sua altura (eixo  $z$ ) expressa a frequência fundamental ( $F_0$ ) das vogais /i/, /a/ e /u/. Assim, cada ponto do gráfico é definido pela coordenada ( $F_2, F_1, F_0$ ). A inclusão da frequência fundamental, para os autores, possibilita a avaliação simultânea da voz e da fala e demonstrou uma acurácia de 81,3% na identificação de falantes com DP, superior ao índice de 79,3% quando apenas medidas relacionadas com a articulação foram utilizadas (OROZCO-ARROYAVE et al., 2014).

Mas a inclusão da frequência fundamental pode ser questionada quando o que se pretende é avaliar a articulação da fala. O estudo longitudinal de Skodda et al. (2013), que acompanhou 80 pessoas com DP, observou a deterioração progressiva da voz e da fala no curso da doença e evidenciou que as medidas de *shimmer* (medida de perturbação na amplitude, ou intensidade, durante a emissão vocal) – e não de *jitter* (medida de perturbação na frequência fundamental, durante a emissão vocal) – eram significativamente mais altas na DP, desde os estágios iniciais da doença.

É provável que os valores de intensidade representem uma medida mais robusta para diferenciar a fala na DP e propõe-se a sua inclusão nos parâmetros já existentes por meio de um novo índice, o Índice de Articulação e Intensidade Vocálica, e da construção de um **prisma acústico-articulatório**, similar ao de Orozco-Arroyave et al. (2014), mas que contemple os valores de intensidade (dB) em seu eixo  $z$ . Sabe-se que há desafios a superar, pois trata-se de uma medida acústica sujeita a interferências, mas acredita-se que avanços só serão possíveis se esses obstáculos forem enfrentados.

Assim, o principal objetivo deste estudo foi investigar as medidas acústicas das vogais produzidas por falantes com doença de Parkinson. Pretende-se também propor novos parâmetros acústicos para avaliação e determinar a eficácia das medidas de intensidade na identificação da fala disártrica.

## 2. METODOLOGIA

Trata-se de um estudo descritivo, observacional e analítico, no qual foram comparadas as medidas acústicas das vogais produzidas por pessoas com doença de Parkinson e falantes neurologicamente saudáveis. Está em consonância com a Resolução CNS 466/12 e foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Sergipe sob o nº CAAE 48290415.6.0000.5546/Parecer nº 1.217.645.

A pesquisa contou com 18 participantes, nove homens e nove mulheres, com idades entre 41 e 71 (IM: 59,4; DP: 9,4), distribuídos em dois grupos: GParkinson, constituído por doze participantes (seis homens e seis mulheres) com doença de Parkinson (IM: 63,9; DP: 7,9), e o GControle, com seis participantes (três homens e três mulheres), neurologicamente saudáveis (IM: 51; DP: 5,9). A fim de classificar o grau de comprometimento da fala, foi utilizada a versão para língua portuguesa da MDS-UPDRS, parte III, Fala (GOETZ et al., 2008).

Para os participantes do GParkinson, o tempo decorrido desde o diagnóstico da doença variou entre dois e quatorze anos – média: 7,5 anos (DP: 3,9). Quanto ao comprometimento da fala, a maioria (dez) apresentou grau 1 ou 2 (leve ou suave).

Observa-se uma diferença significativa para a variável idade ( $p=0,005$ ) entre os grupos, mas esta não apresentou correlação significativa com as medidas acústicas analisadas. O perfil sociodemográfico dos participantes é apresentado no Quadro 1.

Todos foram convidados a gravar três frases-veículo: “Digo **rito** baixinho”; “Digo **rato** baixinho” e “Digo **rude** baixinho”; como proposto por Miranda e Meireles (2012). As gravações foram realizadas em sala acusticamente tratada, com o participante sentado em cadeira estofada, com encosto alto e suporte para cabeça, observada uma distância de 15 cm entre a comissura labial do falante e o microfone.

O microfone utilizado foi um microfone profissional TSI, modelo PROBR, com alta relação sinal/ruído e padrão polar supercardioide, instalado em pedestal e acoplado a um conversor e amplificador de áudio Shure®, o que possibilitou que as amostras de fala fossem captadas, digitalizadas e armazenadas no formato *wav* (monocanal, taxa de amostragem = 44.100 Hz; quantização = 16 bits). O software empregado, tanto para a gravação quanto para a análise posterior, foi o Praat (software de livre acesso, desenvolvido por Boersma e Weenink, em sua versão 6.0.25, 2017).



Quadro 1 – Perfil sociodemográfico dos participantes

Part	Grupo	Sexo	Idade	Tempo de curso da doença (anos)	UPDRS III Fala
01	GParkinson	F	58	9	1
02	GParkinson	M	71	8	2
03	GControle	M	56	--	0
04	GControle	F	53	--	0
05	GParkinson	M	70	10	2
06	GParkinson	M	65	5	1
07	GParkinson	F	66	5	1
08	GParkinson	F	58	6	1
09	GParkinson	M	68	8	2
10	GParkinson	M	66	6	3
11	GControle	M	51	--	0
12	GParkinson	F	43	2	1
13	GControle	M	41	--	0
14	GParkinson	F	62	10	1
15	GParkinson	F	70	14	4
16	GParkinson	M	57	7	1
17	GControle	F	48	--	0
18	GControle	F	57	--	0

Fonte: Dados da pesquisa, Maceió, 2017/2018.

## 2.1 Medidas acústicas

A partir das gravações de cada participante, foram obtidas as medidas de frequência fundamental, formantes F1 e F2 (em Hz) e intensidade (em dB), com o cursor posicionado no ponto médio de cada uma das vogais e realizadas automaticamente pelo software.

Com os valores absolutos dos formantes F1 e F2 das vogais, foi elaborado o triângulo vocálico e calculada a área do triângulo (ou espaço) vocálico (VSA – *vowel space area*), por meio da expressão  $ABS((F1i*(F2a-F2u)+F1a*(F2u-F2i)+F1u*(F2i-F2a))/2)$ , proposta por Liu, Tsao e Kuhl (2005). Com a adição das medidas de intensidade das vogais, foi elaborado o prisma acústico-articulatório e calculado seu volume, com a expressão:

$$ABS(intensidadei*(F2a*F1u-F2u*F1a)+intensidadea*(F2u*F1i-F2i*F1u)+intensidadeu*(F2i*F1a-F2a*F1i))/2$$

Ainda com os valores absolutos dos formantes F1 e F2 das vogais examinadas, foram calculados os índices de articulação vocálica, utilizando-se as fórmulas apresentadas no Quadro 2.

Quadro 2 – Índices de articulação vocálica

Índice	Expressão
FCR (formant centralization ratio)	$(F2u+F2a+F1i+F1u)/(F2i+F1a)$
VAI (vowel articulation index)	$(F2i+F1a)/(F2u+F2a+F1i+F1u)$

Fonte: Sapir et al., 2010; Skodda, Visser e Schlegel, 2011.

Com os valores absolutos dos formantes F1 e F2 e intensidade das vogais /i/, /a/ e /u/, foi calculado o índice de articulação e intensidade vocálica. Trata-se de uma replicação da expressão utilizada para o cálculo do índice de articulação vocálica (Skodda, Visser e Schlegel, 2011), com o acréscimo da média das medidas de intensidade.

Quadro 3 – Novo índice de articulação vocálica proposto

Novo índice proposto	Expressão
IAIV (índice de articulação e intensidade vocálica)	$[(F2i+F1a)/(F2u+F2a+F1i+F1u)]*[MÉDIA \text{ intensidade /a/; intensidade /i/; intensidade /u/}]$

## 2.2 Tratamento estatístico

O tratamento estatístico dos dados foi realizado com o software R (R Core Team, 2018), por meio do pacote RCommander (Fox; Bouchet-Valat, 2018) e do plug-in RcmdrPlugin.EZR (Kanda, 2013). A distribuição dos dados foi analisada por meio do teste Shapiro-Wilk e, nos casos de distribuição não normal, estes foram normalizados por meio de transformação logarítmica. Para comparação entre os grupos, utilizou-se o teste t-Student para amostras independentes. Em todas as situações, a significância foi estabelecida para  $p \leq 0,05$ .

## 3. RESULTADOS E DISCUSSÃO

### 3.1. Frequência fundamental

A frequência fundamental é um parâmetro relevante na avaliação funcional da laringe. Características biodinâmicas e a pressão subglótica determinam a frequência fundamental (F0), que corresponde ao número de ciclos vibratórios

que as pregas vocais realizam por segundo. A F0 varia de acordo com a idade e o sexo, com uma distribuição média entre 80 a 250 Hz, nos indivíduos adultos. Para os homens, encontram-se valores situados entre 80 a 150 Hz e, para mulheres, entre 150 a 250 Hz (BEHLAU et al., 2008).

Uma característica importante da frequência fundamental é o seu papel na diferenciação entre vozes femininas e masculinas, o que pode ser observado mesmo na presença da doença de Parkinson (Tabela 1). Isto significa que, quando a frequência fundamental é considerada na análise, há que se ter parâmetros acústicos diferenciados para homens e mulheres. Esse cuidado foi observado neste estudo e a discussão será retomada, posteriormente, na reflexão sobre as medidas dos formantes e do prisma acústico-articulatório.

Tabela 1 – Valores médios e desvio padrão da medida de frequência fundamental das vogais produzidas por homens e mulheres com DP (GParkinson)

		Frequência fundamental (Hertz)		
		homens	mulheres	p-valor
GParkinson	/a/	124 (14,77)	186 (31,95)	<b>0,002*</b>
	/i/	135 (28,14)	205 (37,58)	<b>0,004*</b>
	/u/	139 (29,55)	199 (35,88)	<b>0,008*</b>
	<b>média (DP)</b>	133 (23,07)	197 (34,47)	<b>0,003*</b>

Teste t-Student para amostras independentes;

\* Valores significativos:  $p \leq 0,05$ .

Legenda: DP = desvio padrão.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Na comparação entre o GParkinson e o GControle, não foram observadas diferenças significativas (Tabela 2). Para os homens com doença de Parkinson, observa-se uma tendência para valores mais altos de F0, resultados semelhantes aos obtidos por Hsu et al., 2017. Outra característica observada, nesse mesmo grupo, são os valores menores para o desvio padrão, indicando menor variabilidade na frequência fundamental.

Tabela 2 – Valores médios e desvio padrão da medida de frequência fundamental das vogais produzidas por homens e mulheres com DP (GParkinson) e homens e mulheres neurologicamente saudáveis (GControle)

		Frequência fundamental (Hertz)			
		GParkinson	GControle	p-valor	effect (r)
<b>homens</b>	/a/	124 (14,77)	116 (35,80)	0,62	0,18
	/i/	135 (28,14)	124 (38,99)	0,64	0,18
	/u/	139 (29,55)	123 (38,37)	0,50	0,25
	<b>média (DP)</b>	133 (23,07)	121 (37,66)	0,57	0,21
<b>mulheres</b>	/a/	186 (31,95)	185 (41,02)	0,94	0,02
	/i/	205 (37,58)	207 (36,25)	0,94	0,02
	/u/	199 (35,88)	197 (40,53)	0,93	0,03
	<b>média (DP)</b>	197 (34,47)	196 (39,20)	0,97	0,01

Teste t-Student para amostras independentes;

\* Valores significativos:  $p \leq 0,05$ .

Legenda: DP= desvio padrão

Fonte: Dados da pesquisa, Maceió, 2017/2018.

### 3.2. Intensidade

A intensidade é uma das grandezas físicas que caracterizam a onda sonora. Define-se pela energia transportada, perpendicularmente por ondas em uma determinada área e unidade de tempo. A intensidade de uma onda é proporcional ao quadrado da amplitude da pressão sonora; logo, intensidade e amplitude estão correlacionadas, mas não são sinônimos. Da percepção da intensidade derivam as qualidades “forte” ou “fraco” de um som, resultado da maior ou menor pressão sonora, que pode ser mensurada pela unidade logarítmica decibel (BARBOSA; MADUREIRA, 2015; PILLING, 2011).

A intensidade da voz/fala é um componente importante para a inteligibilidade e as variações prosódicas. Do ponto de vista fisiológico, a intensidade é dependente da pressão subglótica e requer controles tanto neuromusculares quanto respiratórios. Os ajustes contínuos da tensão da musculatura intrínseca da laringe e o fluxo de ar expiratório são responsáveis pela manutenção de uma adequada intensidade durante a produção da fala.

É extensa a literatura que demonstra a redução e menor variabilidade da intensidade da voz/fala na DP, estimando-se uma diminuição entre 13 a 20% dB. Mesmo quando orientados a falar com maior intensidade, falantes com DP não conseguem um incremento – em dB – similar ao de falantes saudáveis

BEHLAU et al., 2008; FOX; RAMIG, 1997; HO; IANSEK; BRADSHAW, 2001; KNOWLES et al., 2018; LAM; TJADEN, 2016; RUSZ et al., 2015; TJADEN; LAM; WILDING, 2013; VAICIUKYNAS et al., 2017).

O decréscimo da intensidade, observado na DP, poderia ser explicado pela diminuição na amplitude dos movimentos respiratórios e reduzida adução das pregas vocais, que resultam em menor pressão subglótica. O controle da força necessária para a complexa sequência de movimentos da fala parece ser afetado pelo circuito frontoestriado, na DP. Uma das consequências da queda da pressão subglotal é a gradual diminuição da intensidade (CRUTTENDEN, 1995; GRICE; BAUMAN, 2007; HO; IANSEK; BRADSHAW, 2001; RUSZ et al., 2017).

Para os participantes deste estudo, a média das medidas de intensidade foi de 70,68 dB ( $\pm 6,21$ ) para os falantes neurologicamente saudáveis e 64,74 dB ( $\pm 6,34$ ) para os falantes com DP, uma diferença de quase 6 dB, que representa uma redução de cerca de 8%. Apesar das diferenças metodológicas, variação idêntica foi notada em outros estudos (DYKSTRA; ADAMS; JOG, 2015; LIRANI-SILVA; MOURÃO; GOBBI, 2015; WALSH; SMITH, 2012).

Foram observadas características peculiares nas falas masculinas e femininas que merecem destaque, nesta discussão. Em uma análise intragrupo, a despeito da ausência de diferenças significativas, percebe-se que os homens do GParkinson apresentam valores de intensidade mais altos do que as mulheres, em todas as vogais examinadas. Comportamento inverso ocorre no GControle, no qual as mulheres revelam médias em torno de 5 dB acima dos homens (Tabela 3).

Na comparação entre os grupos (Tabela 4), notam-se diferenças significativas apenas para as mulheres, com uma redução em torno de 10 dB nos valores médios obtidos para as falantes com DP. A intensidade e a frequência fundamental são dependentes da pressão subglotal e o comportamento das falas masculinas, nos casos de DP, indica que ocorre um maior esforço visando à manutenção da intensidade e os prováveis ajustes realizados resultam em um aumento dos valores de F0, nesse grupo.

Tabela 3 – Valores médios e desvio padrão das medidas de intensidade (dB) obtidas no ponto central das vogais produzidas por homens e mulheres do GParkinson e GControle

	Vogais	Homens	Mulheres	p-valor
<b>GParkinson</b>	/a/	65,04 (4,76)	62,57 (5,20)	0,40
	/i/	68,29 (6,67)	63,04 (7,67)	0,23
	/u/	65,72 (6,73)	63,80 (8,41)	0,67
	<b>média (DP)</b>	66,35 (5,88)	63,14 (6,89)	0,40
<b>GControle</b>	/a/	67,58 (7,64)	72,36 (2,51)	0,36
	/i/	68,71 (9,41)	74,30 (2,67)	0,38
	/u/	68,01 (8,98)	73,12 (1,49)	0,38
	<b>média (DP)</b>	68,10 (8,63)	73,26 (1,40)	0,36

Teste t-Student para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: dB = decibéis; DP = desvio padrão.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Tabela 4: – Valores médios e desvio padrão das medidas de intensidade (dB) obtidas no ponto central das vogais produzidas por homens e mulheres com DP ou neurologicamente saudáveis

	vogais	GParkinson	GControle	p-valor	effect (r)
<b>homens</b>	/a/	65,04 (4,76)	67,58 (7,64)	0,55	0,22
	/i/	68,29 (6,67)	68,71 (9,41)	0,93	0,03
	/u/	65,72 (6,73)	68,01 (8,98)	0,68	0,16
	<b>média DP)</b>	66,35 (5,88)	68,10 (8,63)	0,72	0,13
<b>mulheres</b>	/a/	62,57 (5,20)	72,36 (2,51)	<b>0,01*</b>	0,75
	/i/	63,04 (7,67)	74,30 (2,67)	<b>0,04*</b>	0,67
	/u/	63,80 (8,41)	73,12 (1,49)	0,10	0,57
	<b>média (DP)</b>	63,14 (6,89)	73,26 (1,40)	<b>0,04*</b>	0,67

Teste t-Student para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: dB = decibéis; DP = desvio padrão.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

### 3.3 Formantes vocálicos, triângulo e área do espaço vocálico

A posição dos articuladores supraglotais – especialmente a posição da língua, o arredondamento ou não dos lábios e o grau de abertura da mandíbula – determina a identidade dos sons da fala. A posição da língua, durante a produção das vogais, permite a configuração de tubos acústicos que geram ressonâncias

– os formantes – e aumentam a amplitude do som em torno desses formantes (BARBOSA; MADUREIRA, 2015).

Os valores dos formantes estão estreitamente relacionados com os movimentos do corpo da língua nos eixos vertical e anteroposterior. O primeiro formante (F1) varia de acordo com a altura da língua em uma relação inversa, ou seja, quanto mais alta estiver a língua, mais baixo será o valor da frequência de F1; o segundo formante (F2), por sua vez, responde à posição da língua no eixo sagital de forma que quanto mais anterior for a sua posição, maior será o valor de F2. O clássico triângulo vocálico é uma representação gráfica dessa relação (BARBOSA; MADUREIRA, 2015; KENT; READ, 2015).

A configuração do tubo acústico é fundamental na determinação dos formantes e, além da forma, seu comprimento também deverá ser considerado, pois quanto menor o comprimento do trato, maiores serão os valores das frequências dos formantes. Nesse sentido, os valores absolutos das frequências dos formantes são influenciados por diferenças no tamanho do trato vocal, decorrentes do sexo e da idade do falante (BARBOSA; MADUREIRA, 2015; KENT; READ, 2015).

As medidas dos formantes F1 e F2 das vogais /i/, /a/ e /u/ possibilitam a construção do triângulo vocálico, em um plano bidimensional  $xy$ , no qual as coordenadas ( $F2_{\text{vogal}}$ ,  $F1_{\text{vogal}}$ ) determinam os vértices do triângulo. Considerando que os formantes são sensíveis à variável sexo, o mesmo pode ocorrer com o triângulo vocálico sendo, por isso, necessário separar falantes masculinos e femininos.

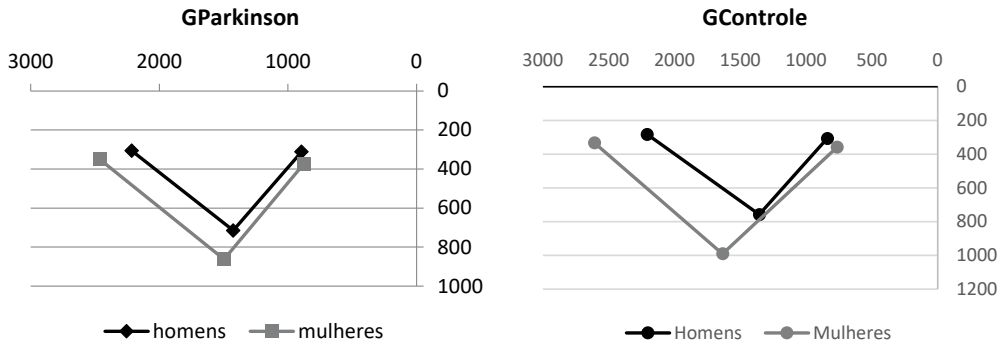
Com os dados obtidos neste estudo, foram elaborados os triângulos apresentados nas Figuras 1 e 2. Na Figura 1, é possível visualizar a diferença entre os triângulos obtidos para homens e mulheres, independente do fato de tratar-se de um participante do GParkinson ou do GControle. Assim, os triângulos vocálicos de falantes masculinos são, visualmente, menores do que aqueles obtidos para falantes femininos, corroborando resultados de outros estudos, como o de Lansford e Liss (2014b), destacando-se que, para falantes saudáveis, essa diferença é bastante acentuada.

A mera inspeção visual do triângulo pode não ser suficiente para identificar um falante com DP, sendo necessário representar a redução da amplitude dos movimentos articulatorios e consequente centralização do triângulo por meio do cálculo da área do espaço vocálico.

A área do espaço vocálico (VSA – *vowel space área*) pode ser considerada um índice da precisão articulatória, da habilidade de coordenação entre os movimentos da língua e a abertura/fechamento da mandíbula. É calculada por meio da expressão:

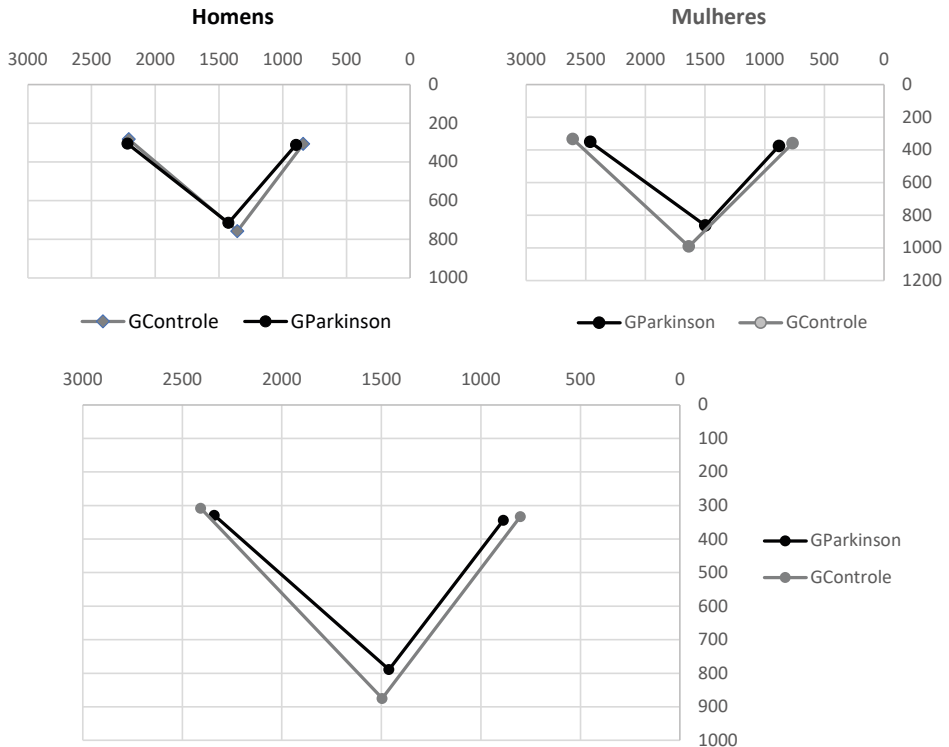
$$VSA = \text{ABS}((F1i * (F2a - F2u) + F1a * (F2u - F2i) + F1u * (F2i - F2a)) / 2) \quad (\text{LIU; TSAO; KUHL, 2005; SAPIR et al., 2010}).$$

Figura 1 – Triângulos vocálicos elaborados a partir dos valores médios de F1 e F2 das vogais /i/, /a/ e /u/ produzidas por homens e mulheres dos grupos estudados.



Fonte: Dados da pesquisa, Maceió, 2017/2018.

]Figura 2 – Triângulos vocálicos construídos a partir dos valores médios de F1 e F2 das vogais /i/, /a/ e /u/ produzidas por falantes com doença de Parkinson e falantes neurologicamente saudáveis.



Fonte: Dados da pesquisa, Maceió, 2017/2018.



Com o cálculo do VSA (Tabela 5), confirmam-se as observações sobre a configuração gráfica do triângulo: a sutil diferença entre os homens dos dois grupos não é significativa ( $p=0,46$ ), enquanto para as mulheres obtém-se um valor  $-p=0,03$ , mesmo para uma amostra pequena e participantes com DP sem disartria visível. Esse achado é coincidente com o estudo de Skodda, Visser e Schlegel (2011), no qual a área do espaço vocálico foi significativamente menor apenas para as mulheres e sem relação com os sintomas motores da DP.

Tabela 5 – Médias e desvio padrão da área do espaço vocálico (VSA)

	Área do espaço vocálico (Hz <sup>2</sup> )		p-valor	Cohen -d	(r)
	média (desvio padrão)				
	GParkinson	GControle			
<b>homens</b>	267341 (85254,35)	324107 (139826)	0,46	0,54	0,28
<b>mulheres</b>	397464,3 (113158,7)	594193,7 (72843,98)	<b>0,03*</b>	1,90	0,71

Teste t-Student para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: VSA = vowel space área; Hz = Hertz.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Até o momento, não há consenso quanto à eficácia da área do triângulo vocálico para a adequada caracterização da fala disártrica. As razões para as inconsistências encontradas com o uso do cálculo da área do espaço vocálico ainda permanecem no campo das hipóteses. Sapir et al. (2010) sugerem que a medida poderia ser útil apenas para estudos pré e pós-tratamento, em comparações intrafalantes e não interfalantes.

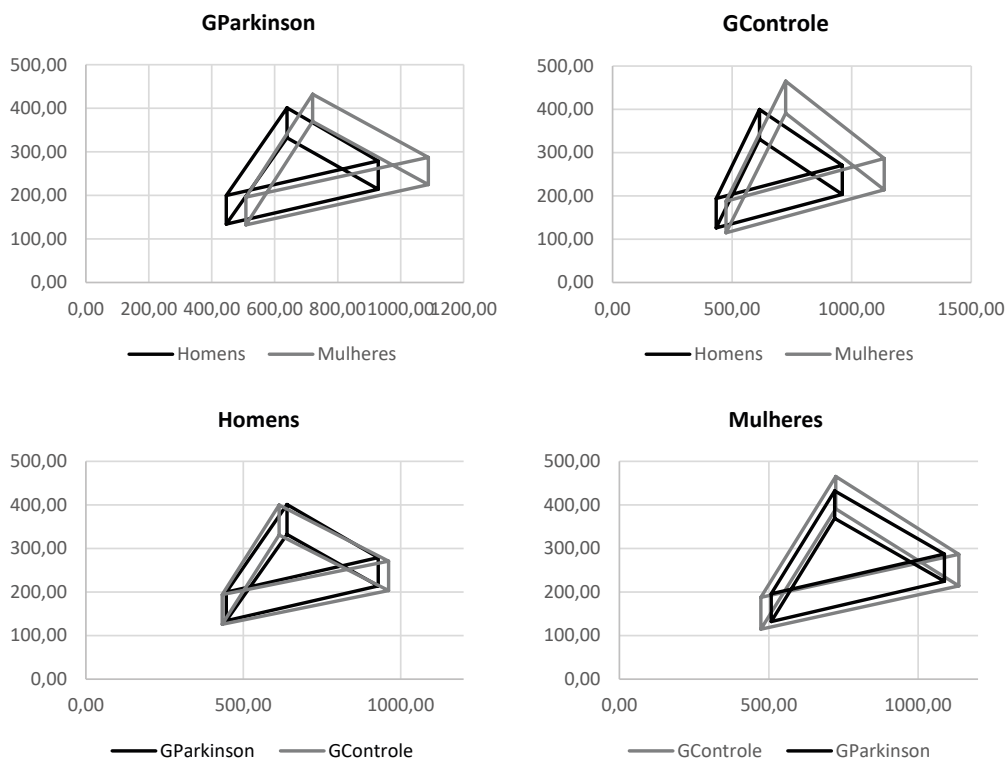
### 3.4 Prisma acústico-articulatório

Uma das metas deste estudo era a proposição de novas medidas que incluíssem os valores de intensidade, pois estes representam uma característica bastante relevante na fala disártrica. Para tal, são apresentados o prisma acústico-articulatório e o índice de articulação e intensidade vocálica.

O ponto de partida do prisma é o clássico triângulo vocálico, representado no plano bidimensional  $xy$ , por meio das coordenadas das vogais /i/, /a/ e /u/. As coordenadas são obtidas por meio dos valores de F2 (que correspondem aos valores de  $x$ ) e F1 (correspondentes aos valores de  $y$ ). Com a adição do eixo  $z$ , cada par ordenado foi convertido em um terço ordenado com a inclusão do valor de intensidade como a terceira variável.

Assim, com os valores absolutos dos formantes F1 e F2 e intensidade das vogais /a/, /i/ e /u/, foram obtidos os prismas acústico-articulatórios. A fim de investigar se o prisma acústico-articulatório também é sensível ao sexo do falante, foram comparados falantes femininos e masculinos, dentro de cada grupo (Figura 3).

Figura 3 – Projeção do prisma acústico-articulatório das vogais produzidas por homens e mulheres com doença de Parkinson (GParkinson) ou neurologicamente saudáveis (GControle).



Fonte: Dados da pesquisa, Macció, 2017/2018.

Entre homens e mulheres, há diferença significativa para o volume do prisma no GControle (Tabelas 6 e 7). Na comparação entre os grupos, percebe-se que a diferença ocorre para as mulheres, com resultados muito próximos daqueles obtidos com a área do espaço vocálico. Um exame dos resultados permite verificar que a inclusão da medida de intensidade das vogais (representada no eixo z) no prisma, proporcionou um aumento da sensibilidade desse parâmetro, em relação à área do triângulo vocálico.

Tabela 6 – Médias e desvio padrão do volume do prisma acústico-articulatório (VPAA)

VPAA					
média (desvio padrão)					
	Homens	Mulheres	p-valor	Cohen -d	(r)
<b>GParkinson</b>	17210222 (4251718)	26214452 (10047637)	0,07	1,16	0,53
<b>GControle</b>	22101390(9798563)	43653479 (5434966)	<b>0,03*</b>	2,72	0,85

Teste t-Student para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: VPAA = volume do prisma acústico-articulatório.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Tabela 7 – Médias e desvio padrão do volume do prisma acústico-articulatório (VPAA)

VPAA					
média (desvio padrão)					
	GParkinson	GControle	p-valor	Cohen -d	(r)
<b>homens</b>	17210222 (4251718)	22101390(9798563)	0,31	0,77	0,38
<b>mulheres</b>	26214452 (10047637)	43653479 (5434966)	<b>0,02*</b>	1,94	0,72

Teste t-Student para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: VPAA = volume do prisma acústico-articulatório.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

### 3.5 Índices

O índice de centralização formântica (FCR – *formant centralization ratio*) foi proposto por Sapir, Ramig, Spielman e Fox, em 2010, e é calculado por meio da expressão  $(F2u+F2a+F1i+F1u)/(F2i+F1a)$ . Seu valor deve situar-se entre zero e 2,0, sendo que os valores maiores a 1,0 são indicadores de centralização. Quanto maior o valor, maior a constrição do triângulo vocálico.

O fato de os valores dos formantes poderem ser expressos por meio de uma razão possibilita que as variações interfalantes sejam superadas e o índice seja minimamente sensível a estas. Na pesquisa de Sapir e colaboradores (2010), o FCR mostrou-se eficaz, e mais sensível que a área do espaço vocálico, na identificação de uma fala disártrica, independente da idade e do sexo do falante.

Esse desempenho, entretanto, não pode ser confirmado por outros estudos, nos quais o FCR apresentou boa especificidade (sucesso ao identificar falantes saudáveis), mas baixa sensibilidade (falha ao identificar falantes disártricos) (LANSFORD e LISS, 2014a e 2014b).

O índice de articulação vocálica (VAI – *vowel articulation index*) também foi desenvolvido com o propósito de superar a sensibilidade interfalantes observada nas medidas da área do espaço vocálico. Quanto maior o valor do índice, menor a centralização do triângulo e mais precisa é a articulação (KNOWLES et al., 2018; ROY et al., 2009; RUSZ et al., 2017; RUSZ et al., 2013).

Pode ser calculado por meio da expressão  $(F2i+F1a)/(F2u+F2a+F1i+F1u)$ , que é o inverso do índice de centralização formântica. No índice de articulação vocálica, espera-se que os valores dos formantes, no numerador, diminuam enquanto os valores, no denominador, aumentem, quanto maior a constrição do espaço vocálico (KNOWLES et al., 2018; SKODDA; VISSER; SCHLEGEL, 2011).

Os resultados obtidos são apresentados na Tabela 8. Como os dois índices são considerados insensíveis à variável sexo (*gender-independent*), foram calculadas a média e desvio padrão para falantes masculinos e femininos agrupados.

Tabela 8 – Médias e desvio padrão dos índices obtidos para falantes com doença de Parkinson (GParkinson) e falantes saudáveis (GControle)

	<b>GParkinson</b>	<b>GControle</b>	<b>p-valor</b>	<b>Cohen -d</b>	<b>(r)</b>
<b>ICF</b>	0,97 (0,09)	0,90 (0,09)	0,15	0,73	0,34
<b>IAV</b>	1,04 (0,10)	1,12 (0,11)	0,14	0,76	0,35

Teste t-Student; para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: ICF = índice de centralização formântica; IAV = índice de articulação vocálica.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Para o índice de centralização formântica (ICF), quanto maior o valor, maior a constrição do espaço vocálico. Observam-se valores maiores para os falantes do GParkinson sem, entretanto, apresentar uma diferença significativa em relação ao GControle. No índice da articulação vocálica, inverso do ICF, constata-se quase o mesmo resultado.

O índice de Articulação e Intensidade Vocálica (Tabela 9), por sua vez, mostrou-se bastante robusto à diferenciação entre falantes com DP e falantes neurologicamente saudáveis, quando comparado ao Índice de Centralização Formântica e ao Índice de Articulação Vocálica.

Tabela 9 – Médias e desvio padrão dos índices obtidos para falantes com doença de Parkinson (GParkinson) e falantes saudáveis (GControle)

	<b>GParkinson</b>	<b>GControle</b>	<b>p-valor</b>	<b>Cohen -d</b>	<b>(r)</b>
<b>ICF</b>	0,97 (0,09)	0,90 (0,09)	0,15	0,73	0,34
<b>IAV</b>	1,04 (0,10)	1,12 (0,11)	0,14	0,76	0,35
<b>IAIV</b>	67,31 (9,04)	79,17 (11,54)	<b>0,02*</b>	1,19	0,51

Teste t-Student; para amostras independentes;

\* Valores significativos ( $p \leq 0,05$ ).

Legenda: ICF = índice de centralização formântica; IAV = índice de articulação vocálica; IAIV = Índice de articulação e intensidade vocálica.

Fonte: Dados da pesquisa, Maceió, 2017/2018.

Os novos parâmetros, aqui propostos, não requerem equipamentos e softwares sofisticados, nem cálculos complicados e inacessíveis e demonstram que a incorporação das medidas de intensidade das vogais pode contribuir para o aumento da sensibilidade na diferenciação de uma fala disártrica.

#### 4. CONSIDERAÇÕES FINAIS

Falantes com doença de Parkinson apresentam padrões acústicos diferenciados nas medidas de intensidade, frequência fundamental e formantes vocálicos. Devido à redução no controle e mobilidade das estruturas fonoarticulatórias, observa-se uma constrição no triângulo vocálico e diminuição na área do espaço vocálico. Os índices de articulação vocálica também demonstram a tendência à centralização dos movimentos da língua, durante a fala.

A diminuição da intensidade é uma característica marcante da fala disártrica e, até o momento, havia sido negligenciada nas medidas acústicas. O Prisma Acústico-Articulatório e o Índice de Articulação e Intensidade Vocálica incorporam as medidas de intensidade e contribuem para o aumento da sensibilidade desses parâmetros, na identificação da fala na doença de Parkinson.

Os dois novos parâmetros acústico-articulatórios propostos por este estudo, com a inclusão das medidas de intensidade, são promissores e representam um avanço para a avaliação objetiva da fala, na doença de Parkinson, com grande potencial para aplicação em estudos futuros, contemplando outras patologias e demais áreas com interesse no tema.

## REFERÊNCIAS

- BANG, Young-im *et al.* Acoustic characteristics of vowel sounds in patients with Parkinson disease. **Neurorehabilitation**, v. 32, n. 3, p. 649-654, 2013.
- BARBOSA, Plínio A.; MADUREIRA, Sandra. **Manual de fonética acústica experimental: aplicações a dados do português**. São Paulo: Cortez. 2015. 591p.
- BEHLAU, Mara; MADÁZIO, Glaucya; FEIJÓ, Deborah; PONTES, Paulo. Avaliação de voz. *In*: BEHLAU, Mara. (org.). **Voz: o livro do especialista**. Rio de Janeiro: Revinter. v. 1. Cap. 3. 2008. p. 83-245.
- BOERSMA, Paul; WEENINK, David. **Praat: doing phonetics by computer**. Versão 6.0.25. Amsterdam: University of Amsterdam, 2017. Software. Disponível em: <http://www.praat.org>.
- BRASOLOTTO, Alcione Ghedini; WOLF, Aline Epifânio; DIAFÉRIA, Giovana; AZEVEDO, Luciana Lemos de. Voz na senescência e na doença de Parkinson. *In*: MARCHESAN, Irene Queiroz; SILVA, Hilton Justino da; TOMÉ, Marileda Cattelan. (orgs). **Tratado das especialidades em Fonoaudiologia**. São Paulo: Guanabara-Koogan. 2014. pp. 171-177.
- CRUTTENDEN, Alan. **Intonation**. 2. ed. London: Cambridge University Press. 1997.
- DYKSTRA, Allyson D.; ADAMS, Scott G.; JOG, Mandar. Examining the relationship between speech intensity and self-rated communicative effectiveness in individuals with Parkinson's disease and hypophonia. **Journal of Communication Disorders**, v. 56, p. 103-112, 2015.
- FORREST, Karen; WEISMER, Gary. Acoustic analysis of motor speech disorders. *In*: McNEIL, Malcolm R. (ed.). **Clinical management of sensorimotor speech disorders**. 2. ed. New York: Thieme. 2009. pp. 46-63.
- FOX, Cynthia M.; RAMIG, Lorraine Olson. Vocal Sound Pressure Level and Self-Perception of Speech and Voice in Men and Women with Idiopathic Parkinson Disease. **American Journal of Speech-language Pathology**, v. 6, n. 2, p. 85-94, maio 1997.
- FOX, John; BOUCHET-VALAT, Milan. **Rcmdr: R Commander**. Versão 2.2.4. Boca Raton, FL. 2018. R package. Disponível em: <http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/>.

GOETZ, Christopher G. *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. **Movement Disorders**, v. 23, n. 15, p. 2.129-2.170, 2008.

GRICE, Martine; BAUMAN, Stefan. An introduction to intonation – functions and models. *In*: TROUVAIN, Jürgen; GUT, Ulrike. (eds.). **Non-native prosodies: phonetic description and teaching practice**. Berlin: Mouton de Gruyter. 2007. pp. 25-51.

HO, Aileen K.; IANSEK, Robert; BRADSHAW, John L. Motor instability in parkinsonian speech intensity. **Cognitive and Behavioral Neurology**, v. 14, n. 2, p. 109-116, 2001.

HSU, Sih-Chiao *et al.* Acoustic and perceptual speech characteristics of native Mandarin speakers with Parkinson's disease. **The Journal of the Acoustical Society of America**, v. 141, n. 3, p. EL293-EL299, 2017.

KANDA, Yoshinobu. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. **Bone marrow transplantation**, v. 48, n. 3, p. 452, 2013.

KENT, Ray D.; READ, Charles. **Análise acústica da fala**. Trad.: Alexsandro Meireles. São Paulo: Cortez. 2015. 504p.

KNOWLES, Thea *et al.* Deep brain stimulation of the subthalamic nucleus parameter optimization for vowel acoustics and speech intelligibility in Parkinson's disease. **Journal of Speech, Language, and Hearing Research**, v. 61, n. 3, p. 510-524, 2018.

LAM, Jennifer; TJADEN, Kris. Clear speech variants: An acoustic study in Parkinson's disease. **Journal of Speech, Language, and Hearing Research**, v. 59, n. 4, p. 631-646, 2016.

LANSFORD, Kaitlin L.; LISS, Julie M. Vowel acoustics in dysarthria: speech disorder diagnosis and classification. **Journal of Speech, Language, and Hearing Research**, v. 57, n. 1, p. 57-67, 2014a.

LANSFORD, Kaitlin L.; LISS, Julie M. Vowel acoustics in dysarthria: mapping to perception. **Journal of Speech, Language, and Hearing Research**, v. 57, n. 1, p. 68-80, 2014b.

LIRANI-SILVA, Camila; MOURAO, Lúcia Figueiredo; GOBBI, Lilian Teresa Bucken. Disartria e Qualidade de Vida em idosos neurologicamente sadios e pacientes com doença de Parkinson. **CoDAS**, São Paulo, v. 27, n. 3, p. 248-254, 2015.

LIU, Huei-Mei; TSAO, Feng-Ming; KUHL, Patricia K. The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. **The Journal of the Acoustical Society of America**, v. 117, n. 6, p. 3.879-3.889, 2005.

MARTÍNEZ-SÁNCHEZ, Francisco. Trastornos del habla y la voz en la enfermedad de Parkinson. **Revista de Neurología**, v. 51, n. 9, p. 542-550, 2010.

MIRANDA, Irma Iunes; MEIRELES, Alexsandro. Descrição acústica das vogais tônicas da fala capixaba. **Letras de Hoje**, Porto Alegre, v. 47, n. 3, p. 325-332, 2012.

OKADA, Yukihiro; MURATA, M. I. H. O.; TODA, Tatsushi. Effects of levodopa on vowel articulation in patients with Parkinson's disease. **Kobe Journal of Medical Sciences**, v. 61, n. 5, p. E144-54, 2016.

OROZCO-ARROYAVE, Juan Rafael *et al.* Phonation and articulation analysis of spanish vowels for automatic detection of parkinson's disease. *In: International Conference on Speech, Text and Dialogue*, 17<sup>th</sup>, 2014, Czech Republic. SOJKA, P. *et al.* (eds.). **Proceedings** [...], pp. 374-381. 2014.

PILLING, Sergio. **Ondas, som e introdução a bioacústica**. Curso de Biofísica. São José dos Campos: Universidade do Vale do Paraíba. 2011. 23p.

POSTUMA, Ronald B. *et al.* How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. **Brain**, v. 135, n. 6, p. 1.860-1.870, 2012.

R CORE TEAM. **R: a language and environment for statistical computing**. Versão 3.4.2. Software. Disponível em: <http://www.R-project.org>. 2018.

ROY, Nelson *et al.* Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. **Journal of Communication Disorders**, v. 42, n. 2, p. 124-135, 2009.

RUSZ, Jan *et al.* Comparative analysis of speech impairment and upper limb motor dysfunction in Parkinson's disease. **Journal of Neural Transmission**, v. 124, n. 4, p. 463-470, 2016.



RUSZ, Jan *et al.* Speech disorders reflect differing pathophysiology in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. **Journal of Neurology**, v. 262, n. 4, p. 992-1.001, 2015.

RUSZ, Jan *et al.* Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. **The Journal of The Acoustical Society of America**, v. 134, n. 3, p. 2.171-2.181, 2013.

SAPIR, Shimon; SPRECHER, Elliot; SKODDA, Sabine. Early motor signs of Parkinson's disease detected by acoustic speech analysis and classification methods. *In: MAVEBA International Workshop Models and Analysis of Vocal Emissions for Biomedical Applications*, 8<sup>th</sup>, Firenze, 2013. MANFREDI, Claudia. (ed.). **Proceedings and Report** [...], p. 3. 2013.

SAPIR, Shimon *et al.* Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech. **Journal of Speech, Language, And Hearing Research**, v. 53, n. 1, p. 114-125, 2010.

SAPIR, Shimon; RAMIG, Lorraine; FOX, Cynthia. Speech and swallowing disorders in Parkinson disease. **Current Opinion in Otolaryngology & Head And Neck Surgery**, v. 16, n. 3, p. 205-210, 2008.

SAPIR, Shimon *et al.* Effects of Intensive Voice Treatment (the Lee Silverman Voice Treatment [LSVT]) on Vowel Articulation in Dysarthric Individuals with Idiopathic Parkinson Disease: Acoustic and Perceptual Findings. **Journal of Speech, Language, And Hearing Research**, v. 50, n. 4, p. 899-912, 2007.

SKODDA, Sabine *et al.* Progression of Voice and Speech Impairment in the Course of Parkinson's Disease: A Longitudinal Study. **Parkinson's Disease**, v. 2013, p. 1-8, 2013.

SKODDA, Sabine; GRÖNHEIT, Wenke; SCHLEGEL, Uwe. Abnormalities of voice quality in progressive supranuclear palsy (PSP) [abstract]. **Movement Disorders**. v. 27, n. S1, p. 1197, 2012.

SKODDA, Sabine; VISSER, Wenke; SCHLEGEL, Uwe. Vowel Articulation in Parkinson's Disease. **Journal of Voice**, v. 25, n. 4, p. 467-472, 2011.

TJADEN, Kris *et al.* Acoustic and Perceptual Consequences of Clear and Loud Speech. **Folia Phoniatica Et Logopaedica**, v. 65, n. 4, p. 214-220, 2013.

TJADEN, Kris; LAM, Jennifer; WILDING, Greg. Vowel Acoustics in Parkinson's Disease and Multiple Sclerosis: Comparison of Clear, Loud, and Slow Speaking Conditions. **Journal of Speech, Language, And Hearing Research**, v. 56, n. 5, p. 1.485-1.502, 2013.

VAICIUKYNAS, Evaldas *et al.* Detecting Parkinson's disease from sustained phonation and speech signals. **PlosOne**, v. 12, n. 10, p. e185613, 2017.

WALSH, Bridget; SMITH, Anne. Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. **Movement Disorders**, v. 27, n. 7, p. 843-850, 2012.



## A MENTIRA TEM PERNA CURTA? PISTAS PROSÓDICAS NA IDENTIFICAÇÃO DE DISCURSO ENGANOSO

*Remildo Barbosa da Silva<sup>1</sup>*

### 1. INTRODUÇÃO

As atitudes de fala são representadas por comportamentos linguísticos determinados. Vários estudos têm indicado que a prosódia tem um papel fundamental nesse tocante (Oliveira, 2011; Lu, Aubergé e Rilliard, 2012; Paula, 2012; Mitchell e Ross, 2013; Moura, 2016; Barbulescu, Ronfard e Bailly, 2017; Pacheco e Alves, 2017). O trabalho de Alves (2002) sobre persuasão em português ilustra bem esse ponto. Nesse trabalho, a autora observa que determinados aspectos melódicos estão associados a enunciados persuasivos.

Há já um bom número de estudos que apontam essa relação para o português. Em um trabalho que contrastou as atitudes de dúvida e certeza, Silva (2008) verificou que parâmetros como frequência fundamental, duração de enunciado e intensidade apresentaram valores diferentes na elocução dessas atitudes. Oliveira (2011) discorre sobre características prosódicas na atitude da dúvida, certeza e incredulidade. Dentre vários elementos prosódicos analisados, a autora afirma que as medidas de frequência fundamental se apresentaram mais relevantes.

---

<sup>1</sup> Universidade Federal de Alagoas

Por sua vez, Paula (2012), estudou padrões prosódicos na expressão de ironia. Ao analisar frequência fundamental, intensidade e duração, os resultados indicaram valores mais elevados desses constituintes na atitude de ironia, se comparados à leitura. Em estudo semelhante, Almeida (2016) apresenta uma análise de uma série de parâmetros prosódicos na expressão de ironia comparados à expressão de enunciados não irônicos. Seus resultados indicaram uma maior diferença no que se refere aos valores de frequência fundamental e duração de sílabas proeminentes nas duas condições de fala.

O trabalho de Almeida et al. (2015), por sua vez, apresenta uma análise de fenômenos temporais na indicação de atitude. A partir de uma análise intrasujeito, seus resultados apontaram para uma lentificação da velocidade de fala em enunciados associados à emoção de “raiva” quando comparados com expressões “neutras”. Segundo os autores, os resultados se opõem aos que a literatura descreve para outras línguas.

Uma questão importante que deve ser considerada em estudos da relação entre prosódia e sentido é a distinção entre atitude e emoção. Muitos autores divergem sobre a distinção ou não desses dois termos. Mozziconacci e Hermes (1997), por exemplo, tratam sobre essas expressões sem distingui-las. Crystal (1995) ao discorrer sobre as funções da entonação, afirma que a função das entonações emocionais é exteriorizar o significado das atitudes. Por outro lado, há autores que buscam delimitar bem o que é atitude e o que é emoção. Fónagy (1993), por exemplo, evidencia a necessidade de distinguir esses dois termos e afirma que uma atitude representa comportamentos determinados, conscientes, controlados, tendo um componente moral, intelectual, opondo-a às emoções, enquanto descargas espontâneas de uma tensão psíquica.

Seguindo o mesmo posicionamento de Fónagy, Wichmann (2000) afirma que as emoções se diferenciam das atitudes por estarem representadas no sinal de fala de forma mais evidente. Em seu trabalho, Aubergé (2002) ilustra como as atitudes apresentam aspectos diferentes das emoções. Segundo a autora, as atitudes encontram-se em um nível mais cortical e, portanto, mais linguístico e cognitivo, o que possibilita que o falante tenha mais controle sobre elas, fatos que se opõem às emoções, que se revelam menos corticais e distantes das funções linguísticas e, então, menos controladas.

Aspectos cognitivos são levados em conta por quem se dedica a estudar a mentira. Zuckerman et al. (1981) ressaltam que aspectos cognitivamente mais complexos são necessários para a elaboração de mentiras eficazes. No entanto, eles pontuam que emoções podem se manifestar durante a elocução de uma

mentira, o que indica que quem mente teme que sua mentira seja descoberta. Embora este estudo não tenha identificado autores que classificam a mentira como atitude, e diante das colocações feitas até aqui, assumimos que a mentira apresenta traços atitudinais.

Os estudos sobre prosódia e atitude apresentados até aqui revelam uma grande contribuição para a linguística e áreas afins. No entanto, destacamos a necessidade de investigar o comportamento de elementos prosódicos em atitudes pouco exploradas, como é o caso da mentira. A atitude de mentir ainda não vem recebendo a atenção que outras já estudadas despertaram nos pesquisadores, principalmente no Brasil. E embora ainda escassos, alguns estudos sugerem que elementos prosódicos podem funcionar como pistas para identificar se um indivíduo está mentindo ou não (Ekman, 1985; DePaulo et al., 2003; Spence et al., 2012; Benus et al., 2006). Além disso, boa parte dos estudos propõem análises prosódicas majoritariamente relacionadas à entoação, e quando priorizam os parâmetros relacionados ao tempo de enunciação os resultados revelam-se insatisfatórios. Diante disso, o presente trabalho teve por objetivo contribuir com estudos sobre o papel da prosódia na caracterização de atitudes. As análises se concentraram em fenômenos pausológicos presentes na enunciação de declarações mentirosas, em oposição a declarações verdadeiras. Especificamente, este estudo analisou a latência durante a transição de turno de fala e a ocorrência e duração de pausas em enunciados mentirosos e enunciados verdadeiros.

## **2. ESTADO DA ARTE**

A ciência ainda procura meios para identificar características específicas do discurso mentiroso, que já foi alvo de inúmeros experimentos no passado – alguns, inclusive, rodeados de controvérsias (Trovillio, 1939). A seguir, serão apresentados, de maneira bastante sucinta, alguns conceitos-chave para a discussão do presente estudo.

### **2.1. A mentira**

Definir o que é a mentira e delimitar essa atitude em um estudo é uma tarefa muito desafiadora. Isso se justifica pelo fato de que cada mentira ocorre em um contexto diferente com motivos e propósitos bem distintos. Ekman (1985) define mentira como uma situação na qual um indivíduo pretende enganar um outro de forma deliberada, sem que este seja notificado previamente desse propósito e sem

ter sido explicitamente solicitado a fazê-lo pelo alvo. Assim, o autor categoriza duas formas principais de mentir: a falsificação e a dissimulação. Embora não traga uma definição própria, Metts (1989) apresenta uma classificação mais ampla para os tipos de mentira: (i) Falsificação; (ii) Distorção; (iii) Omissão; e (iv) Fuga.

Masip, Garrido e Herrero (2004) entendem a mentira como uma tentativa deliberada de ocultar, fabricar e/ou manipular de qualquer outra forma informações factuais e/ou emocionais, por meios verbais e/ou não verbais, com intuito de criar ou manter em outro ou em outros uma crença que o próprio comunicador considera falsa. Os autores ainda reconhecem que essa atitude pode ser bem-sucedida ou não.

Para Ballone (2006), mentir é dizer ser verdade aquilo que é falso com o intuito de induzir o outro ao erro. O autor ainda alerta afirmando que a mentira não é contada apenas de forma deliberada. Ele faz distinção daquelas mentiras que as pessoas já esperam ser enganadas, como uma ficção ou uma fábula, que nem sempre se apresentam como sinônimo de mentira. Segundo ele, não pode ser mentira a literatura, e a arte, por exemplo.

Carson (2010) afirma que mentir é muito frequentemente confundido com enganar e, apesar de afirmar que muitas vezes esses dois termos são utilizados de formas equivalentes, ele afirma que a mentira se distingue do engano no sentido de que para mentir é necessário que haja uma declaração falsa e essa declaração seja oralizada. O engano, por sua vez, pode ser realizado sem nenhuma declaração por meio da fala oral. Outro ponto que o autor coloca é a probabilidade de sucesso ao mentir e ao enganar, uma vez que enganar implica sucesso, enquanto mentir nem sempre pode ser uma atitude bem-sucedida.

Na concepção de Rodrigues (2016), a mentira pode ocorrer em qualquer contexto social. O autor a define como uma comunicação pela qual um indivíduo, de forma intencional e conveniente, busca controlar o comportamento de outro. Na sua concepção, a mentira não ocorre apenas pela produção de uma fala que o emissor acredita ser falsa, mas também pela ocultação de sinais que revelem a veracidade da mensagem, ou pela estratégia intencional de produzir sinais cujo foco é desviar a atenção do receptor.

Com base nessas considerações, neste trabalho determinamos a mentira como a transmissão consciente e deliberada de uma mensagem completamente falsa por um falante numa situação na qual não foi dada oportunidade para ele pensar. Julgamos que a presença de emoções, como o medo, foi mínima ou inexistente, levando em consideração que o participante deste estudo não corria risco de prejuízo caso deixasse perceber que estava mentindo.

## 2.2. A Prosódia

Para Fujisaki (1997) a prosódia tem manifestações mensuráveis e princípios subjacentes e só é percebida à medida que uma mensagem é produzida com uma sequência coerente de sons da fala. Conforme o autor, a prosódia apresenta organização sistemática de diversas unidades linguísticas presentes em um enunciado ou um grupo coerente de enunciados.

Acusticamente, a fala pode ser analisada a partir de seus três componentes básicos: frequência, amplitude e tempo, conforme é detalhado por Couper-Kuhlen (1986). Esses constituintes permitem ao ouvinte captar pistas linguísticas essenciais para a identificação de atitudes e emoções específicas expressas por seu enunciador.

Ao discorrer sobre funções prosódicas, Barbosa (2012) pontua que é possível identificá-las tanto no plano linguístico quanto expressivo. O autor afirma que os estudos de funções atitudinais (como atitudes, postura interpessoal e estilo de locução) pertencem ao plano expressivo, assim como o estudo das funções afetivas (emoções como tristeza, raiva e alegria) e das funções indiciais (marcas de gênero e sexo, origem social e dialetal).

As investigações de pistas vocais do estudo expressivo da prosódia têm comprovado que diversos padrões comportamentais desses elementos caracterizam tipos distintos de discursos em diferentes contextos de fala. Na seção a seguir, estão detalhadas particularidades da latência de resposta e da pausa, necessárias para uma melhor compreensão deste estudo.

### 2.2.1. A Latência de Resposta

Reynolds e Rendle-Short (2011) afirmam que várias pistas para identificação de mentiras têm sido investigadas em estudos que tratam sobre a comunicação. Uma delas é a *latência de resposta*. Conforme os autores, esse termo é frequentemente usado para descrever o “atraso” de uma resposta. Eles afirmam que o contexto pode ter influência sobre a presença e a duração de latência entre uma pergunta e uma resposta que se crê mentirosa.

Em seus estudos, Benus et al. (2006) advogam que a latência para a enunciação de uma mensagem mentirosa foi mais longa quando comparada à enunciação de mensagens verdadeiras. A diferença encontrada por eles foi de 20 milissegundos antes do próximo enunciado.

Spence et al. (2012) afirmam que há estudos que não reportam diferenças entre a latência que precede a elocução de um fato e a que precede uma mentira.



Por outro lado, há pesquisas que reportam diminuição da latência para respostas mentirosas. Neste último caso, segundo os autores, é sugerido que essa diminuição ocorre devido às crenças dos falantes de que respostas mais rápidas estão associadas a uma impressão mais confiável sobre uma declaração. Contudo, a literatura mostra que isso vai muito além de crenças, pois também são encontrados estudos que defendem a ideia de que o maior período de latência está relacionado a uma possível mentira. Os achados dos autores corroboram com os de Benus et al. (2006). Os resultados mostraram que a latência da resposta era mais longa na condição de fala mentirosa (1200.77 ms) do que na condição de fala verdadeira (775.26 ms). Segundo eles, essa diferença foi significativa.

É notável a discordância entre os estudos. Dessa forma, é importante considerar o que é declarado por Reynolds e Rendle-Short (2011): alguns estudos apontam para a uma inconsistência no que se refere à relação da latência com a mentira devido ao contexto de fala no qual ela está inserida.

### 2.2.2. A Pausa

De acordo com O'Connell e Kowal (1983), a pausa ocorre por vários fatores, tais como ansiedade, ênfase, interrupções e respiração. Esse constituinte prosódico ainda é classificado em pausas preenchidas e pausas não preenchidas, levando-se em consideração a ausência ou presença de vocalização em sua produção.

Para Oliveira Jr. (2000), a pausa pode ser entendida como uma pista para a segmentação de enunciados. As pausas mais curtas são pistas de conexão semântica entre essas unidades de sentido, ou seja, indicam que não aconteceu ainda o término de uma determinada seção num discurso. Estudos de caráter psicolinguísticos, como os de Goldman-Eisler (1968), consideram a pausa como um fator importantíssimo para o processamento da fala. A autora observou que as pausas mais longas ocorrem com mais frequência quando há exigência cognitiva maior, como a imprevisibilidade de uso de certas palavras que podem acarretar a dificuldade do processamento da fala. Seguindo esse mesmo raciocínio, o estudo de Ford et al. (1982) sugere que pausas que apresentam duração mais longa podem fornecer o tempo que o falante utiliza para organizar a estrutura do próximo enunciado.

Oliveira Jr. (2000) adverte sobre problemas metodológicos relacionados à medição de pausas no discurso. O autor apresenta uma série de estudos e destaca o desacordo no que diz respeito ao ponto de corte que define uma pausa silenciosa. Duez (1985) realizou um teste de percepção com um grupo de voluntários e determinou cinco classes de duração para a pausa silenciosa: (i) menores que

250 ms; (ii) de 250 ms a 400 ms; (iii) de 400 ms a 600 ms; (iv) de 600 ms a 900 ms; e (v) pausas com duração superior a 900 ms. Em um estudo com narrativas espontâneas do português brasileiro, Oliveira Jr. (2000) verificou que pausas silenciosas apresentaram duração média de 240 ms. Em média, a presença delas ocupa cerca de 1/4 da totalidade das narrativas. Segundo o autor, esse achado confirma o que foi constatado em estudos anteriores de Parmenter e Treviño (1935) e Brotherton (1979), que observaram que pausas silenciosas ocupavam cerca de 25% da duração total da fala. Glukhov (1975) realizou um estudo que incluiu línguas românicas, como espanhol, português, francês, italiano. Seus resultados constataram uma grande diferença na frequência em que elas ocorrem num intervalo entre 50 a 150 ms.

Embora haja esse impasse sistemático, o ponto de corte mínimo mais comum é de 250 ms, conforme sustentado nos estudos de Goldman-Eisler (1968). De acordo com a autora, 71,5% de todas as pausas ocorrem no período entre 250 ms e 1 segundo. Sua justificativa é que delimitar um ponto de corte menor pode trazer implicações para um estudo, pois pausas menores que 250 ms são majoritariamente articulatórias. Por outro lado, Hieke, Kowal e O'Connell (1983) questionam os argumentos de Goldman-Eisler (1968) e estabelecem um ponto de corte mínimo de 130 ms. Os autores afirmam que, embora não seja comum encontrar pausa com duração entre 130 ms e 250 ms, fatos fonotáticos e análises empíricas de vários *corpora* contradizem a alegação de que há predominância de pausas articulatórias nesse intervalo.

Apesar de esporádicos, os estudos sobre a pausa no discurso mentiroso apresentam uma grande contribuição e destacam a importância de sua análise. Ekman (1985) chama a atenção para a dificuldade no controle da fala durante o ato de mentir. Ele afirma que as pausas são as pistas mais comuns na detecção de mentiras por meio da análise acústica da fala, tanto em relação à frequência que ocorrem, quanto à duração acima do normal. O autor ainda reforça dizendo que o mentiroso hesita antes de responder a uma pergunta. Ainda segundo o autor, no decorrer da fala, as pausas são mais breves que na hesitação. No entanto, ele esclarece que, em alguns casos, por tentarem ter um controle de suas palavras e acabarem revelando verdades, a fala de mentirosos pode acumular uma maior quantidade de pausas.

Vrij et al. (1999) observaram que para a formulação de uma mentira os indivíduos utilizam mais pausas preenchidas que indivíduos instruídos a dizer a verdade. E que pessoas que foram instruídas a contar mentiras apresentaram mais disfluências. Por outro lado, DePaulo et al. (2003) afirmam que em discurso

mentiroso há menos pausa preenchida que em discurso verdadeiro. No entanto, os resultados foram insatisfatórios para assumir que pausas são pistas para detecção de discurso mentiroso.

Embora os resultados de Benus et al. (2006) concordem com os de DePaulo et al. (2003), estudos que investigaram a presença da pausa em enunciados mentirosos divergem acerca de sua importância como pista para descoberta de mentiras, como é pontuado pelos próprios autores. Seus achados sugerem que, em geral, a utilização da pausa está mais relacionada com o discurso verdadeiro do que com o discurso mentiroso. Tanto as pausas preenchidas quanto as não preenchidas se manifestaram mais em discursos classificados como verdadeiros.

Embora apresentem divergências, os estudos sobre a latência e a pausa como pista para identificação de mentiras se tornam importantes devido à grande variedade de contextos de fala nos quais esses elementos ainda não foram investigados ou nos quais os estudos ainda são insuficientes. É possível que haja disparidades em relação a padrões de ocorrência e duração desses constituintes entre as línguas, mesmo que elas pertençam à mesma família linguística. Esse dado é importante, pois enfatiza ainda mais a necessidade de explorar o comportamento desses parâmetros. Na seção seguinte apresentaremos a metodologia deste estudo.

### 3. METODOLOGIA

Esta seção detalha todo o procedimento metodológico realizado no desenvolvimento deste estudo. Nas seções subsequentes estão descritas informações acerca dos participantes que cederam dados de fala para esta pesquisa e a caracterização do *corpus* utilizado.

#### 3.1. Perfil dos participantes

Os participantes que produziram os enunciados para este estudo tinham pelo menos o Ensino Médio completo e tinham idade entre 18 e 40 anos (média = 24.4375; desvio padrão = 4.15768), sendo estes compostos por 19 homens e 11 mulheres, totalizando 30 participantes, todos falantes do português brasileiro.

#### 3.2. *Corpus*

O *corpus* deste estudo é composto por trinta enunciados que foram proferidos como respostas a perguntas previamente elaboradas. Para cada pergunta captamos 15 respostas verdadeiras e 15 respostas mentirosas.

### 3.3. Procedimentos de coleta de dados

A coleta de dados foi feita com um gravador digital de *flash* Marantz PMD661, utilizando-se um microfone tipo *headset*. Os arquivos de áudio estão em formato PCM, não comprimido (gravados em formato WAV, ou mais especificamente RIFF), com taxa de amostragem de 96 kHz e 32 bits por amostra. A adoção de tais medidas proporcionou arquivos de áudio de altíssima qualidade e ótimas condições para análises acústicas aprimoradas.

Reynolds e Rendle-Short (2011) alertam para a possibilidade de pistas vocais em enunciados com conteúdo mentiroso variarem de acordo com o contexto em que a mentira é contada. Os participantes deste estudo assistiram ao curta-metragem em CGI (*computer-generated imagery*) *Brain Divided*<sup>2</sup> (Song, Song e Haworth, 2013), para posteriormente, em uma conversa com um assistente de coleta de dados, responderem perguntas sobre elementos ligados à narrativa do vídeo. Cada pergunta pode ser respondida enunciando verdades ou mentiras. No caso desta última, a mentira deveria ser uma falsificação. O contexto da conversa seguiu os moldes de uma arguição. No espaço havia uma mesa e cadeiras. O participante sentou-se à mesa com um *notebook* à sua frente. Duas instruções foram transmitidas aos participantes, uma para cada etapa dos procedimentos. A primeira delas instruiu sobre o curta-metragem: *Você assistirá a um curta-metragem. Assista com bastante atenção. Você deverá assistir duas vezes, mas caso ache necessário, poderá assistir novamente. Logo após, você irá interagir com um entrevistador como em uma entrevista sobre o filme que você assistiu. Durante a entrevista, fale da maneira mais natural possível, não se preocupe em ser formal.* A segunda etapa foi constituída pela entrevista. Durante toda essa fase o participante ainda tinha à sua frente o notebook no qual havia assistido ao filme. O propósito foi transmitir futuras orientações. O entrevistador<sup>3</sup> sentou de frente para o participante. A segunda instrução foi remetida: *Durante a entrevista, o entrevistador fará perguntas sobre o filme. Para respondê-las você seguirá as orientações dos slides. Caso o slide mostre “Mentira”, você deverá mentir para a pergunta ou comando; caso mostre “Verdade”, você deverá falar a verdade;*

<sup>2</sup> Sugerimos que o vídeo seja assistido para uma melhor compreensão dos procedimentos executados neste estudo. A versão disponibilizada aqui é dublada – a mesma apresentada para os participantes: <https://www.youtube.com/watch?v=b9VUFT5b7kE&t>.

<sup>3</sup> O assistente de coleta de dados assumiu o papel de entrevistador. Foi dito ao participante que o entrevistador não sabia dos fatos do filme. É importante destacar que enquanto o participante estava assistindo ao vídeo, o assistente estava numa sala à parte e sua entrada na sala onde o participante estava para a coleta se deu logo após o término do vídeo.

*se em ambos os casos você não souber a resposta verdadeira, basta responder que não sabe ou não lembra. Lembre-se: em ambas as situações você deverá ser convincente. O entrevistador não sabe que você irá mentir para algumas perguntas ou comandos”*. Dúvidas foram sanadas.

As instruções em *slides* orientavam o participante a dar uma versão de resposta para as perguntas feitas pelo assistente: mentir ou falar a verdade. O pesquisador controlou o dispositivo a distância. As instruções contidas nele foram antecipadas em relação às perguntas do entrevistador, de modo que, para o participante, não gerasse desconfiança para o entrevistador de que ele estava sendo guiado por elas. Em outras palavras, antes do entrevistador realizar as perguntas que mais interessavam a esta pesquisa, o participante já deveria saber qual versão utilizar para responder à pergunta. Desse modo pudemos fazer com que a latência pudesse ser analisada, pois caso a instrução no *slide* fosse visualizada após a pergunta, o tempo de ação para conduzir o dispositivo e o tempo de leitura da instrução pelo participante interfeririam diretamente no valor da latência.

Dez perguntas previamente selecionadas fizeram parte da conversa. Sete dessas perguntas foram distratores, enquanto as outras três tinham por objetivo elicitare respostas para análise. As perguntas experimentais foram: (i) *Qual o nome do rapaz?*; (ii) *Como o rapaz estava vestido?*; e (iii) *Descreva o ambiente do encontro*. Essas perguntas são denominadas no decorrer deste artigo como *Pergunta 1*, *Pergunta 2* e *Pergunta 3*.

### 3.4. Análises

Utilizamos o aplicativo computacional Praat para fazer os recortes, segmentações e anotações necessárias. Dos arquivos de áudio original, foram recortadas as três perguntas e respostas de cada participante postas em análise, visto que elas não ocorreram uma após a outra durante a entrevista. Após recortadas as três partes do áudio original, elas foram concatenadas e formaram um segundo arquivo de áudio. Esse arquivo foi utilizado no processo de segmentação e anotação.

Salientamos que este estudo não presumiu coletar dados de fala longos, como narrativas. Assim, considerando os estudos de Glukhov (1975) e Hieke, Kowal e O’Connell (1983), decidimos optar por um ponto de corte mínimo de 150 ms por ocorrência para ser caracterizado como pausa, independentemente do tipo: preenchida ou silenciosa. Não definimos um ponto de corte para a latência.

Todas as informações prosódicas necessárias para a análise deste estudo foram organizadas em tabelas de Excel para fins de análise. As análises estatísticas foram realizadas automaticamente no *software Jamovi*, versão 0.8.6.0.

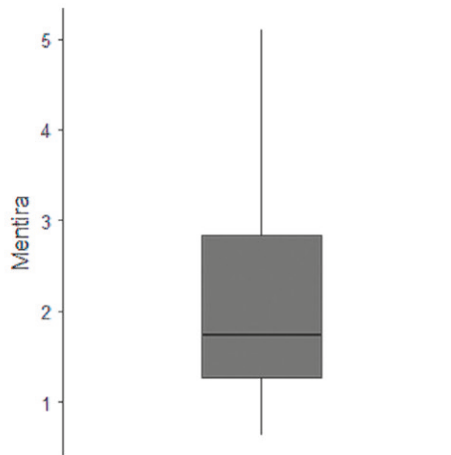
## 4. RESULTADOS

Nesta seção são apresentados os resultados deste estudo para posterior discussão. Primeiramente tratamos do valor da latência nas duas versões de enunciados. Em seguida trazemos resultados referentes à pausa.

### 4.1. Latência

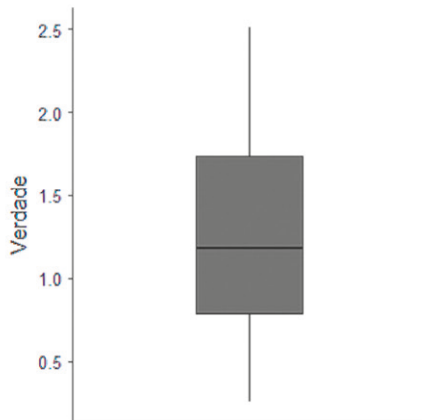
Neste estudo, a latência de resposta foi analisada nas respostas das três perguntas já descritas. As análises foram feitas em 90 registros em áudio, dos quais 45 eram respostas mentirosas e 45 verdadeiras. Em cada versão de respostas foram encontrados dois valores *outliers*. Esses dados foram descartados da análise para que eles não interferissem nos resultados.

Figura 1 – Valores de latência para resposta mentirosa.



A média de duração da latência para uma resposta mentirosa foi de 2.111 ms (desvio padrão: 1.14). O menor valor foi de 0.623 ms, enquanto que o maior valor atingiu 5.100 ms. A figura a seguir detalha os valores encontrados para a latência de resposta verdadeira:

Figura 2 – Valores de latência para resposta verdadeira.



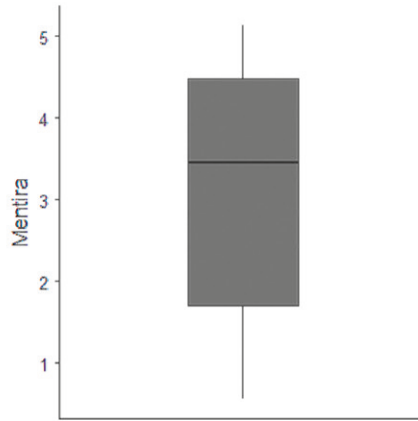
Valores mais baixos foram encontrados para a versão verdadeira de respostas. O valor mínimo foi de 0.250 ms. A média atingiu um valor de 1.260 ms (desvio padrão: 0.619). O valor mais alto de latência para resposta verdadeira foi de 2.510 ms. A diferença entre as médias de latência das duas versões é significativa. O resultado do teste *t* de *student*, com os valores da estatística, o grau de liberdade e o *p*-valor comprova esse achado:  $\text{stat.} = 4.30$ ;  $\text{df} = 39.0$ ;  $p = <.001$ . Desta forma, as respostas falsas ou mentirosas no presente estudo foram acompanhadas por uma latência significativamente maior que as respostas verdadeiras.

## 4.2. Pausa

Os resultados reportados a seguir tratam acerca dos valores de duração encontrados para a pausa nos enunciados referentes à *Pergunta 3*, já que esses permitiram uma melhor análise, tendo em vista que foram os enunciados mais longos obtidos como resposta. Todos os dados de fala analisados apresentaram apenas pausas silenciosas.

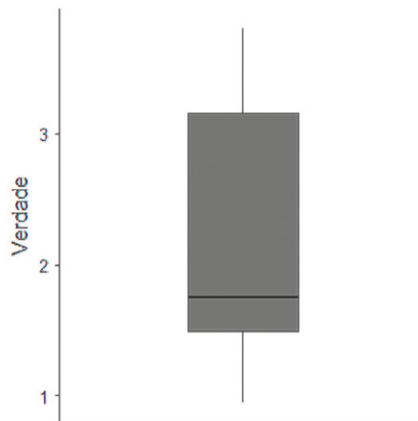
Em relação à duração da pausa, os enunciados mentirosos apresentaram um valor mínimo de inatividade vocal de 0.552 ms, enquanto o maior valor encontrado atingiu 5.130 ms. A média da duração total da pausa foi de 3.106 ms (desvio padrão: 1.59). Esses valores foram reportados tendo como base enunciados que tiveram em média 10.729 ms (desvio padrão: 5.19). A Figura 3 ilustra os valores correspondentes às pausas.

Figura 3 – Duração da pausa para resposta mentirosa.



As análises referentes à duração da pausa na versão verdadeira são detalhadas na figura a seguir:

Figura 4 – Duração da pausa para resposta verdadeira.

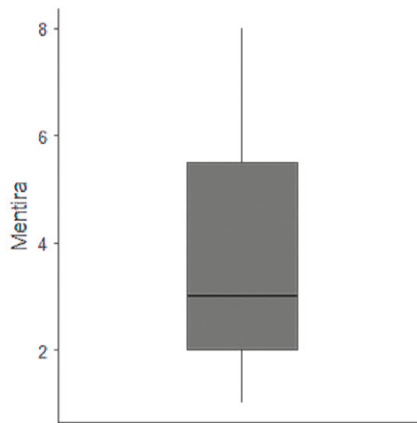


O valor médio da pausa para a versão verdadeira foi menor quando comparado ao da versão mentirosa. Sua duração foi de 2.174 ms (desvio padrão: 1.04). Valores de duração mínima de pausa no tipo de enunciado em questão atingiram 0.939 ms, enquanto que a máxima atingiu 3.806 ms. A diferença da duração média da pausa não é significativa de acordo com um teste *t*:  $stat. = 1.98$ ;  $df = 9.00$ ;  $p = 0.079$ . Esses valores foram reportados para enunciados que tiveram média de 12.117 ms (desvio padrão: 3.68).



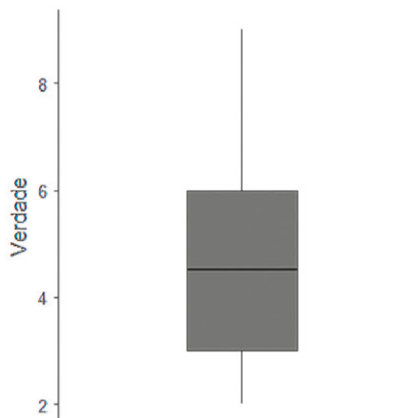
A média de ocorrência de pausa no enunciado mentiroso foi de 3.91 ocorrências (desvio padrão: 2.30). A ocorrência mínima de pausa foi uma e a máxima de até oito pausas. Assim, este estudo sugere que a média de pausas encontradas represente, em média, 3.106 ms da duração de um enunciado com duração média de 10.729 ms, o que representa 28,9% da totalidade do enunciado sendo ocupado por pausas.

Figura 5 – Ocorrência da pausa para resposta mentirosa.



Em enunciados verdadeiros, verificou-se uma média de 4.75 pausas (desvio padrão: 2.18). Nessa versão de enunciado a ocorrência mínima de pausa foram duas (desvio padrão: 2.18), enquanto o máximo de pausa foi 9.

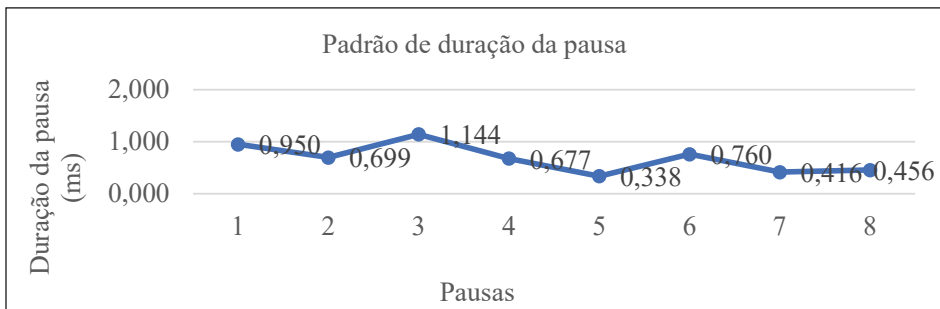
Figura 6 – Ocorrência da pausa para resposta verdadeira.



Assim como para respostas mentirosas, as verdadeiras também apresentaram pausas, e em média a ocorrência é maior. Entretanto, essa diferença não é significativa:  $\text{stat.} = -1.17$ ;  $\text{df} = 10$ ;  $p = 0.271$ . Os resultados sugerem que a média de pausa encontrada representa, em média, 2.174 ms da duração de um enunciado com duração média de 12.117 ms, o que representa 17,9% da totalidade do enunciado sendo ocupado por pausas. Os resultados reforçam a ideia de que os participantes que relataram verdades ocupam menos o seu tempo de fala com pausas do que aqueles que relataram mentiras.

A partir das análises de duração e ocorrência da pausa nas duas versões de resposta, foi possível fazer uma análise de como esses valores das pausas se comportam no enunciado. O gráfico a seguir ilustra o padrão de valores da pausa para os enunciados mentirosos:

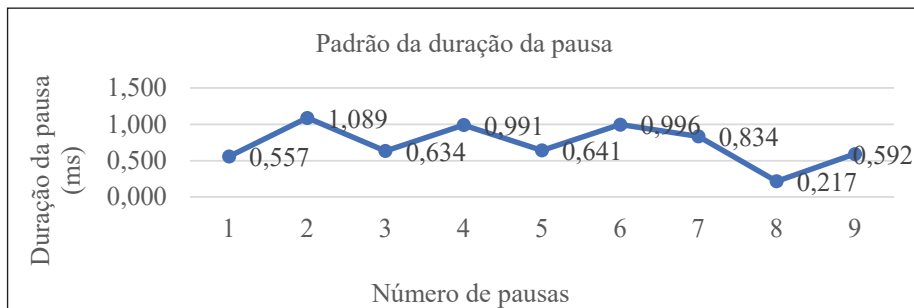
Figura 7 – Padrão duracional da pausa em respostas mentirosas.



Conforme o gráfico da Figura 7, é possível perceber que um padrão de ocorrência começa a se desenhar até a quarta pausa, no qual temos um padrão alto-baixo, isto é, a primeira pausa é maior que a segunda, que é menor que a terceira, que é maior que a quarta. Ao fazer uma análise individual da duração da pausa, este padrão alto-baixo no início dos enunciados foi característico em 61,5% deles.

Na versão verdadeira é possível observar um padrão mais consistente e que se estende até a ocorrência das últimas pausas. Um fato curioso ainda pode ser observado: o padrão alto-baixo não ocorre na versão verdadeira; o que ocorre é justamente o contrário: um padrão baixo-alto. A figura a seguir ilustra esse padrão:

Figura 8 – Padrão duracional da pausa em respostas verdadeiras.



Percebe-se que há uma frequência considerável de alternância entre o valor da pausa, isto é, parece ser padrão que, quando os participantes falavam a verdade, os valores subiam e desciam, formando um desenho de vale e pico até próximo das últimas pausas, nessa altura o padrão já começa a ganhar outra forma, não seguindo a tendência inicial. Essa característica de valor baixo-alto foi observada isoladamente em cada um dos enunciados, e o resultado foi o mesmo: 69,2% dos enunciados apresentaram essa peculiaridade.

## 5. DISCUSSÕES

A latência aparece como o elemento que se mostra mais confiável quando se trata de identificar mentiras. Ela foi o parâmetro que se apresentou estatisticamente significativo em relação às pausas analisadas dentro dos enunciados. Os dados mostram que padrões de duração mais longos desse parâmetro estão relacionados à elocução de informações fabricadas. Provavelmente isso se deve ao fato de que, para dizer algo que não acredita ser verdadeiro, o indivíduo ainda não tenha em mente uma versão alternativa e distorcida dos fatos e por isso necessite de um período maior para fabricar seu álibi. Entretanto, alguns poucos participantes deste estudo se mostraram muito eficientes, no que se refere à latência, quando sentiram a necessidade de mentir. É possível que isso seja resultado de uma capacidade cognitiva maior. Mas por outro lado, também é possível que isso seja justificado pela maior experiência do falante em situações que exijam que ele falsifique informações com certa frequência. No entanto, essas colocações são ainda suposições, pois seria necessário, por exemplo, um questionário com perguntas relacionadas a mentiras como, por exemplo, “Você costuma mentir?” ou “Com que frequência você costuma mentir?”, para que, dessa forma, as informações fossem associadas.

Ainda sobre a latência, um detalhe percebido nas respostas para a *Pergunta 3* (descreva o ambiente do encontro), considerada o estímulo que exigiria um esforço cognitivo maior para responder à pergunta, foi que as respostas mentirosas de menor latência foram em geral mais curtas, enquanto que as mais extensas apresentaram uma latência maior. É possível que essa diferença tenha ocorrido por causa do entendimento subjetivo da pergunta pelos participantes, já que alguns simplesmente responderam caracterizando o ambiente com poucas palavras, enquanto outros responderam com descrições breves, como “um restaurante” ou “uma lanchonete”. Houve também descrições bem mais detalhadas, com descrição de todo o espaço.

Durante a latência, há também a possibilidade de que tenha ocorrido lapsos de memória, e mesmo assim o participante tenha insistido em responder. Isso reforça o fato de que para esse tipo de estudo as perguntas pré-selecionadas precisam ser muito bem formuladas.

Pausas são comuns tanto em enunciados que relatam a verdade, quanto nos que relatam mentiras. No que se refere à sua duração, a pausa não apresentou resultados estatisticamente significativos para a distinção de discurso mentiroso e discurso verdadeiro. Embora os resultados tenham sido insuficientes, quando computada em conjunto, a duração das pausas é maior em enunciados que relatam mentiras do que em enunciados que relatam verdades. É possível que isso se justifique da seguinte forma: quando sentem necessidade de mentir, os participantes precisam de uma latência significativamente fora do comum para responderem a uma pergunta. É possível que essa latência incomum tenha oferecido o tempo necessário para o participante idealizar seu álibi, embora ainda pouco consistente. Ao perceber que está demorando mais que o habitual para responder a uma pergunta, o falante se vê obrigado a iniciar sua resposta para que não dê indícios de que estará mentindo. E então, só a partir do momento de fala, utilizando pausas mais longas que o comum, é que o participante consegue ter sucesso em sua resposta. Dessa forma, isso pode ter exigido mais de suas capacidades cognitivas para elaborar versões de respostas que não fossem verdadeiras. Ao que tudo indica, isso pode corroborar com o que diz Goldman-Eisler (1968) ao afirmar que a duração mais longa das pausas está associada à imprevisibilidade de uso de certas palavras, resultando numa exigência cognitiva maior.

Assim como a duração, a ocorrência da pausa também apresentou resultados não significativos, mas com um resultado curioso. Embora haja a crença de que um número fora do comum de pausas ocorre mais em enunciados que relatam mentiras – e algumas pesquisas já comprovaram isso, gerando divergência

com estudos que diziam o contrário – ocasionando disfluências, esse fato não foi confirmado neste estudo, cujos resultados associam um maior número de pausa em enunciados que relatam a verdade. É possível que isso tenha ocorrido em razão da extensão da resposta dos participantes instruídos a dizer a verdade, pois, como relatado no parágrafo anterior, pessoas instruídas a dizer mentiras costumam proferir respostas menos extensas que pessoas instruídas a dizer a verdade.

O fato de ter ocorrido mais pausas em relatos verdadeiros também pode se justificar pelo que já foi postulado por Ekman (1985), que diz que em algumas situações as pessoas conseguem monitorar suas falas para não acabarem revelando verdades.

É possível que uma decisão metodológica também tenha implicado nesse resultado. A duração da pausa neste estudo foi estipulada em um valor de 150 ms acima. É possível que pausas com essa duração ou até um pouco maior seja característica de enunciados que relatam verdades. Então, seria interessante que o ponto de corte para a consideração da pausa fosse maior, tendo em vista que pausas ocupam mais tempo no enunciado – apesar de resultados não significativos – e em geral, esse tempo é ocupado mais em favor da duração das pausas que da ocorrência.

Há também um padrão específico da pausa em relação à forma como ocorrem e a duração (Figuras 7 e 8). Enunciados que relataram mentiras apresentaram uma pausa curta, seguida de uma longa, assumindo um padrão alto-baixo. Em enunciados que relatam a verdade foi constatado o contrário: a primeira pausa é mais curta que a segunda, que é maior que a terceira, que é mais curta que a quarta e assim por diante, o que remete a um efeito baixo-alto. Neste último aspecto, o padrão se apresentou mais longo e padronizado. É possível que isso reflita o controle do discurso por parte do falante em situações como a que foi utilizada neste estudo. Em enunciados que relatam mentiras, o efeito é contrário, embora seja menos consistente. Esse tipo de ocorrência menos padronizada talvez reflita a falta de controle da situação que o participante tinha quando estava mentindo.

## 6. CONCLUSÃO

Este estudo apresentou resultados de análises da latência de resposta e de duração e ocorrência de pausas em enunciados com conteúdo falso em oposição a enunciados com conteúdo que se crê verdadeiro. Foram objetos de análise a

latência e a pausa. As análises indicaram que a latência se caracteriza como um elemento estatisticamente significativo na distinção de um enunciado que relata mentiras em oposição a um que relata verdades. Por outro lado, embora haja, em análises preliminares, a indicação de que a pausa apresenta duração maior em enunciados associados a uma mentira e duração menor em enunciados associados a uma verdade, os testes estatísticos mostraram que os resultados não são significativos. Em relação à ocorrência da pausa, os resultados revelaram que menos pausas são utilizadas durante um enunciado que relata mentiras, enquanto que a maior incidência desse elemento está associada a enunciados que relatam a verdade. No entanto, nesse ponto os resultados também não se mostraram estatisticamente significativos. Esse constituinte prosódico ainda apresentou um padrão de comportamento distinto nos dois tipos de enunciado. Enunciados que relatam mentira apresentam um padrão alto-baixo, já os que relatam verdade têm um aspecto baixo-alto.

Este trabalho apresenta uma contribuição para os estudos linguísticos, sobretudo àqueles ligados à Fonética Forense. Além da contribuição para a linguística, outras áreas também se beneficiam de estudos como este, como, por exemplo, a das artes dramáticas, além de serem também um recurso bastante expressivo que pode ser explorado pela retórica e pela oratória.

## REFERÊNCIAS

ALMEIDA, P. T. A. **A prosódia nas atitudes dos falantes: o caso da ironia**. Orientador: Vera Pacheco. 2016. 91 f. Dissertação (Mestrado) – Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, 2016. Disponível em: [https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id\\_trabalho=4912349](https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=4912349). Acesso em: 18 dez. 2020.

ALMEIDA, A. N. S; OLIVEIRA JR., M; ALMEIDA, R. A. S. A velocidade de fala como pista acústica da emoção básica de raiva. **Diadorim**, Rio de Janeiro, v. 17, n. 2, p. 198-211, 2015. DOI 10.17074/1980-2552.2016n18v2p(198). Disponível em: <https://revistas.ufrj.br/index.php/diadorim/article/view/4076>. Acesso em: 27 dez. 2020.

ALVES, L. M. **O estudo da persuasão na fala do vendedor**. Orientador: César Augusto da Conceição Reis. 2002. 137 f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Minas Gerais, 2002.

AUBERGÉ, V. Morphology of Prosody Directed by Functions: the Example of a Step Model Developed at ICP. **Proceedings of the 1st Conference on Speech Prosody**, [s. l.], 2002. Disponível em: <http://sprosig.org/sp2002/pdf/auberge.pdf>. Acesso em: 26 dez. 2020.

BALLONE, G. J. Sobre a Mentira. **PsiquWeb**, Internet, p. 1-4, 16 fev. 2010. Disponível em: <http://www.psiqweb.med.br/site/?area=NO/LerNoticia&idNoticia=246>. Acesso em: 10 nov. 2020.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. **Revista de estudos da linguagem**, Belo Horizonte, v. 20, n. 1, p. 11-27, 2012. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/2571/0>. Acesso em: 10 dez. 2020.

BARBULESCU, A.; RONFARD, R.; BAILLY, G. Which prosodic features contribute to the recognition of dramatic attitudes? **Speech commun. – Speech Communication**, [s. l.], v. 95, p. 78-86, 2017. Disponível em: <https://hal.inria.fr/hal-01643330/file/prosodic-features-contribute.pdf>. Acesso em: 9 jan. 2021.

BENUS, S. *et al.* Pauses in deceptive speech. **Proceedings of the 3rd speech prosody conference**, Nova York, p. 1-4, 2006. Disponível em: [http://www.cs.columbia.edu/~sbenus/Research/Benus\\_et\\_al\\_fps\\_SP06.pdf](http://www.cs.columbia.edu/~sbenus/Research/Benus_et_al_fps_SP06.pdf). Acesso em: 2 dez. 2020.

BOERSMA, P.; WEENINK, D. Praat: Doing phonetics by computer [Computer program]. Version 6.0.36. 2017. Disponível em: <http://www.Praat.org/>. Acesso em: 21 nov. 2017.

*BRAIN Divided*. Direção: Joon Soo Song, Joon Shik Song, Josiah Haworth. Estados Unidos: TheCGBros, 2013. Disponível em: <https://www.youtube.com/watch?v=JuyB7NO0EYY>. Acesso em: 15 mar. 2017.

BROTHERTON, P. Speaking and not speaking: process for translating ideas into speech. In: SIEGMAN, A.; FELDESTSTEIN, S. **Of Time and Speech**. Hillsdale, Nova Jersey: Lawrence Erlbaum, 1979. p. 178-209.

CARSON, T. L. **Lying and deception**. Oxford: Oxford University Press, 2010. 288p.

COUPER-KUHLEN, E. **An introduction to English prosody**. Tübingen: Max Niemeyer, 1986. 239 p.

CRYSTAL, D. **Encyclopedia of the English Language**. Cambridge: Cambridge University Press, 1995. 489 p.

DEPAULO, B. M. *et al.* Cues to deception. **Psychological Bulletin**, Santa Bárbara, v. 129, n. 1, p. 74-118, 2003. DOI 10.1037/0033-2909.129.1.74. Disponível em: <https://smg.media.mit.edu/library/DePauloEtAl.Cues%20to%20Deception.pdf>. Acesso em: 2 dez. 2020.

DUEZ, D. Perception of silent pauses in continuous speech. **Language and Speech**, [s. l.], v. 28, n. 4, p. 377-389, 1985. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/002383098502800403>. Acesso em: 26 nov. 2020.

EKMAN, P. **Telling lies: clues to deceit in the marketplace, politics, and marriage**. Nova York: Norton & Company, 1985. 366 p. ISBN 0-393-30872-3.

FÓNAGY, I. As funções modais da entoação. **Cadernos de Estudos Linguísticos**, Campinas, v. 25, p. 25-65, 1993. Disponível em: <https://www.periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8636884/4606>. Acesso em: 30 dez. 2020.

FORD, M. Sentence planning units The: implications for the speaker's representation of meaningful relations underlying sentences. *In*: BRESNAN, J. **Mental Representation of Grammatical Relations**. Cambridge, Massachusetts: MIT Press, 1982. p. 797-827.

FUJISAKI, H. Prosody, models, and spontaneous speech. *In*: SAGISAKA, Y.; CAMPBELL, N.; HIGUCHI, N. **Computing Prosody: Computational Models for Processing Spontaneous Speech**. Nova York: Springer, 1997. p. 27-42. ISBN 978-1-4612-7476-6.

GLUKHOV, A. A. Statistical analysis of speech pauses for Romance and Germanic languages. **Soviet Physics**, [s. l.], v. 21, p. 71-72, 1975.

GOLDMAN-EISLER, F. **Psycholinguistics: experiments in spontaneous speech**. Nova York: Academic Press, 1968. 169 p.

HIEKE, A. E, KOWAL, S; O'CONNELL, D. C. The trouble with "articulatory" pauses. **Language and Speech**, [s. l.], v. 26, n. 3, p. 203-214, 1983. DOI 10.1177/002383098302600302. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/002383098302600302>. Acesso em: 21 nov. 2020.



LOVE, J; DROPMANN, D; SELKER, R. Jamovi Project. Jamovi (Version 0.8) [Computer Software]. 2018. Disponível em: <https://www.jamovi.org>. Acesso em: 01 jan. 2018.

LU, Y.; AUBERGÉ, V.; RILLIARD, A. Do you hear my attitude? Prosodic perception of social affects in Mandarin. **6th International Conference on Speech Prosody (Speech Prosody 2012)**, Shanghai, China, p. 685-688, 2012. Disponível em: <https://hal.archives-ouvertes.fr/hal-00744696/document>. Acesso em: 9 jan. 2021.

MASIP, J., GARRIDO, E., AND HERRERO, C. Defining deception. **Anales de Psicología**, Espanha, v. 20, n. 1, p. 147-171, 2004. Disponível em: [https://www.um.es/analesps/v20/v20\\_1/12-20\\_1.pdf](https://www.um.es/analesps/v20/v20_1/12-20_1.pdf). Acesso em: 21 dez. 2020.

METTS, S. An exploratory investigation of deception in close relationships. **Journal of Social and Personal Relationships**, Londres, v. 6, p. 159-179, 1989. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/026540758900600202>. Acesso em: 14 out. 2020.

MITCHELL, R. L. C.; ROSS, E. D. Attitudinal prosody: what we know and directions for future study. **Neuroscience & Biobehavioral Reviews**, [s. l.], v. 37, n. 3, p. 471-479, 2013.

MOURA, L. S. **O papel da prosódia na expressão de atitudes de ataque ao ethos no discurso político**. Orientador: Leandra Batista Antunes. 2016. 151 p. Dissertação (Mestrado) – Universidade Federal de Ouro Preto, Mariana, Minas Gerais, 2016. Disponível em: [https://www.repositorio.ufop.br/bitstream/123456789/6496/1/DISSERTA%C3%87%C3%83O\\_PapelPros%C3%B3diaExpress%C3%A3o.pdf](https://www.repositorio.ufop.br/bitstream/123456789/6496/1/DISSERTA%C3%87%C3%83O_PapelPros%C3%B3diaExpress%C3%A3o.pdf). Acesso em: 9 jan. 2021.

MOZZICONACCI, S. J. L.; HERMES, D. J. A study of intonation patterns in speech expressing emotion or attitude: Production and perception. **IPO Annual Progress Report**, [s. l.], v. 32, p. 154-160, 1997. Disponível em: <https://pdfs.semanticscholar.org/d0ac/15125a839a7ff7dd1c7a2487957129ba4039.pdf>. Acesso em: 3 dez. 2020.

O'CONNELL, D. C.; KOWAL, S. Pausology. In: SEDELOW, W. A.; SEDELOW, S.Y. **Computers in Language Research 2**. Nova York: Mouton Publishers, 1983, p. 221-301.

OLIVEIRA, B. F. V. **A prosódia na expressão das atitudes de dúvida, incerteza e incredulidade no português brasileiro**. Orientador: César Reis. 2011. 195 p. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2011. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/DAJR-8MZPMC/1/1392m.pdf>. Acesso em: 21 nov. 2020.

OLIVEIRA JR., M. **Prosodic features in spontaneous narratives**. Orientador: Zita McRobbie. 2000. 275 p. Tese (Doutorado) – Simon Fraser University, Canadá, 2000. Disponível em: <https://www.collectionscanada.gc.ca/obj/s4/f2/dsk2/ftp02/NQ61670.pdf>. Acesso em: 26 nov. 2020.

PACHECO, V.; ALVES, P. T. A prosódia nas atitudes dos falantes: o caso da ironia. **Confluência**, Rio de Janeiro, n. 53, 2017. Disponível em: <http://lp.bibliopolis.info/confluencia/rc/index.php/rc/article/view/189>. Acesso em: 9 jan. 2021.

PARMENTER, C. E; TREVIÑO, S. N. The length of the sounds of a middle westerner. **American Speech**, [s. l.], v. 10, n. 2, p. 129-133, 1935.

PAULA, K. M. **O papel da prosódia na ironia como expressão de atitude**. Orientador: César Reis. 2012. 90 p. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2012. Disponível em: [https://repositorio.ufmg.br/bitstream/1843/LETR-96NNDK/1/disserta\\_\\_o\\_final\\_karen\\_26\\_09\\_2012.pdf](https://repositorio.ufmg.br/bitstream/1843/LETR-96NNDK/1/disserta__o_final_karen_26_09_2012.pdf). Acesso em: 26 dez. 2020.

REYNOLDS, E; RENDLE-SHORT, J. Cues to deception in context: Response latency/gaps in denials and blame shifting. **British Journal of Social Psychology**, [s. l.], v. 50, n. 3, p. 431-449, 2011. DOI 10.1348/014466610X520104. Disponível em: <https://bpspsychub.onlinelibrary.wiley.com/doi/epdf/10.1348/014466610X520104>. Acesso em: 30 out. 2020.

RODRIGUES, H. Algumas verdades sobre a mentira. **Revista Polêm!ca**, Rio de Janeiro, v. 16, n. 1, p. 42-62, 2016. DOI 10.12957/polemica.2016.21335. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/polemica/article/view/21335>. Acesso em: 28 out. 2020.

SILVA, J. P. G. **Análise dos aspectos prosódicos na expressão da certeza e da dúvida no português brasileiro**. Orientador: César Reis. 2008. 172 f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

SPENCE, K; ARCIULI, J; VILLAR, G. Markers of deception in Italian speech. **Frontiers in Psychology**, [s. l.], v. 3, p. 1-9, 2012. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00453/full>. Acesso em: 9 nov. 2020.

TROVILLIO, P. V. A history of lie detection. **Journal of Criminal Law and Criminology**, [s. l.], v. 29, n. 6, p. 848-881, 1939. Disponível em: <https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=2844&context=jclc>. Acesso em: 8 out. 2020.

VRIJ, A; HEAVEN, S. Vocal and verbal indicators of deception as a function of lie complexity. **Psychology, Crime, & Law**, [s. l.], v. 5, p. 203-215, 1999.

WICHMANN, A. The attitudinal effects of prosody, and how they relate to emotion. **Proceedings of the ISCA Workshop on Speech and Emotion**, Newcastle, 2000. Disponível em: [https://isca-speech.org/archive\\_open/archive\\_papers/speech\\_emotion/spem\\_143.pdf](https://isca-speech.org/archive_open/archive_papers/speech_emotion/spem_143.pdf). Acesso em: 29 dez. 2020.

ZUCKERMAN, M; DEPAULO, B. M; ROSENTHAL, R. Verbal and nonverbal communication of deception. *In*: BERKOWITZ, L. (ed.). **Advances in experimental social psychology**. Nova York: Academic Press, 1981. v. 14, p. 1-59.

# CORRELATOS ELETROFISIOLÓGICOS DO PROCESSAMENTO DE FRONTEIRA PROSÓDICA NO PORTUGUÊS BRASILEIRO

*Musiliyu Oyedeji*

## 1. INTRODUÇÃO

Linguagem escrita geralmente tem uma estrutura clara, devido ao uso de convenções tipográficas, tais como sinais de pontuação (i.e., vírgulas, pontos). A fala, no entanto, envolve uma série efêmera de sons conectados sem marcadores de fronteiras óbvios. Não obstante, há mecanismos específicos que sinalizam a estrutura do discurso falado. Diversos estudos mostraram que a prosódia é frequentemente usada para organizar a fala em um arranjo hierárquico, coerente com macro unidades (Geluykens & Swerts, 1994; Grosz & Hirschberg, 1992; Passonneau & Litman, 1993). Tentativas de explicar aspectos da fonologia que não podiam ser capturados por um modelo linear da fonologia levaram ao desenvolvimento da teoria da fonologia prosódica (Lieberman & Prince, 1977; Selkirk, 1980, 1984; Nespor & Vogel, 1983; Beckman & Pierrehumbert, 1986; Ladd, 1986; Ladd & Campbell, 1991), que propõe que o fluxo do discurso é organizado em um número finito de unidades fonológicas, ou constituintes prosódicos. Esses constituintes formam uma hierarquia de componentes arranjados do menor para o maior, como se segue: sílaba ( $\sigma$ ), pé ( $\Sigma$ ), palavra fonológica ( $\omega$ ), grupo clítico (C), frase fonológica ( $\Phi$ ), frase entoacional (IP) e enunciado fonológico (U).

Com respeito à existência de uma diferença sistemática entre IP e U, Nespor e Vogel (1986) assinalam-nos como distintos na aplicação diferencial de certas regras fonológicas. Notadamente, U consiste em, pelo menos, um IP e usualmente se estende ao longo do comprimento da cadeia dominada pelo maior nó na árvore sintática, referido como  $X^n$ . Entretanto, U não é simplesmente a contrapartida fonológica de  $X^n$ , desde que combina duas ou mais sentenças em uma unidade de mais alto nível. Quando comparado com IP, U se caracteriza pelo contorno entoacional da forma linguística que sempre tem pausas tanto no início quanto no fim, bem como significado completo (Zhang, 2017). Por outro lado, um conjunto relativo de parâmetros acústicos para constituintes prosódicos em diferentes linguagens é reportado na fronteira do IP (daqui por diante, *IPB*), ressaltando o comprimento dos segmentos pré-fronteira (e.g., Beckman & Edwards, 1990; Beckman et al., 1992; Wightman et al., 1992; Cambier-Langeveld, 2000; Turk & Shattuck-Hufnagel, 2000, 2007; Byrd et al., 2006), variação de pitch e pausa (e.g., Nespor & Vogel, 1986; Venditti et al., 1996; Hirst & Di Cristo, 1998; Féry et al., 2011).

Estudos sobre o português brasileiro propõem que o *IPB* e a fronteira do U (doravante *UB*) são marcados principalmente por três pistas prosódicas principais: alongamento final, isto é, um aumento na duração da vogal tônica das palavras imediatamente anteriores às fronteiras (Moraes, 1995; Serra, 2009; Serra & Frota, 2009), variação de pitch, por exemplo, tons de fronteira caracterizados por um aumento ou possivelmente queda de pitch para frase entoacional e principalmente uma queda de pitch para enunciado (Frota & Vigário, 2000; Tenani, 2002; Fernandes, 2007; Serra, 2009) e, embora não seja obrigatório, uma pausa (Chacon and Fraga, 2014). Esses estudos relataram que *IPB* e *UB* são definidos por um contorno nuclear, combinando um acento pitch (H + L \* ou L + H \*) e tom de fronteira (L% ou H%), alcançado na sílaba final do IP, seguido por uma pausa potencial (em línguas tônicas como inglês e português, o contorno nuclear é a melodia na sílaba nuclear e sílaba(s) pós-tônica subsequente(s) e como nas línguas românicas, proeminência nuclear está mais à direita dentro do IP) (Nespor & Vogel, 2007; Ladd, 2008). Ressalte-se que a escolha do acento de pitch e tons de fronteira depende do tipo de frase e da posição IP em relação a U (Frota et al., 2015). Tenani (2002) observou que a última sílaba tônica do IP que não coincide com o final do U é marcada por um evento tonal L + H \*, e seguida por um tom de fronteira alto H% quando ocorre uma pausa (caracterizando uma “subida continuativa”). Em palavras trissilábicas pré-fronteira e pseudopalavras, com um padrão de tonicidade paroxítona, Santana and Leal (2011) descobriram que sílabas tônicas, seguidas por sílabas pós-tônicas são mais longas no *IPB* em comparação com constituintes prosódicos de nível inferior. Serra (2016) revelou

que os *IPBs* esperados são percebidos como quebras mais fortes na leitura da fala em comparação com a fala espontânea. Soncin et al. (2017) observou que a percepção da pausa pode depender exclusivamente do contorno nuclear que caracteriza o *IPB*. Uma vez que o papel da prosódia como dispositivo na estruturação do discurso falado é bem descrito no que diz respeito à produção, o presente estudo visa antes investigar seu impacto na percepção, especialmente quando se trata da percepção de uma perspectiva neurocognitiva.

Com o advento de novas técnicas como a eletroencefalografia (doravante EEG), a propagação da atividade nervosa no cérebro pode ser medida, em tempo real, de forma não invasiva (Araújo et al., 2004), durante a produção ou percepção da fala. O sinal de EEG gerado em resposta à estimulação externa ou interna é chamado de Potencial Relacionado a Eventos (doravante ERP) (Kutas & Hillyard, 1980, 1984). As técnicas de extração de ERP fornecem informações valiosas sobre a natureza e o curso de tempo do processamento da linguagem.

No passado recente, diferentes componentes de ERP relacionados à língua foram identificados. Os ERPs têm sido associados a diferentes processos de compreensão apresentados em modelos de compreensão de sentenças auditivas (por exemplo, Friederici, 2011). Steinhauer et al. (1999) encontraram pela primeira vez a resposta de ERP de *Closure Positive Shift* (doravante CPS) para o processamento de *IPBs*. Esse efeito do ERP revelou uma mudança positiva centroparietalmente distribuída, cerca de 500 milissegundos (doravante ms) após o início da *IPB*.

O CPS foi encontrado em estudos de várias outras línguas, como holandês (Kerkhofs et al., 2007; Bögels et al., 2010), japonês (Wolff et al., 2008), chinês (Li & Yang, 2009), inglês (Itzhak et al., 2010), francês (Annie C. Gilbert et al. 2010) e português europeu (Leone-Fernandez et al., 2016a, b). O CPS não é acionado pela pausa no *IPB* per se; em vez disso, a variação do tom e o alongamento da sílaba imediatamente anterior à pausa são suficientes para evocar o componente. Isso foi demonstrado em um experimento no qual a pausa no *IPB* foi excluída (Steinhauer et al., 1999). A última suposição foi verificada para ouvintes como crianças mais velhas e adultos, com conhecimento sintático suficiente para maior habilidade de compreensão da linguagem, mas não para bebês e crianças (Männel & Friederici, 2009, 2011). Isso sugere que a pausa serve inicialmente como uma dica alvo para estruturar a entrada de fala, mas que não é necessária para o fraseado entoacional, uma vez que conhecimento suficiente sobre a estrutura prosódica e sintática é adquirido (Friederici, 2011). Experimentos adicionais com adultos mostraram que o CPS também pode ser obtido quando apenas a informação prosódica de uma sentença é entregue (ou seja, quando a

informação segmentada não está disponível). Sob esta condição, onde os investigadores usaram material experimental apresentando informações semânticas, sintáticas e fonêmicas decrescentes (por exemplo, frases jabberwocky, em que todas as palavras de conteúdo foram substituídas por palavras sem sentido; frases de pseudopalavra, em que todas as funções e todas as palavras de conteúdo são substituídas por palavras sem sentido; e frases deslexicalizadas, contorno de entonação cantarolada de uma frase removendo todo o conteúdo segmentar), o CPS é relatado sobre o hemisfério direito (Pannekamp et al., 2005). Independentemente de pistas de pitch ou formas linguísticas significativas, o CPS também foi evocado especificamente por trechos de fala marcados por alongamento (Gilbert et al., 2010, 2012, 2015). Além disso, o CPS é relatado para a leitura de sentenças desencadeada por uma vírgula indicando um limite de frase sintática (Steinhauer & Friederici, 2001; Steinhauer, 2003; Kerkhofs et al., 2008). Assim, o CPS pode ser considerado um componente ERP que está correlacionado com o fraseado prosódico tanto quando realizado explicitamente no fluxo de voz, quanto quando realizado secretamente em sentenças escritas (Friederici, 2011).

Estudos anteriores de ERP levaram a um perfil geral do CPS. Em termos de distribuição no couro cabeludo, o CPS é encontrado bilateralmente e é maior nos eletrodos da linha média (Bögels et al., 2011). Alguns estudos relataram uma extensão para eletrodos anteriores (por exemplo, Bögels et al., 2010; Pannekamp et al., 2005). Itzhak et al. (2010) também relataram uma distribuição CPS frontocentral. O CPS geralmente começa por volta ou mesmo antes do início da pausa (Steinhauer, 2003), embora Kerkhofs et al. (2007) tenham encontrado um início posterior. Seu deslocamento parece ser acionado pelo início da palavra imediatamente após a pausa e dura em torno de 500-700 ms (Pauker et al., 2011). Estudos anteriores localizaram o início de um CPS observando um pico negativo máximo inicial (Pannekamp et al., 2005; Kerkhofs et al., 2008; Bögels et al., 2010; Pauker et al., 2011).

Os ERPs são geralmente calculados tomando a média de vários alcances do EEG bloqueado pelo tempo até o início do evento de interesse. No caso do CPS, uma vez que o tom limite começa na última palavra imediatamente anterior à pausa, e esse alongamento pré-final da sílaba está claramente presente apenas a partir do início da última sílaba tônica antes da pausa, Bögels et al. (2011) consideraram este início como o “início da fronteira prosódica”, e como o ponto de bloqueio de tempo teoricamente mais apropriado para a análise CPS.

O grau em que pistas acústicas e linguísticas determinam a explicitação do CPS ainda não foi estabelecido (Peter et al., 2014). A maioria dos estudos CPS

apoia a ideia de que as pistas prosódicas acústicas são os principais responsáveis pela geração do CPS (Brouwer et al., 2012), enquanto as pistas linguísticas modulam sua amplitude e topografia do couro cabeludo (Kerkhofs et al., 2007). Além disso, Kerkhofs et al. (2007) argumentou que um CPS maior é evocado por uma fronteira prosódica menos esperada, que é, portanto, mais saliente. Isso indica que a amplitude do CPS depende da saliência dos marcadores de fronteira prosódica (Kerkhofs et al., 2008). Segundo Li and Yang (2009), embora a pausa não seja um fator necessário para evocar o CPS, ela pode modular sua latência e amplitude de início. Steinhauer (2003) relata que a amplitude e a duração do CPS são maiores na apresentação da linguagem auditiva do que visual e parecem se correlacionar positivamente com o grau em que, durante a leitura silenciosa, as representações fonológicas (ou seja, ouvir a “voz interior”) são ativadas.

Para investigar melhor a modulação (em termos de latência de início, amplitude, duração e distribuição no couro cabeludo) do CPS, medimos o CPS no *IPB* e *UB*. Em uma sentença declarativa, uma vez que o *IPB* interno é marcado por diferentes pistas prosódicas em comparação com o *IPB* que coincide com *UB* (Tenani, 2002), formulamos a hipótese de que (i) o CPS será eliciado em resposta ao processamento de fronteiras prosódicas em português brasileiro, como foi demonstrado para várias outras línguas, e (ii) o CPS será modulado em função das diferenças entre as pistas prosódicas que marcam as duas fronteiras.

## 2. MÉTODO

### 2.1 Material

Os itens experimentais consistem em enunciados que contêm um *IPB* interno e um *UB* interno (coincidindo também com um *IPB*). Três versões de cada declaração foram criadas. A primeira versão (Tipo A) foi a base para as outras duas versões (Tipo B e Tipo C) conforme segue no Quadro 1, onde (#) significa *IPB*, (# \*) significa “sem *IPB*”, (%) significa *UB* e (% \*) significa “sem *UB*”.

Quadro 1 – Exemplo de material experimental

<b>Tipo A</b>
<i>[Assim que Paula viu sua aMIga # ela fechou a jaNEla #]#[Foi abrir a porta]#.</i>
<b>Tipo B</b>
<i>[Assim que Paula viu sua aMIga #* de infância # ela fechou a janela #]#[Foi abrir a porta]#.</i>
<b>Tipo C</b>
<i>[Assim que Paula viu sua amiga # ela fechou a jaNEla %* da sala #]#[Foi abrir a porta]#.</i>



É importante ressaltar que as análises neste presente estudo focaram apenas nas “bordas alvo” caracterizadas no Quadro 1 em fonte itálico e negrito. No Tipo A, a fronteira após a oração subordinada coincide com um *IPB* e a fronteira após a oração principal com *UB*. No Tipo B, a porção da frase correspondente à oração subordinada (no Tipo A) não é seguida imediatamente por um *IPB* – “sem *IPB*” (*NIPB*). E no Tipo C, a porção da frase correspondente à oração principal (no Tipo A) não é seguida imediatamente por um *UB* – “sem *UB*” (*NUB*). As últimas palavras da oração subordinada e da oração principal foram sempre um complemento verbal trissilábico (indicado em negrito e itálico) doravante denominado “palavras-alvo”, apresentando um padrão de acentuação paroxítona (indicado em maiúsculas na penúltima sílaba). As três versões do material experimental diferiram apenas no que diz respeito aos adjuntos adnominais (indicados pelas palavras sublinhadas) adicionados à oração subordinada no Tipo B e à oração principal no Tipo C. Em termos de número de sílabas, o valor médio de comprimento da oração subordinada foi  $M = 10$  sílabas,  $DP = 0,75$ , e para a oração principal,  $M = 8$  sílabas,  $DP = 0,89$ .

Um total de 134 itens foram criados. Destes, 120 itens experimentais foram selecionados após um teste de aceitabilidade. Além disso, foram construídos 120 “preenchedores”. Os “preenchedores” eram semelhantes em comprimento aos itens experimentais. Também apresentaram três tipos, cada um deles caracterizado por diferentes padrões sintáticos, conforme segue no Quadro 2, onde (#) significa *IPB* e (%) para *UB*.

Quadro 2 – Exemplo de preenchedores

<b>Tipo F<sub>a</sub> (1-40)</b>
<i>[Quando viram o pescador chegando # lhe ajudaram a tirar o barco da água #]%</i> .
<b>Tipo F<sub>b</sub> (41-80)</b>
<i>[O prefeito da cidade fraturou a perna direita durante a festa de fim de ano]%</i> .
<b>Tipo F<sub>c</sub> (81-120)</b>
<i>[Berto revelou o segredo da família no tribunal]%</i> [ <i>O pai tinha outra esposa</i> ]%

Para garantir a naturalidade e aceitabilidade dos itens experimentais, conduzimos um teste de aceitabilidade com um grupo de 30 alunos, todos falantes nativos do português brasileiro. Um total de 402 itens (incluindo as três versões dos 134 itens originais) foi usado no referido estudo teste de aceitabilidade. Eles foram classificados nos três tipos de enunciados (A, B e C) descritos no Quadro 2. A fim de evitar um efeito de familiaridade ao apresentar itens semelhantes ao

mesmo participante, dividimos os participantes em três grupos de dez indivíduos e usamos um *Latin square design* para criar três listas de materiais. Os “preenchedores” não foram incluídos no estudo de normatização.

O teste de aceitabilidade foi conduzido on-line usando a plataforma *Qualtrics* na qual a aceitabilidade dos materiais foi julgada usando uma escala numérica ou rótulos categóricos (uma escala do tipo *Likert*), que poderia ser posteriormente convertida em números (Schmidt-Nielsen, 1992). Os participantes foram instruídos a atribuir a cada um dos enunciados uma escala numérica de 1 (menos aceitável) a 7 (mais aceitável). Eles começaram com tentativas de prática (dez novas sentenças foram usadas para evitar efeitos de repetição) antes do experimento de normatização real. O valor médio dos números obtidos na escala numérica para cada um dos 134 conjuntos de enunciados foi calculado e, em seguida, relatamos a Média (média de todas as 134 médias) que foi  $M = 6$ ,  $DP = 0,39$ . Isso mostrou que, em grande medida, os enunciados foram considerados aceitáveis. Com base nos resultados do experimento de normatização, um total de 120 conjuntos de itens experimentais que receberam a maior taxa de aceitabilidade foi selecionado para nosso experimento ERP.

## 2.2 Estímulos: Descrição geral e análise acústica

Um locutor profissional do sexo masculino e falante nativo do Português Brasileiro gravou os estímulos de fala usando um microfone omnidirecional em condições acústicas adequadas (uma cabine acústica) com uma velocidade de fala normal ( $M = 5,589$  sílabas por segundo (sílabas / s),  $DP = 0,4$ ). Essa velocidade de fala foi medida dividindo-se o número de sílabas do estímulo por sua duração (excluindo a duração da pausa). O resultado corrobora achados anteriores, que relataram velocidade média de fala variando de 3,2 a 5,5 Syll / s para um falante nativo fluente do Português Brasileiro (Oliveira, 2000; Martins et al., 2008). O locutor foi instruído a repetir cada enunciado pelo menos três vezes, para que pudéssemos escolher a melhor gravação possível para os itens experimentais e preenchedores. Esses itens foram selecionados com base no julgamento de um grupo de 5 alunos, todos falantes nativos do português brasileiro. A gravação foi digitalizada em 44.100 Hz com uma profundidade de 16 bits por amostra.

As gravações selecionadas, 360 itens experimentais, foram analisadas usando o software Praat (Boersma and Weenink, 2007) para primeiro verificar quais pistas prosódicas foram associadas ao *IPB* e *UB* (no Tipo A) em comparação com o *NIPB* (no Tipo B) e *NUB* (no Tipo C). A medição da duração foi extraída automaticamente executando um script Praat chamado

“Analyse\_tier” (Hirst, 2012). Em segundo lugar, verificamos as diferenças acústicas entre *IPB* e *UB*.

A medição da pista acústica foi realizada nas três sílabas imediatamente anteriores a cada fronteira, ou seja, a sílaba pré-tônica (Pre\_), a sílaba tônica (Str\_) e a sílaba pós-tônica (Pos\_). O alongamento pré-final da sílaba foi determinado pela medição da duração média das vogais das sílabas em cada posição. Para essa medição, os valores de duração foram normalizados usando uma fórmula de z-score para minimizar os efeitos da duração intrínseca e da variabilidade da duração das vogais. Para cada sílaba, a variação do pitch foi determinada pelo cálculo do intervalo médio do núcleo da vogal. Os valores médios de F0 e intensidade também foram medidos no núcleo vocálico das sílabas.

A pausa foi avaliada medindo-se a duração do “intervalo de silêncio” marcando *IPB* e *UB* nos itens experimentais, quando ocorria. Esses intervalos de silêncio foram categorizados em dois tipos: (i) *P\_IPB*, intervalo de silêncio marcando *IPB*; (ii) *P\_UB*, intervalo de marcação de silêncio *UB*. Qualquer intervalo de silêncio superior a 150 ms foi considerado uma pausa, seguindo Kowal et al. (1983).

As comparações dos dados da análise acústica associados ao *IPB*, *NIPB*, *UB* e *NUB* foram realizadas da seguinte forma: (i) *IPB* vs. *NIPB*; (ii) *UB* vs. *NUB*; (iii) *IPB* vs. *UB*. Os dados foram submetidos a t-tests de duas amostras. Relatamos apenas diferenças estatisticamente significativas entre pistas acústicas.

### 2.2.1 Relatórios da análise acústica

Com relação à pausa, os “intervalos de silêncio” *P\_IPB* variaram entre 34 ms e 150 ms, e *P\_UB* variou entre 260 e 388 ms. Assim, *P\_UB* foi sempre longo o suficiente para ser considerado uma pausa; não foi o caso do *P\_IPB*. A comparação revelou que *P\_UB* foi mais longo em duração [*P\_UB*, M = 305 ms (SE = 0,55); *P\_IPB*, M = 113 ms (SE = 1,21); t (718) = -143,8, p < 0,001].

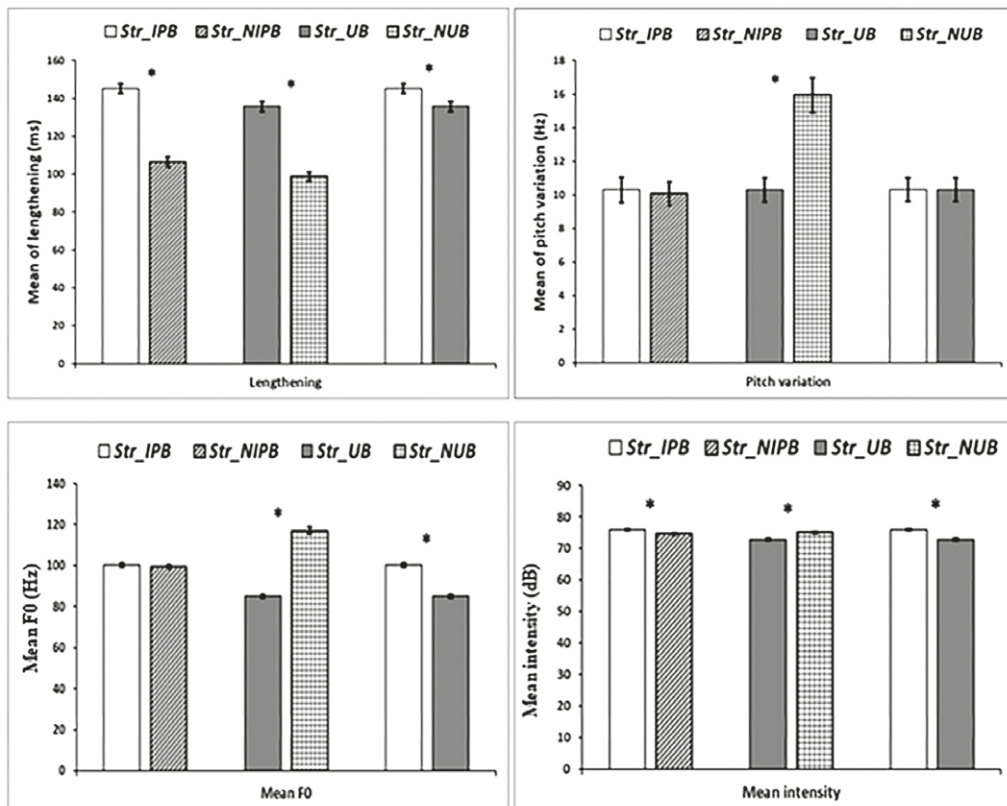
Os dados da análise acústica das três sílabas anteriores ao *IPB* e *UB* foram comparados com os anteriores ao *NIPB* e ao *NUB*, respectivamente. Também comparamos os resultados da análise acústica das três sílabas anteriores ao *IPB* com as anteriores ao *UB*. Os dados para cada sugestão foram submetidos a t-tests de duas amostras. Uma visão geral dos resultados é apresentada no Quadro 3, onde (\*) indica  $p \leq 0.05$ ; (\*\*) indica  $p \leq 0.01$  e (\*\*\*) indica  $p \leq 0.001$ .

Quadro 3 – Resultados da análise estatística de comparação dos dados da análise acústica associados ao *IPB*, *NIPB*, *UB* e *NUB*

	Alongamento (z-score)	Varição do Pitch (Hz)	Média de F0 (Hz)	Média de intensidade (dB)
Comparação das sílabas	t (df) Média (SE)	t (df) Média (SE)	t (df) Média (SE)	t (df) Média (SE)
<i>Pre_IPB</i> vs. <i>Pre_NIPB</i>			6.7 (237)*** 114.8 (23) > 96.9 (1.1)	2.2 (238)* 75.3 (0.3) > 74.3 (0.3)
<i>Str_IPB</i> vs. <i>Str_NIPB</i>	10.4 (238)*** 1.1 (0.05) > 0.2 (0.05)			3.5 (238)*** 75.9 (0.2) > 74.6 (0.2)
<i>Pos_IPB</i> vs. <i>Pos_NIPB</i>	5.2 (238)*** -0.3 (0.3) > -0.6 (0.03)	2.3 (229)** 8 (0.49) > 6.4 (0.4)		
<i>Pre_UB</i> vs. <i>Pre_NUB</i>		3 (232)** 11.9 (0.9) < 16 (1)	-2.5 (232)** 122.8 (2.2) > 115.5 (1.6)	
<i>Str_UB</i> vs. <i>Str_NUB</i>	6.9 (238)*** 0.7 (0.05) > 0.2 (0.03)	4.3 (236)*** 10.3 (0.7) < 15.9 (1)	16.6 (236)*** 84.7 (0.7) < 116 (1.7)	5.9 (237)*** 72.9 (0.2) < 75.1 (0.2)
<i>Pos_UB</i> vs. <i>Pos_NUB</i>	2.9 (238)*** -0.4 (0.03) > -0.6 (0.03)	3 (232)** 5.9 (0.6) < 8.7 (0.7)	18.9 (202)*** 78.1 (0.7) < 127.6 (2.1)	18.5 (238)*** 62.8 (0.4) < 73.5 (0.3)
<i>Pre_IPB</i> vs. <i>Pre_UB</i>		2.3 (239)** 8.7(0.6) < 11.3 (0.9)	-2.4 (239)** 114.8 (2.3) < 122.8 (2.2)	
<i>Str_IPB</i> vs. <i>Str_UB</i>	2.5 (238)** 1.1 (0.05) > 0.9 (0.05)		12.1 (236)*** 100.4 (0.9) > 84.7 (0.7)	8.1 (239)*** 75.9 (0.2) > 72.9 (0.2)
<i>Pos_IPB</i> vs. <i>Pos_UB</i>		2.5 (202)** 8 (0.4) > 5.9 (0.6)	9.1 (202)*** 104.1 (2.3) > 78.1 (0.7)	14.6 (239)*** 71.3 (3.5) > 62.8 (0.4)

As dicas acústicas associadas às sílabas tônicas pré-fronteira são representadas na Gráfico 1. As sílabas foram comparadas da seguinte forma: *Str\_IPB* vs. *Str\_NIPB*; *Str\_UB* vs. *Str\_NUB*; *Str\_IPB* vs. *Str\_UB*. Essa ilustração está relacionada ao fato de que as respostas do ERP são fixadas no tempo para o início das sílabas tônicas pré-fronteira, conforme detalhado na seção “Paradigma experimental EEG” do presente estudo.

Gráfico 1 – Pistas acústicas associadas ao *IPB*, *NIPB*, *UB* e *NUB*. As sílabas tônicas são comparadas como segue: *Str\_IPB* vs. *Str\_NIPB*; *Str\_UB* vs. *Str\_NUB* and *Str\_IPB* vs. *Str\_UB*. Os asteriscos (\*) indicam diferenças significativas.



É importante ressaltar que a análise mostra que *IPB* e *UB* são sinalizados por diferentes pistas acústicas. O *IPB* foi marcado por alongamento pré-final da sílaba mais longa, maior média de F0 e intensidade média em relação ao *UB*.

### 2.3 Participantes

Participaram do experimento ERP 30 alunos voluntários (15 homens; idade média: 24,31 anos; DP: 3,1), falantes nativos do PB, da Universidade Federal de Alagoas. Eles eram destros, conforme avaliado por uma versão em português do Brasil do inventário de lateralidade de Edinburgh (*Edinburgh Handedness Inventory*) (Oldfield), sem deficiência auditiva ou história prévia de distúrbios neurológicos ou psiquiátricos baseados em autodeclaração.

## 2.4 Paradigma experimental de EEG

Três listas de estímulos foram usadas para o experimento. Cada lista continha apenas uma versão das falas que serviram de estímulo. Uma ordem pseudoaleatória de materiais experimentais e preenchedores foi projetada com a restrição de que o mesmo tipo de estímulo não fosse apresentado duas vezes consecutivas. Os participantes foram divididos em três grupos de dez indivíduos. Cada grupo teve acesso a uma lista.

Em consonância com a abordagem metodológica sugerida em um estudo anterior (Bögels, Schriefers, Vonk, and Chwilla, 2011), e considerando os resultados da análise acústica do material experimental do presente estudo, optamos por fixar o tempo de nossa resposta ERP para o início das sílabas tônicas pré-fronteira. Os códigos de gatilho de evento foram definidos no início das sílabas tônicas pré-fronteira nas “palavras-alvo” no Tipo A (*IPB*, para fronteira prosódica IP; *UB*, para fronteira prosódica U) e no início das mesmas sílabas tônicas nos correspondentes “palavras-alvo” que não precedem imediatamente as fronteiras prosódicas no Tipo B e Tipo C (*NIPB*, para nenhuma fronteira prosódica IP; e *NUB*, para nenhuma fronteira prosódica U).

Cada lista continha 240 estímulos divididos em 6 blocos experimentais de 40 estímulos de no máximo 5 minutos com pausas de descanso entre os blocos. Os estímulos foram fornecidos por meio de fones de ouvido (*Sennheiser hd280 pro*) usando o software E-Prime.

Para o experimento, os participantes estavam sentados em uma cadeira confortável colocada a 1 metro de distância da tela do computador. No início, uma instrução escrita os informava sobre o experimento. Eles foram solicitados a ouvir atentamente as sentenças. O experimento começou com uma sessão prática de 10 novas sentenças construídas para os ensaios. Uma tentativa sempre era iniciada com um prompt escrito “próximo estímulo” de 100 ms, seguido de um ponto de fixação (+) no monitor. A frase começou 200 ms após a fixação. Os participantes foram solicitados a olhar para o ponto de fixação para evitar movimentos dos olhos e piscar até o deslocamento de uma frase. Os estímulos foram seguidos aleatoriamente por um prompt escrito (uma palavra). Na tarefa, os participantes foram solicitados a indicar por meio de um pressionamento de tecla se a palavra solicitada estava presente ou ausente no estímulo imediatamente anterior (em 5% dos estímulos para cada participante). Essa tarefa foi dada para garantir que os participantes prestassem atenção ao ouvir as frases.

## 2.5 Aquisição dos sinais digitais de EEG

O experimento foi conduzido no Laboratório de Psicolinguística da Universidade Federal de Alagoas, Brasil. Para o registro de EEG, usamos um sistema Geodesic EEG 400 (*Electrical Geodesics, Inc*), com um HydroCel Sensor Net de 256 canais, uma série Net Amps 400 de amplificadores de EEG e um conjunto completo de software Net Station para aquisição e análise de EEG. Verificamos se os sensores estavam em bom contato com o couro cabeludo, se o sinal do EEG não continha ruído elétrico e se os eletrodos foram ajustados até que suas impedâncias fossem mantidas abaixo de 5 k $\Omega$ . Os dados do EEG foram registrados com um filtro passa-alta em 0,1 Hz e a taxa de amostragem definida em 1.000 Hz. A referência de gravação on-line para todos os sensores nos dados de EEG foi Cz. O software Net Station gravou dados de EEG continuamente junto com códigos de gatilho de evento. Os dados EEG foram exportados em formato bruto, usando ferramentas Net Station, para análises off-line.

## 2.6 Extração de ERPs (Potenciais Relacionados a Evento)

Para análise EEG off-line, rodamos EEGLAB (Delorme and Makeig, 2004) no MATLAB e importamos os dados brutos de EEG. Em seguida, carregamos um arquivo de localização de canal para o HydroCel GSN de 256 canais e referenciamos os dados novamente para a referência média. Em seguida, os dados foram filtrados usando um filtro passa-baixa de 30 Hz. No ERPLAB (Lopez-Calderon and Luck, 2014), dividimos os segmentos de EEG em seções (épocas) que começaram 200 ms antes do início das sílabas tônicas nas “palavras-alvo” e terminaram 2.000 ms após esse início. Épocas de EEG que continham grandes artefatos musculares foram removidas dos dados por inspeção visual. A correção do artefato ocular foi realizada usando análise de componente independente (ICA), conforme implementado no EEGLAB (função ‘eeg\_runica’). Componentes independentes com características conhecidas de piscar de olhos (com base no espectro de energia da atividade, topografia do couro cabeludo e atividade durante as tentativas) foram identificados visualmente para cada participante. As contribuições desses componentes foram então removidas do EEG contínuo. Os artefatos foram detectados e removidos automaticamente usando um procedimento de pico a pico de janela móvel, com uma janela móvel de 200 ms, uma etapa de janela de 100 ms e um limite de voltagem de 100  $\mu$ V.

As épocas foram calculadas em média para produzir um ERP para cada condição de contorno prosódica e cada participante teve pelo menos 70% de tentativas aceitas por condição [*IPB*,  $M = 94,25$  ( $SD = 7,46$ ); *NIPB*,  $M = 93,08$

(SD = 8,7); *UB*, M = 93,85 (SD = 5,69); *NUB*, M = 94,08 (SD = 8,1)]. Formas de onda de ERP individuais foram calculadas para obter a média geral de ERPs para cada condição. Uma vez que os pares de condições – *IPB* vs. *NIPB* e *UB* vs. *NUB* – continham as mesmas informações lexicais e prosódicas até o início das sílabas tônicas das “palavras-alvo” (ver Figura 1 para *IPB* vs. *NIPB* e Figura 2 para *UB* vs. *NUB*), as comparações de ERPs de grande média em nossas janelas de tempo [0-2000 ms] devem refletir o processamento de pistas prosódicas marcando o *IPB* e *UB*.

Figura 1 – Exemplo de pares de condições para frase entoacional. *IPB* no Tipo A e *NIPB* no Tipo B. Os retângulos pretos tracejados indicam o contorno entoacional do IP até o início da sílaba tônica pré-fronteira no Tipo A, e o contorno da mesma parte do IP sem fronteira prosódica no Tipo B. O retângulo preto sólido indica os tons de fronteira do *IPB*. Os círculos tracejados no Tipo A indicam: I1, I2 e I3, representando *Pre\_IPB*, *Str\_IPB* e *Pos\_IPB* respectivamente; U1, U2 e U3, representando *Pre\_UB*, *Str\_UB* e *Pos\_UB* respectivamente. O círculo tracejado no Tipo B indica: I1 \*, I2 \* e I3 \* simbolizando *Pre\_NIPB*, *Str\_NIPB* e *Pos\_NIPB* respectivamente. PI e PU, indicados pelas setas, representam *P\_IPB* e *P\_UB* respectivamente.

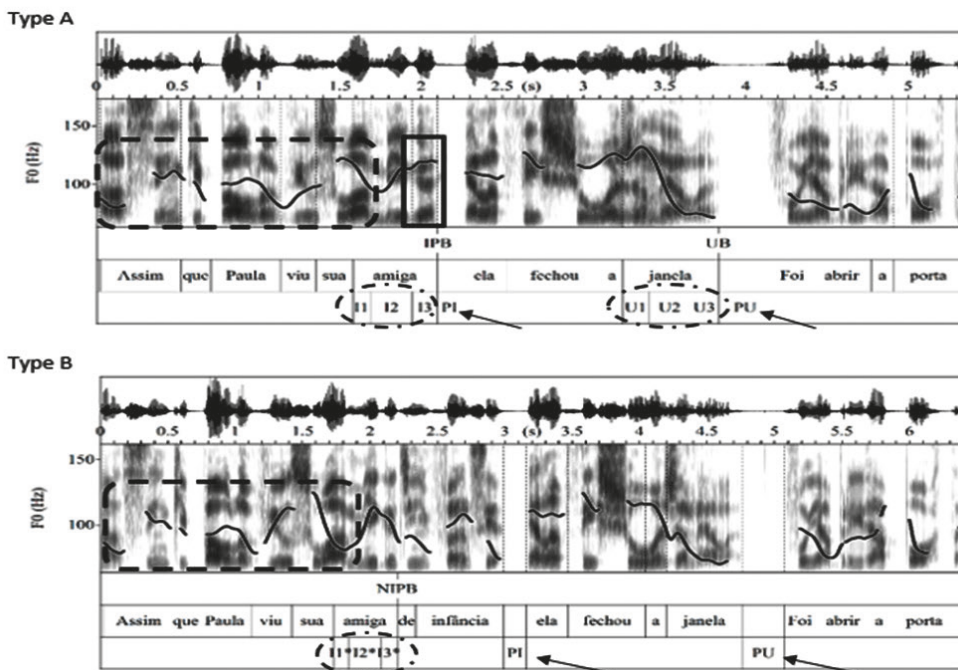
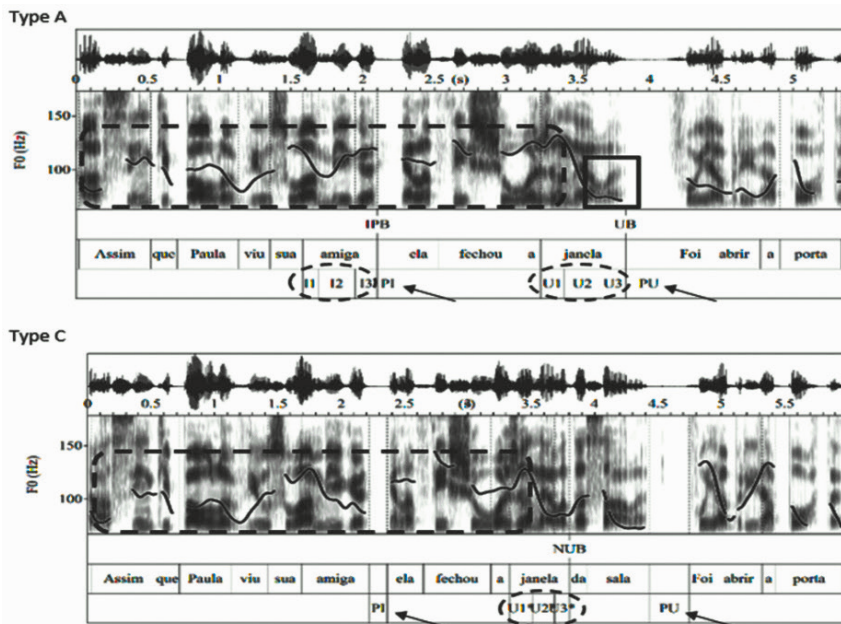




Figura 2 – Exemplo de pares de condições para frase entoacional. *UB* no Tipo A e *NUB* no Tipo C. Os retângulos pretos tracejados indicam o contorno entoacional de U até o início da sílaba tônica pré-fronteira no Tipo A, e o contorno da mesma porção de U sem fronteira prosódica no Tipo C. O retângulo preto sólido indica tons de limite de *UB*. Os círculos tracejados no Tipo A indicam: I1, I2 e I3, representando *Pre\_IPB*, *Str\_IPB* e *Pos\_IPB* respectivamente; U1, U2 e U3, representando *Pre\_UB*, *Str\_UB* e *Pos\_UB* respectivamente. O círculo tracejado no Tipo C indica: U1\*, U2\* e U3\*, simbolizando *Pre\_NUB*, *Str\_NUB* e *Pos\_NUB* respectivamente. PI e PU, indicados pelas setas, representam *P\_IPB* e *P\_UB* respectivamente.



## 2.7 A análise de ERPs

Para as respostas de ERP fixadas no tempo para o início das sílabas tônicas das “palavras-alvo”, calculamos a amplitude média em seis janelas de tempo consecutivas de 200 ms (de 0 a 1.200 ms, considerando a possível ocorrência dos efeitos CPS). Seguindo estudo prévio de ERP auditivo (Peter et al., 2014), consideramos 19 eletrodos organizados da seguinte forma: Eletrodos de linha média: Cz, Pz e Fz; Os eletrodos laterais foram agrupados em seis regiões de interesse (ROIs): anterior direito (F4, F8, FT8), posterior direito (P4, P8, CP4, TP8), anterior esquerdo (F3, F7, FT7), posterior esquerdo (P3, P7, CP3, TP7), central esquerda (FC3) e central direita (FC4).

Duas ANOVA de medidas repetidas totalmente cruzadas foram calculadas separadamente nas janelas de tempo: (i) sobre os eletrodos da linha média, incluindo os fatores Fronteiras (IP vs. U), Condições (fronteira vs. “sem fronteira”)

e Eletrodos (Fz vs. Cz vs. Pz): (ii) e sobre os eletrodos laterais, incluindo fatores Fronteiras (IP vs. U), Condições (fronteira vs. “sem fronteira”), Hemisférios (direito vs. esquerdo) e Localizações (anterior vs. central vs. posterior).

Apenas relatamos amplitudes médias envolvendo o fator “Condições” (fronteira vs. “sem fronteira”), uma vez que está diretamente relacionado aos objetivos deste estudo. Se uma interação significativa foi encontrada entre “Condições” e qualquer outro fator, ANOVA post-hoc foram calculados para entender o efeito desse fator para cada janela de tempo e condição separadamente.

### **3. RESULTADOS**

As formas de onda de ERP da média geral nos 19 eletrodos, incluindo o eletrodo Cz representativo, para *IPB* vs. *NIPB* e *UB* vs. *NUB*, são ilustradas nas Figuras 3 e 4, respectivamente.

Figura 3 – Grande média dos ERPs para as condições *IPB* e *NIPB* nos 19 eletrodos, incluindo o eletrodo Cz. O retângulo preto tracejado indica um “pico negativo inicial”. Nos enunciados ilustrativos, a seta sólida indica o ponto de bloqueio de tempo para medir os ERPs e as sílabas tônicas das “palavras-alvo”, que estão em maiúsculas. Hash (#) significa *IPB* e asterisco hash (# \*) significa *NIPB*.

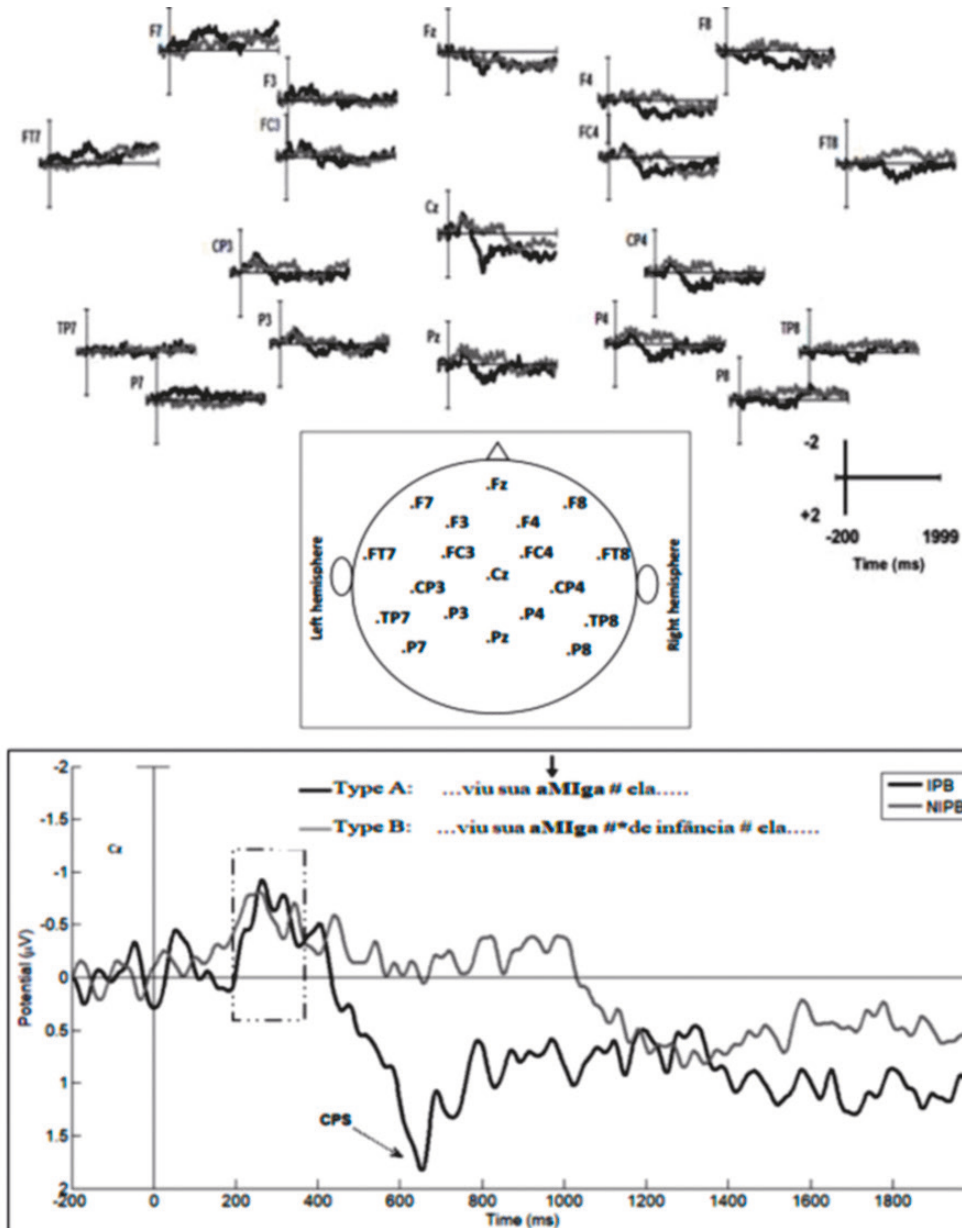
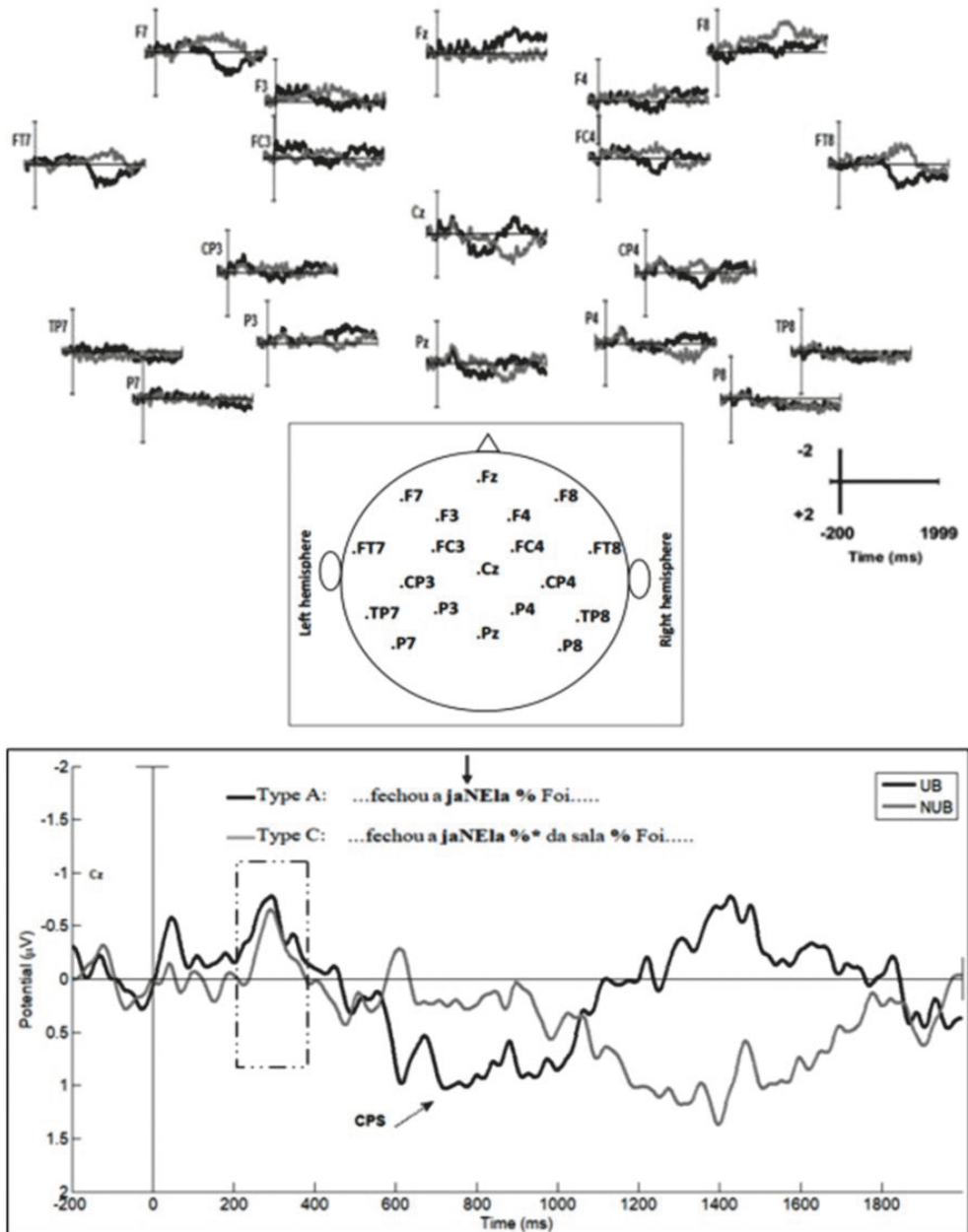


Figura 4 – ERPs de média geral para as condições *UB* e *NUB* nos 19 eletrodos, incluindo o eletrodo Cz. O retângulo preto tracejado indica um “pico negativo inicial”. Nos enunciados ilustrativos, a seta sólida indica o ponto de bloqueio de tempo para medir os ERPs e as sílabas tônicas das “palavras-alvo”, que estão em maiúsculas. Porcentagem (%) significa *UB* e asterisco de porcentagem (%\*) para *NUB*.



Para todas as condições, nos eletrodos Cz (parte inferior das Figuras 3 e 4), observamos um pequeno “pico negativo inicial” indicado com um retângulo preto tracejado, que atinge sua amplitude de pico em ~250 ms após o início das sílabas tônicas das “palavras-alvo”. Essa observação foi corroborada pela análise estatística apresentada no Quadro 4.

Quadro 4 – Efeitos significantes (ANOVAs) em amplitudes médias através da janela de tempo de [0–400 ms] para o componente “pico negativo inicial” de ERPs. Tws = Janelas temporais; B = Condição de fronteira; NB = Condição “sem fronteira”; Cond = Condição; Hem = Hemisfério; × = interação. (\*) indica  $p \leq 0.05$ ; (\*\*) indica  $p \leq 0.01$ ; (\*\*\*) indica  $p \leq 0.001$ ; (>) indica Maior do que; (<) indica Menor do que

Tws (ms)	Eletrodos centrais			Regiões de interesse lateral (ROIs)		
	Efeito	F(df)	Média (SE)( $\mu$ V) B vs. NB	Efeito	F(df)	Média (SE)( $\mu$ V) B vs. NB
0 – 200				Cond	6.9 (1,1908)**	-0.26 (0.03) < -0.14 (0.03)
200 – 400				Cond × Hem	18.9 (1,1908)***	
				Hem esquerdo	25.4 (1,954)***	-0.36 (0.04) < -0.06 (0.04)

O resultado mostra que o “pico negativo inicial” foi mais significativo nos eletrodos laterais. Além disso, era mais proeminente na localização do hemisfério esquerdo para condições de contorno (*IPB* e *UB*), em comparação com as condições sem contorno (*NIPB* e *NUB*).

Após o “pico negativo inicial”, observamos uma ampla deflexão positiva, indicada pelas setas tracejadas na Figura 3 (de ~ 500 a 1.200 ms) e na Figura 4 (de ~ 600 a 1.200 ms), para as condições *IPB* e *UB*, respectivamente. A análise estatística apresentada no Quadro 5 fundamentou esta observação.

Quadro 5 – Efeitos significantes (ANOVAs) em amplitudes médias através da janela de tempo de [400–1.200 ms] para o componente CPS de ERPs. Tws = Janelas temporais; B = Condição de fronteira; NB = Condição “sem fronteira”; Cond = Condição; Hem = Hemisfério; Elec = eletrodos; Fron = Fronteira; Loca = localização; × = interação. (\*) indica  $p \leq 0.05$ ; (\*\*) indica  $p \leq 0.01$ ; (\*\*\*) indica  $p \leq 0.001$ ; (>) indica Maior do que; (<) indica Menor do que.

Tws (ms)	Eletrodos centrais			Regiões de interesse lateral (ROIs)		
	Efeito	F(df)	Média (SE)( $\mu$ V) B vs. NB	Efeito	F(df)	Média (SE)( $\mu$ V) B vs. NB
400 – 600	Cond	4.2 (1,348)*				
	Cond × Fron	5.5 (1,348) **	0.2 (0.08) vs. -0.1 (0.07)			
	IP	10.5 (1,178)**	0.3 (0.1) vs. -0.14 (0.11)			
600 – 800	Cond	23.2 (1,348)***	0.6 (0.10) vs. 0.03 (0.08)	Cond	10.2 (1,1896)**	0.05 (0.0) vs. -0.14 (0.04)
	Cond × Elec	3.9 (2,348)*		Cond × Hem	11.9 (1,1896)***	
	Cz	19.4 (1,118)***	1.05 (0.1) vs. 0.01 (0.1)	Hem direito	19.9 (1,958)***	0.24 (0.0) vs. -0.17 (0.06)
	Pz	12.6 (1, 118)**	0.5 (0.14) vs. -0.1 (0.1)	Cond × Fron	6.6 (1,1896)**	
800–1.000	Cond	9.7 (1,348)***	0.4 (0.1) vs. 0.007 (0.08)	IP	9.9 (1,958)***	0.14 (0.07) vs. -0.14 (0.05)
	Cond × Elec	4.4 (2,348)**		Cond	31.6 (1,1896)***	0.16 (0.05) vs. -0.2 (0.04)
	Cz	11.2 (1,118)**	0.7 (0.1) vs. -0.02 (0.1)	Cond × Hem	14.6 (1,1896)***	
	Pz	9 (1,118)**	0.4 (0.15) vs. -0,1 (0.14)	Hem direito	40.5 (1,958)***	0.37 (0.07) vs. -0.27 (0.06)
1.000–1.200				Cond	32.3 (1,1896)***	0.24 (0.05) vs. -0.19 (0.05)
				Cond × Hem	6.7 (1,1896)**	
				Hem direito	31.4 (1,958)***	0.38 (0.08) vs. -0.24 (0.07)
				Cond × Loca	9.18 (2,1896) ***	
				Loca anterior	25.8 (1,718)***	0.34 (0.1) vs. -0.4 (0.1)
				Loca central	10.4 (1,238)**	0.4 (0.1) vs. -0.13 (0.1)

Para as condições de fronteira em comparação com as condições “sem fronteira”, o resultado mostra que através da janela de tempo de [400-600] o amplo efeito de deflexão positiva foi apenas significativo nos eletrodos da linha média e mais proeminente no IP. Em [600-800] e [800-1.000] o efeito foi significativo para ambos (i) eletrodos de linha média (mais proeminentes em Cz e Pz) e (ii) eletrodos laterais (mais proeminentes no hemisfério direito para IP em [600-800] e apenas mais proeminente no hemisfério direito em [800-1.000]). No que diz respeito à janela de tempo de [1.000-1.200], o efeito foi apenas significativo nos eletrodos laterais e mais proeminente no hemisfério direito, localizações anterior e central.

## 4. DISCUSSÃO

Os resultados da análise acústica mostraram diferenças na força das pistas prosódicas marcando *IPB* e *UB* (também coincidindo com um *IPB*) em comparação com *NIPB* e *NUB*. Principalmente, essas diferenças também foram observadas na comparação entre *IPB* e *UB*. Em um enunciado (agrupando mais de um *IPB*), esses resultados indicam claramente, em consonância com a literatura (Moraes, 1995; Wichmann, 2000; Oliveira, 2000; Frota & Vigário, 2000; Tenani, 2002; Fernandes, 2007; Serra, 2009; Serra & Frota, 2009; Chacon & Fraga, 2014), que, do ponto de vista da produção, as fronteiras prosódicas de hierarquia diferente são caracterizadas de forma diferente. Eles também demonstram que muitas dessas diferenças aparecem antes do início do próprio limite.

Os achados comportamentais revelaram respostas de ERP significativamente mais positivas ao processamento de *IPB* interno e *UB* interno (também coincidindo com um *IPB*) quando comparados com a ausência dessas fronteiras. Foi observada que a detecção das fronteiras prosódicas envolveu as regiões da linha média, hemisférica direita e esquerda do cérebro. No entanto, essa detecção envolveu predominantemente o hemisfério direito, localizações central e anterior. Presumimos que a ampla deflexão positiva de ERPs que observamos em nossos resultados é o CPS em resposta ao processamento de *IPB* e *UB* nos enunciados usados no presente estudo.

Além disso, antes da resposta CPS, de ~ 200 a 400 ms após o início das sílabas tônicas de “palavras-alvo” para *IPB* e *UB*, encontramos um “pico negativo inicial” que foi mais pronunciado sobre os eletrodos do hemisfério esquerdo. Consideramos que esse efeito ERP é a negatividade pré-CPS também observada em estudos anteriores (Pannekamp et al., 2005; Kerkhofs et al., 2008; Bögels et al., 2010; Pauker et al., 2011). Essa “negatividade precoce” pode ser uma consequência do processamento de pistas prosódicas iniciais que marcam as fronteiras prosódicas. Como também foi observado para *NIPB* e *NUB*, pode-se sugerir que seja desencadeado pelo esforço neural associado à expectativa de uma fronteira prosódica, devido ao conhecimento linguístico dos ouvintes. A negatividade menos pronunciada que observamos para *NIPB* e *NUB* é certamente devido ao fato de que as fronteiras prosódicas eventualmente não ocorreram.

A distribuição mais pronunciada da “negatividade inicial” no couro cabeludo no hemisfério esquerdo corrobora achados anteriores (Pannekamp et al., 2005; Kerkhofs et al., 2008; Bögels et al., 2010), embora, notavelmente, Pauker et al. (2011) tenham encontrado uma negatividade pré-CPS lateralizada

à direita. Em consonância com estudos anteriores de ERP (Steinhauer et al., 1999; Steinhauer & Friederici, 2001; Pannekamp et al., 2005; Peter et al., 2014) sobre processamento de fronteira prosódica em sentenças normais, o CPS foi elicitado em condições de escuta como um marcador de *IPB* e *UB* (também coincidindo com um *IPB*). A comparação deste componente ERP para o *IPB* interno e o *UB* interno revelou descobertas interessantes. A distribuição do CPS no couro cabeludo relatada para ambas as fronteiras prosódicas na linha média e predominantemente sobre o hemisfério direito, locais central e anterior, é uma indicação de que há uma semelhança no processamento deles. Curiosamente, a distribuição mais pronunciada do CPS no hemisfério direito que encontramos está de acordo com o modelo de via dupla (Friederici & Alter, 2004), que assume que a prosódia linguística é preferencialmente processada no hemisfério direito do cérebro porque transmite informações em um nível suprasegmental, enquanto os processos de fala que requerem análise de informações segmentais (por exemplo, reconhecimento de fonemas) são assumidos como envolvendo predominantemente o hemisfério esquerdo. Além disso, a distribuição frontocentral relatada para sentenças em inglês (por exemplo, Itzhak et al., 2010, p. 10), distribuição do couro cabeludo bilateral e amplamente na linha média (Bögels et al., 2011, p. 428) do CPS corroboram a distribuição do couro cabeludo na linha média, localização anterior e central que encontramos. Embora a distribuição do CPS no couro cabeludo varie até certo ponto ao longo dos estudos – provavelmente em função da variação experimental e das condições de estímulo (Pannekamp et al., 2005, p. 6) – podemos afirmar que os efeitos do CPS não são apenas relatados com uma ampla distribuição em todo o couro cabeludo (por exemplo, Steinhauer et al., 1999; Kerkhofs et al., 2007; Leone-Fernandez et al., 2016a, b), mas também com localizações mais específicas, incluindo uma distribuição centroparietal para sentenças alemãs (Steinhauer and Friederici).

Em relação à latência, amplitude e duração do início, as diferenças na força das pistas prosódicas que marcam *IPB* e *UB* interagiram com a elicitação do efeito CPS. Notadamente, as estimativas de latência de início vieram de uma inspeção visual das grandes médias que substanciamos com análise estatística. Nos eletrodos da linha média, o CPS para o *IPB* pareceu mostrar um início precoce (~ 400 ms), duração mais longa (~ 600 ms) e amplitude mais alta em comparação com *UB* com início tardio (~ 600 ms) e duração mais curta (~ 400 ms) do efeito CPS. Nos eletrodos laterais também, sobre o hemisfério direito em locais central e anterior, o CPS para o *IPB* pareceu mostrar um início precoce



(~ 600 ms), maior duração (~ 600 ms) e maior amplitude em comparação com *UB* com um início (~ 800 ms), duração mais curta (~ 400 ms) do efeito CPS. Essas latências de início do CPS estão, em ter grande medida, de acordo com estudos anteriores (Steinhauer et al., 1999; Pannekamp et al., 2005), onde o CPS estava presente cerca de 500 ms após o início do limite do IP. Além disso, as durações estão, em grande medida, de acordo com a duração de 500 a 700 ms do efeito CPS observado em pesquisas anteriores (Steinhauer et al., 1999; Pauker et al., 2011). Notavelmente, essas comparações foram estabelecidas com cautela, uma vez que, ao contrário do estudo atual, a maioria das outras pesquisas não bloqueou o tempo do CPS para o início da sílaba tônica pré-fronteira.

É importante ressaltar que essas observações fundamentaram a hipótese que formulamos sobre a modulação (no que diz respeito à latência de início, amplitude e duração) do CPS, em função das diferenças na força das pistas prosódicas que marcam *IPB* interno e *UB* interno nos enunciados que nós usamos. A latência de início precoce, duração mais longa e amplitude relativamente maior do efeito CPS para o *IPB* em comparação com a latência de início tardio, duração reduzida e amplitude menor para o *UB* são uma indicação de que as pistas prosódicas acústicas que sinalizam ambas as fronteiras variaram e foram processadas de maneiras diferentes.

Crucialmente, o *IPB* foi marcado pelo alongamento da sílaba pré-final e variação do tom, enquanto o *UB*, além dessas pistas, foi marcado por uma pausa. Além da pausa – que não aciona o CPS per si (Steinhauer et al., 1999; Leone-Fernandez et al., 2016a, b) – a análise acústica mostrou que as pistas acústicas que marcam as duas fronteiras eram diferentes. *IPB* foi marcado por alongamento de sílaba pré-final mais longo e um contorno nuclear ascendente caracterizando uma “ascensão contínua”, em comparação com *UB*, que foi marcado por alongamento de sílaba pré-final menor e um contorno nuclear decrescente (ou queda final) caracterizando um “declaração neutra” (Tenani, 2002; Frota and Moraes, 2016). Presumivelmente, essas pistas acústicas guiaram os ouvintes na detecção de que *IPB* e *UB* coincidiam com as fronteiras da oração e do enunciado, respectivamente, nas declarações que usamos. Uma vez que estudos anteriores apoiaram a ideia de que o CPS é principalmente elicitado por pistas prosódicas acústicas (Brouwer et al., 2012; Leone-Fernandez et al., 2016a, b), enquanto as pistas linguísticas modulam sua amplitude e topografia do couro cabeludo (Kerkhofs, Vonk, Schriefers, and Dorothee J Chwilla), formulamos a hipótese de que os efeitos CPS encontrados no *IPB* e *UB* podem ter sido elicitados por pistas prosódicas acústicas, enquanto diferenças de informações linguísticas

(isto é, sintáticas e semânticas), caracterizando orações subordinadas e principais, modulavam a latência, amplitude e duração do início. Isso poderia ter sido amplamente comprovado neste estudo se IP e U fossem semanticamente e sintaticamente possíveis e apenas pela prosódia. Notavelmente, embora o *IPB* não tenha sido marcado por uma pausa, foi sinalizado por intervalos de silêncio que não foram manipulados. Sugerimos que a presença dessas pequenas pausas pode ter impactado na modulação do CPS que observamos. Pesquisas futuras, onde controlamos esses intervalos de silêncio, devem determinar seu impacto sobre os efeitos do CPS.

De acordo com Kerkhofs et al. (2008), a amplitude do CPS é modulada em função da saliência dos marcadores de contorno prosódicos. Esta presunção sugere que a “subida contínua” caracterizada pelo tom de fronteira – alongamento pré-final mais longo (na sílaba tônica pré-fronteira), aumento do pitch e intensidade mais alta (na sílaba tônica pré-fronteira e pós-tônica) – marcando o *IPB* desencadeou a maior amplitude do CPS que observamos em comparação com a menor amplitude do CPS para *UB*, que foi marcada pelo tom de fronteira – menos alongamento pré-final (na sílaba tônica pré-fronteira), queda do tom e menor intensidade (na sílaba tônica pré-fronteira e sílaba pós-tônica) – caracterizando um “enunciado neutro”. Portanto, essa diferença na força das pistas prosódicas parece ser o fator mais provável para explicar a diferença de amplitude relativa do CPS existente entre *IPB* e *UB*. Aparentemente, os ouvintes eram sensíveis às diferenças de padrão prosódico entre *IPB* e *UB*.

Uma explicação plausível para a diferença de latência de início pode, também, estar relacionada às diferenças na força das pistas prosódicas. A latência de início precoce do CPS para *IPB* em comparação com *UB* pode refletir um processamento mais rápido para *IPB*. Isso pode ser explicado principalmente pelo fato de a sílaba tônica pré-fronteira – marcando o início da fronteira prosódica – para *IPB*, em comparação com *UB*, ter pistas acústicas mais evidentes incluindo alongamento mais longo, maior F0 média e maior intensidade média. Isso parece ter desencadeado um efeito CPS inicial para *IPB* em comparação com *UB*.

Quanto à diferença de duração, sugerimos que a duração do CPS variou em função de sua modulação em amplitude. Conforme refletido na análise acústica, a manifestação em pistas prosódicas é que *IPB* foi marcado por um alongamento de sílaba pré-final mais longo (em sílaba tônica pré-fronteira) em comparação com *UB*, resultando em uma saliência mais alongada de marcadores prosódicos acústicos relatados para *IPB*. Assim, a diferença de duração do CPS pode ser

atribuída à diferença de alongamento da sílaba pré-final entre *IPB* e *UB*. Levando toda a discussão em consideração, podemos concluir que a amplitude, latência de início e duração dos efeitos CPS refletiram a extensão em que as pistas prosódicas acústicas foram ativadas, de acordo com achados anteriores (Steinhauer & Friederici, 2001; Steinhauer, 2003).

## 5. CONCLUSÃO

Este estudo apresenta evidências para o fato de que, na compreensão da fala, os ouvintes são sensíveis a pistas prosódicas de diferentes níveis constituintes, a saber, *IPB* e *UB* no português brasileiro. Isso foi observado por meio do exame das diferenças prosódicas entre as fronteiras que sinalizam os dois constituintes e da análise das respostas neurofisiológicas ao processamento de pistas que as sinalizam por meio de um experimento que inspeciona o componente de ERP *Closure Positive Shift* (CPS). O estudo contribui para os campos de pesquisa do processamento da fala e compreensão da linguagem falada, fornecendo evidências para a diferença no processamento de frases internas e fronteiras prosódicas de enunciados internos (conforme refletido nos diferentes parâmetros dos componentes do CPS observados nas duas fronteiras distintas). A comparação das duas respostas CPS é uma contribuição relevante para o campo, fornecendo insights sobre os processos neurocognitivos em constituintes frasais hierarquicamente e prosodicamente diferentes.

## REFERÊNCIAS

Araújo, Dráulio B. de, et al. “Localizando a Atividade Cerebral via Magnetoencefalografia”. *Ciência e Cultura*, vol. 56, no. 1, 2004, pp. 38-40.

Barbara Grosz; Julia Hirschberg. “Some Intonational Characteristics Of Discourse Structure”. In *Proceeding of the International Conference on Spoken Language Processing*, 1992, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.37.8065>.

Beckman, Mary, et al. “Prosodic Structure and Tempo in a Sonority Model of Articulatory Dynamics”. *Gesture, Segment, Prosody*, edited by Gerard J. Docherty and D. Robert Ladd, Cambridge University Press, 1992, pp. 68-89, doi:10.1017/CBO9780511519918.004.

Beckman, Mary E., and Jan Edwards. “Lengthenings and Shortenings and the Nature of Prosodic Constituency”. *Papers in Laboratory Phonology*, edited by

John Kingston and Mary E. Beckman, Cambridge University Press, 1990, pp. 152-78, doi:10.1017/CBO9780511627736.009.

Beckman, Mary E., and Janet B. Pierrehumbert. “Intonational Structure in Japanese and English”. *Phonology*, vol. 3, May 1986, p. 255, doi:10.1017/S095267570000066X.

Boersma, P., and D. Weenink. “Praat: Doing Phonetics by Computer (Version 4.5.) [Computer Program]”. Retrieved from <Http://Www.Praat.Org/>, vol. 5, no. 9/10, 2007, pp. 341-45, papers3://pUBlication/uuid/AF582E4D-2F7A-409E-B-4F1-7A10385D9135.

Bögels, Sara, Herbert Schriefers, Wietske Vonk, and Dorothee J. Chwilla. “Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials”. *Linguistics and Language Compass*, vol. 5, no. 7, 2011, pp. 424-40, doi:10.1111/j.1749-818X.2011.00291.x.

Bögels, Sara, Herbert Schriefers, Wietske Vonk, Dorothee J. Chwilla, et al. “The Interplay between Prosody and Syntax in Sentence Processing: The Case of SUBject- and Object-Control Verbs”. *Journal of Cognitive Neuroscience*, vol. 22, no. 5, 2010, pp. 1036-53, doi:10.1162/jocn.2009.21269.

Brouwer, Harm, et al. “Getting Real about Semantic Illusions: Rethinking the Functional Role of the P600 in Language Comprehension”. *Brain Research*, vol. 1.446, 2012, pp. 127-43, doi:10.1016/j.brainres.2012.01.055.

Byrd, Dani, et al. “How Far, How Long: On the Temporal Scope of Prosodic Boundary Effects”. *The Journal of the Acoustical Society of America*, vol. 120, no. 3, Sept. 2006, pp. 1589-99, doi:10.1121/1.2217135.

Cambier-Langeveld, GM. *Temporal Marking of Accents and Boundaries*. Den Haag: Holland Academic Graphics, 2000, <https://dare.uva.nl/search?identifier=e26b40a8-a920-453b-9fc2-589d0656c350>.

Chacon, Lourenço, and Milena Fraga. “Pausas na Interpretação Teatral: Delimitação de Constituintes Prosódicos”. *Filologia e Linguística Portuguesa*, vol. 16, no. 1, July 2014, p. 121, doi:10.11606/issn.2176-9419.v16i1p121-146.

Delorme, Arnaud, and Scott Makeig. “EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis”. *Journal of Neuroscience Methods*, vol. 134, no. 1, 2004, pp. 9-21, doi:10.1016/j.jneumeth.2003.10.009.

Fernandes, Flaviane Romani. *Ordem, Focalização e Preenchimento em Português: Sintaxe e Prosódia*. Universidade Estadual de Campinas, 2007, [http://www.tycho.iel.unicamp.br/~tycho/pesquisa/monografias/FERNANDES\\_FR-Dr.pdf](http://www.tycho.iel.unicamp.br/~tycho/pesquisa/monografias/FERNANDES_FR-Dr.pdf).

Féry, Caroline, et al. “Correlates of Phrasing in French and German from an Experiment with Semi-Spontaneous Speech”. *Intonational Phrasing in Romance and Germanic*, 2011, pp. 11-41, <https://benjamins.com/#catalog/books/hsm.10/main>.

Friederici, Angela D. “The Brain Basis of Language Processing: From Structure to Function”. *Physiological Reviews*, vol. 91, no. 4, Oct. 2011, pp. 1357-92, doi:10.1152/physrev.00006.2011.

Frota, Sónia, et al. *P-ToBI: Tools for the Transcription of Portuguese Prosody*. 2015.

Frota, Sónia, and João Moraes. “Intonation in European and Brazilian Portuguese”. *The Handbook of Portuguese Linguistics*, edited by Sergio Menuzzi W. Leo Wetzels, João Costa, 2016.

Frota, Sónia, and Marina Vigário. “Aspectos de Prosódia Comparada: Ritmo e Entoação no PE e no PB”. *Actas Do XV Encontro da Associação Portuguesa de Linguística*, 2000, pp. 533-55, <http://labfon.letras.ulisboa.pt/SonseMelodias/APLPEPB.pdf>.

Geluykens, Ronald, and Marc Swerts. “Prosodic Cues to Discourse Boundaries in Experimental Dialogues”. *Speech Communication*, vol. 15, no. 1-2, Elsevier Science PUBLishers B. V., Oct. 1994, pp. 69-77, doi:10.1016/0167-6393(94)90042-6.

Gilbert, Ac, et al. “Effects of Temporal Chunking on Speech Recall”. *Speech Prosody*, 2012, pp. 2-5, [www.phonetique.info](http://www.phonetique.info).

Gilbert, Annie C., et al. “Exploring the Rhythmic Segmentation of Heard Speech Using Evoked Potentials”. *Proceedings of the 5th International Conference on Speech Prosody (SP2010)*, no. JANUARY 2010, 2010.

Gilbert, Annie C, et al. “The Perceptual Chunking of Speech: A Demonstration Using ERPs”. *Brain Research*, vol. 1603, Elsevier, 2015, pp. 101-13, doi:10.1016/j.brainres.2015.01.032.

Hirst, Daniel, and Albert Di Cristo. "A Survey of Intonation Systems". *Intonation Systems A Survey of Twenty Languages*, no. January 1998, 1998, pp. 1-44, <http://books.google.com/books?hl=en&lr=&id=LClvNiI4k0sC&pgis=1>.

Itzhak, Inbal, et al. "Event-Related Potentials Show Online Influence of Lexical Biases on Prosodic Processing". *Neuroreport*, vol. 21, no. 1, 2010, pp. 8-13, doi:10.1097/WNR.0b013e328330251d.

Kerkhofs, Roel, Wietske Vonk, Herbert Schriefers, and Dorothee J Chwilla. "Discourse, Syntax, and Prosody: The Brain Reveals an Immediate Interaction". *Journal of Cognitive Neuroscience*, vol. 19, no. 9, 2007, pp. 1.421-34, doi:10.1162/jocn.2007.19.9.1421.

Kerkhofs, Roel, Wietske Vonk, Herbert Schriefers, and Dorothee J. Chwilla. "Sentence Processing in the Visual and Auditory Modality: Do Comma and Prosodic Break Have Parallel Functions?" *Brain Research*, vol. 1224, 2008, pp. 102-18, doi:10.1016/j.brainres.2008.05.034.

Kowal, Sabine, et al. "The Use of Time in Storytelling". *Language and Speech*, vol. 26, no. 4, 1983, pp. 377-92, doi:10.1177/002383098302600405.

Kutas, M., and S. A. Hillyard. "Brain Potentials during Reading Reflect Word Expectancy and Semantic Association". *Nature*, vol. 307, no. 5.947, 1984, pp. 161-63, doi:10.1038/307161a0.

Kutas, M., and S. A. Hillyard. "Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity". *Science (New York, N.Y.)*, vol. 207, no. 4.427, 1980, pp. 203-05, doi:10.1126/science.7350657.

Ladd, D. Robert. "Intonational Phrasing: The Case for Recursive Prosodic Structure". *Phonology*, vol. 3, 1986, p. 311, doi:10.1017/S0952675700000671.

Ladd, D. Robert. "Introduction to Intonational Phonology". *Intonational Phonology*, Cambridge University Press, 2008, pp. 3-42, doi:10.1017/CBO9780511808814.002.

Ladd, D. Robert, and Nick Campbell. "Theories of Prosodic Structure: Evidence from Syllable Duration". *Proceedings of the 12th International Congress of Phonetic Sciences*, vol. 2, 1991, pp. 290-93.

Li, W., and Y. Yang. "Perception of Prosodic Hierarchical Boundaries in Mandarin Chinese Sentences". *Neuroscience*, vol. 158, no. 4, 2009, pp. 1.416-25, doi:10.1016/j.neuroscience.2008.10.065.

Liberman, Mark, and Alan Prince. "On Stress and Linguistic Rhythm". *Linguistic Inquiry*, vol. 8, no. 2, 1977, pp. 249-336, doi:10.2307/4177987.

Lopez-Calderon, Javier, and Steven J. Luck. "ERPLAB: An Open-Source Toolbox for the Analysis of Event-Related Potentials". *Frontiers in Human Neuroscience*, vol. 8, Frontiers, Jan. 2014, p. 213, doi:10.3389/fnhum.2014.00213.

Männel, Claudia, and Angela D. Friederici. "Intonational Phrase Structure Processing at Different Stages of Syntax Acquisition: ERP Studies in 2-, 3-, and 6-Year-Old Children". *Developmental Science*, vol. 14, no. 4, 2011, pp. 786-98, doi:10.1111/j.1467-7687.2010.01025.x.

Männel, Claudia, and Angela D. Friederici. "Pauses and Intonational Phrasing: ERP Studies in 5-Month-Old German Infants and Adults". *Journal of Cognitive Neuroscience*, vol. 21, no. 10, 2009, pp. 1.988-2.006, doi:10.1162/jocn.2009.21221.

Martins, Vanessa de Oliveira, and Claudia Andrade Furquim de Regina. *Perfil Evolutivo da Fluência da Fala de Falantes do Português Brasileiro*. no. 1, 2008, pp. 7-12, doi:10.1590/S0104-56872008000100002.

Moraes, J. A. *Lexical e Acentuação Frasal Em Português. Um Estudo Acústico-Perceptivo*. Estudos Lingüísticos e Literários 17, pp. 39-57, [http://www.scielo.br/scielo.php?script=sci\\_nlinks&ref=000098&pid=S0102-4450199800030000500024&lng=pt](http://www.scielo.br/scielo.php?script=sci_nlinks&ref=000098&pid=S0102-4450199800030000500024&lng=pt).

Nespor, Marina, and Irene Vogel. "Prosodic Phonology :With a New Foreword". *Studies in Generative Grammar*, vol. 28, 2007.

Nespor, Marina, and Irene Vogel. "Prosodic Phonology". *Linguistic Inquiry*, 1986.

Nespor, Marina, and Irene Vogel. *Prosodic Structure Above the Word*. Springer, Berlin, Heidelberg, 1983, pp. 123-40, doi:10.1007/978-3-642-69103-4\_10.

Oldfield, R. C. "The Assessment and Analysis of Handedness: The Edinburgh Inventory". *Neuropsychologia*, vol. 9, no. 1, 1971, pp. 97-113, doi:10.1016/0028-3932(71)90067-4.

Oliveira, Miguel. *Prosodic Features in Spontaneous Narratives*. Simon Frazer University, Vancouver, Canada, 2000.

Pannekamp, Ann, et al. "Prosody-Driven Sentence Processing: An Event-Related Brain Potential Study". *Journal of Cognitive Neuroscience*, vol. 17, no. 3, 2005, pp. 407-21, doi:10.1162/0898929053279450.

Passonneau, Rebecca J., and Diane J. Litman. "Intention-Based Segmentation: Human Reliability and Correlation with Linguistic Cues". *Proceeding of the 31st Annual Meeting of the Association for Computational Linguistics (ACL-93)*, 1993, p. 8, <http://arxiv.org/abs/cmp-lg/9405015>.

Pauker, Efrat, et al. "Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences: ERP Evidence for the Boundary Deletion Hypothesis". *Journal of Cognitive Neuroscience*, vol. 23, no. 10, 2011, pp. 2.731-51, doi:10.1162/jocn.2011.21610.

Peter, Varghese, et al. "Using Event-Related Potentials to Measure Phrase Boundary Perception in English". *BMC Neuroscience*, vol. 15, no. 1, 2014, p. 129, doi:10.1186/s12868-014-0129-z.

Santana, Santos Raquel, and Eneida Goes Leal. "Os Domínios Prosódicos e a Duração de Sílabas no Português Brasileiro". *Estudos da Língua(Gem)*, vol. 8, no. 2, 2011, pp. 133-71.

Schmidt-Nielsen, Astrid. *Intelligibility and Acceptability Testing for Speech Technology*. 1992, <http://www.dtic.mil/dtic/tr/fulltext/u2/a252015.pdf>.

Selkirk, Elisabeth O. "Phonology and Syntax: The Relation between Sound and Structure". *Current Studies in Linguistics Series*, 1984, doi:10.1.1.78.8619.

Selkirk, Elisabeth O. "The Role of Prosodic Categories in English Word Stress". *Linguistic Inquiry*, vol. 11, no. 3, 1980, pp. 563-605.

Serra, Carolina. "A interface prosódia-sintaxe e prosódico no português do brasil. Prosody-Syntax Interface and Prosodic Phrasing in Brazilian Portuguese". *Journal of Speech Sciences*, vol. 5, no. 2, 2016, pp. 47-86.

Serra, Carolina. *Realização e Percepção de Fronteiras Prosódicas no Português do Brasil: Fala Espontânea e Leitura*. Universidade Federal do Rio de Janeiro – UFRJ, 2009.



Serra, Carolina, and Sónia Frota. “Fraseamento Prosódico no Português do Brasil: Pistas para a Percepção”. *ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA, XXV*, [http://labfon.letras.ulisboa.pt/texts/APLSerra\\_Frota 2009.pdf](http://labfon.letras.ulisboa.pt/texts/APLSerra_Frota 2009.pdf).

Soncin, Geovana, et al. *Percepção de Pausa Em Fronteira Prosódica*. 2017, pp. 143-64.

Steinhauer, K., et al. “Brain Potentials Indicate Immediate Use of Prosodic Cues in Natural Speech Processing”. *Nature Neuroscience*, vol. 2, no. 2, 1999, pp. 191-96, doi:10.1038/5757.

Steinhauer, K., and A. D. Friederici. “Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers”. *Journal of Psycholinguistic Research*, vol. 30, no. 3, 2001, pp. 267-95, doi:10.1023/A:1010443001646.

Steinhauer, Karsten. “Electrophysiological Correlates of Prosody and Punctuation”. *Brain and Language*, vol. 86, no. 1, 2003, pp. 142-64, doi:10.1016/S0093-934X(02)00542-4.

Tenani, Luciani Ester. *Domínios Prosódicos no Português do Brasil: Implicações para a Prosódia e para a Aplicação de Processos Fonológicos*. Universidade Estadual de Campinas, 2002.

Turk, Alice E., and Stefanie Shattuck-Hufnagel. “Multiple Targets of Phrase-Final Lengthening in American English Words”. *Journal of Phonetics*, vol. 35, no. 4, Academic Press, Oct. 2007, pp. 445-72, doi:10.1016/j.wocn.2006.12.001.

Turk, Alice E., and Stefanie Shattuck-Hufnagel. “Word-Boundary-Related Duration Patterns in English”. *Journal of Phonetics*, vol. 28, no. 4, Academic Press, Oct. 2000, pp. 397-440, doi:10.1006/jpho.2000.0123.

Venditti, Jennifer J., et al. “Prosodic Cues to Syntactic and Other Linguistic Structures in Japanese, Korean, and English”. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisitions*, edited by J. L. Morgan and K. Demuth, 1996, pp. 287-311.

Wichmann, Anne. “Intonation in Text and Discourse: Beginnings, Middles, and Ends”. *Studies in Language and Linguistics*, 2000.

Wightman, C. W., et al. “Segmental Durations in the Vicinity of Prosodic Phrase Boundaries”. *The Journal of the Acoustical Society of America*, vol. 91, no. 3, 1992, pp. 1.707-17, doi:10.1121/1.402450.

Wolff, Susann, et al. “The Neural Mechanisms of Word Order Processing Revisited: Electrophysiological Evidence from Japanese”. *Brain and Language*, vol. 107, no. 2, 2008, pp. 133-57, doi:10.1016/j.bandl.2008.06.003.

Zhang, Hongming. *Syntax-Phonology Interface: Argumentation from Tone Sandhi in Chinese Dialects*. Routledge, Routledge, an imprint of the Taylor & Francis Group, 2017.



# ASPECTOS PROSÓDICOS DA LIBRAS EM NÚMEROS DE TELEFONE MÓVEL

## PROSODIC ASPECTS OF LIBRAS IN MOBILE PHONE NUMBERS

*Ewerton Carlos Matos Marques<sup>1</sup>*

*Miguel Oliveira Jr.<sup>2</sup>*

*Ayane Nazarela Santos de Almeida<sup>3</sup>*

**RESUMO:** O presente artigo objetiva descrever aspectos prosódicos de duração e pausa na composição da estrutura de números de telefone móvel em Libras (Língua de Sinais Brasileira). Investiga a forma mais recorrente para números de telefone móvel a partir de características prosódicas específicas que se revelam na segmentação dos agrupamentos numéricos. A metodologia utilizada organiza-se em três etapas: (i) coleta de dados por meio de entrevista em Libras através de sinalização espontânea e leitura de agrupamentos numéricos projetados através de *slides*; (ii) segmentação e anotação das entrevistas no *ELAN* (HELLWIG e GEERTS, 2013); (iii) transferência das anotações realizadas no *ELAN*, para arquivo com extensão *textgrid* do *Praat* (BOERSMA e WEENINK, 2017) a fim de automatizar, por meio do *script Analyse tier* (HIRST, 2012), os parâmetros prosódicos necessários para a análise de dados dos números de telefone móvel. Identificamos, tanto nos dados de leitura, quanto nos espontâneos, a sinalização dos números localizados em posição pré-fronteira com uma duração significativamente maior do que os números que não estavam localizados

---

<sup>1</sup> Universidade Federal de Alagoas (UFAL).

<sup>2</sup> Universidade Federal de Alagoas (UFAL).

<sup>3</sup> Universidade Federal do Recôncavo da Bahia (UFRB);

em posição de fronteira, o que indica um alongamento do sinal como pista de marcação da fronteira prosódica. Por outro lado, nossos achados revelam que a pausa não parece ser parâmetro prosódico relevante para marcar fronteira entre unidades prosódicas na sinalização de agrupamentos numéricos de celular.

**Palavras-chave:** Prosódia; Números de celular; Libras.

**ABSTRACT:** The present paper aims to describe prosodic aspects of duration and pause in the composition of the structure of mobile phone numbers in Libras (Brazilian Sign Language). It investigates the most recurrent form for mobile phone numbers from specific prosodic features that are revealed in the segmentation of numeric groupings. The methodology used is organized in stages: (i) data collection through interview in Libras through spontaneous signaling and reading of numerical groupings projected through slides; (ii) segmentation and annotation of interviews at *ELAN* (HELLWIG e GEERTS, 2013); (iii) transfer of the annotations made in *ELAN*, to a file with Praat textgrid extension (BOERSMA e WEENINK, 2017) in order to automate, through the Analyze tier script (HIRST, 2012), the prosodic parameters required for data analysis of mobile phone numbers. We identified both in the reading data and in the personal data, the signaling of numbers located in a pre-boundary position with a significantly longer duration than the numbers that were not located in a boundary position, which indicates an lengthening of the signal as a cue of prosodic boundary marking. On the other hand, our findings reveal that the pause does not seem to be a relevant prosodic parameter to mark the boundary between prosodic units in signaling mobile phone numerical groupings.

**Keywords:** Prosody; Mobile phone numbers; Libras.

## 1. INTRODUÇÃO

Em Libras, como em qualquer outra língua, oral ou de sinais, existem estruturas linguísticas regulares ou fixas que constituem expressões cujo significado depende da observação de características segmentais e suprasegmentais de seus constituintes. Os agrupamentos numéricos, utilizados para nomear entidades, distinguindo-as entre si (WIESE, 2003), são exemplos dessas estruturas regulares, tais como números telefônicos, Cadastro de Pessoa Física (CPF) e Código de Endereçamento Postal (CEP).

Avanços nessa área têm sido obtidos em decorrência da descrição da estrutura prosódica de números naturais nas mais diversas línguas orais, tais como

o inglês (RAHIM et al., 2001), o alemão (BAUMANN e TROUVAIN, 2001), o espanhol (LOPEZ et al., 1998), o francês (MOKBEL et al., 1997), o japonês (AMINO e OSANAI, 2011), o húngaro (OLASZI, 2000), o chinês (WANG e SENEFF, 1998) e o português europeu (RODRIGUES, 2001). Até o momento, poucos são os trabalhos sobre o português do Brasil que descrevem, de forma sistemática e abrangente, as várias características acústicas da organização de números naturais em estruturas pré-estabelecidas (ALMEIDA, 2017; MUSILYU, 2014). Nas línguas de sinais, tanto quanto sabemos, há apenas uma pesquisa, em Libras (MARQUES, 2020), que descreveu aspectos prosódicos na sinalização de números naturais, tais como números de telefone móvel, como nos propomos a fazer na presente investigação.

Estudos sobre a prosódia em línguas de sinais, como a Libras (LEITE, 2008; SANTOS, 2018; CASTRO, 2019), apontam que, pelo fato de estas se constituírem como línguas de modalidade visual-espacial, apresentam sistemas prosódicos que devem ser observados a partir dos movimentos corporais, das mãos, dos braços, tronco, cabeça e das expressões faciais. Outras pesquisas (SANDLER, 2012; PFAU e QUER, 2010; HANSEN e HESSMANN, 2008; DENMARK; CAMPBELL e WOLL, 2008; NICODEMUS e SMITH, 2006) ratificam que meios não manuais como as expressões dos olhos, das sobrancelhas, da boca, da cabeça e da pálpebra também desempenham funções prosódicas, sobretudo no que se refere à entoação e à intensidade, em línguas de sinais. Portanto, a expressão facial e alguns outros marcadores, como a ausência de expressões, têm importantes funções para as línguas de sinais, assim como a entoação para as línguas orais.

No entanto, a presente pesquisa não investiga aspectos prosódicos de entoação e intensidade, mas aspectos duracionais como a pausa e a duração. Assim, apresentamos uma descrição de aspectos prosódicos de duração e pausa na composição da estrutura de números de telefone móvel em Libras. Embasados na literatura prévia para sentenças não numéricas nas Línguas de Sinais, Brasileira (LEITE, 2008; CASTRO, 2019), Americana (COULTER, 1993; NICODEMUS e SMITH, 2006; SANDLER, 2012), Alemã (HANSEN e HESSMANN, 2008), Britânica e Sueca (DENMARK, CAMPBELL e WOLL, 2008), partimos das hipóteses de que a pausa e o alongamento são pistas produtivas na marcação de fronteiras prosódicas na sinalização de agrupamentos numéricos de telefones móveis em Libras.

O presente estudo visa a contribuir nesse sentido, apresentando uma descrição da prosódia de agrupamentos numéricos em uma estrutura fixa (números

de telefone celular), tal como sinalizada espontaneamente e a partir de estímulos de leitura. Ademais, na maioria das línguas, o sistema numérico tem grande importância no funcionamento da vida prática, na nossa interação cotidiana. A despeito disso, os estudos em linguística que analisam esses elementos ainda são restritos. Analisar como tal sistema se estrutura em Libras (Língua de Sinais Brasileira) é observar o conjunto específico de características que organizam seu ordenamento, contribuindo para a expansão dos estudos linguísticos a respeito de uma língua ainda pouco explorada cientificamente, sobretudo no que se refere aos seus aspectos prosódicos.

## 2. PROCEDIMENTOS METODOLÓGICOS

Os dados utilizados no presente estudo são resultantes de uma coleta realizada em vídeo e obtidos através de uma pequena entrevista durante a qual os participantes foram solicitados a sinalizar suas informações pessoais, tais como sua cor favorita, sua preferência no esporte, local de residência, os números de seu telefone celular, do Cadastro de Pessoa Física (CPF) e do Código de endereçamento postal (CEP) do seu local de residência. Apesar do objetivo dessa pesquisa ser a análise dos números de celular, solicitamos também CPF e CEP apenas como elementos distratores.

Para essa tarefa, os participantes não consultaram seus dados pessoais, eles sinalizaram os números da forma como haviam memorizado. Por essa razão, consideramos que os dados pessoais foram sinalizados pelos participantes de forma natural, espontânea.

Posteriormente, foi solicitado aos informantes que sinalizassem os agrupamentos numéricos exibidos em *slides*, que continham dois estímulos de cada um dos três agrupamentos, a saber: número de celular, CPF e CEP. Os voluntários foram instruídos a ler e sinalizar os números exibidos através dos *slides*. A duração das etapas de entrevista e leitura de slides foi de aproximadamente 7 (sete) minutos para cada informante.

Participaram da entrevista dezoito informantes surdos, estudantes universitários, alagoanos e sinalizantes fluentes da Libras. A média de idade dos participantes era de 28 anos, sendo nove do sexo feminino e nove do sexo masculino. Todos se apresentaram como sinalizantes nativos da Libras, embora não tenham nascido em famílias de pais surdos.

As gravações dos dados foram realizadas no estúdio do Laboratório de Gravação do curso de Letras-Libras da Universidade Federal de Alagoas – UFAL,

fazendo uso de câmeras, iluminação artificial e parede azul como tela de fundo, recurso utilizado para atender às normas técnicas apropriadas para gravação de vídeos.

Após gerar os arquivos de vídeo de todos os dados, realizamos a segmentação e a anotação dos mesmos, utilizando o *ELAN* (HELLWIG e GEERTS, 2013). Para Reed (2011), o primeiro passo, em qualquer análise prosódica, é a identificação das unidades prosódicas. Então, a segmentação e anotação dos agrupamentos numéricos em unidades menores foram feitas a partir da identificação dessas unidades.

Após realizadas as anotações, exportamos as anotações do *ELAN*, como arquivo com extensão *textgrid* do *Praat* (BOERSMA e WEENINK, 2017) para que fosse possível automatizar, por meio do *script Analyse tier* (HIRST, 2012), os valores de duração necessários para análise dos dados. Todos os testes estatísticos foram realizados por meio do *software* estatístico *R* (R Core Team, 2016).

Diante dos dados devidamente anotados, a primeira etapa das análises dos agrupamentos consistiu em verificar a distribuição de frequências, numérica e decimal, dos agrupamentos numéricos sinalizados. Inicialmente, verificamos a preferência dos participantes por sinalizarem os números utilizando a forma decimal em unidade (U), e não em dezena (D), centena (C) ou milhar (M).

Então, procedemos à verificação com relação à distribuição numérica utilizada para sinalizar cada agrupamento, ou seja, como os participantes estruturaram o agrupamento numérico de celular em unidades prosódicas. Por exemplo, uma distribuição numérica 3-2-2-2 significa uma estruturação em uma unidade prosódica ternária (composta por três números) seguida de três unidades prosódicas binárias (composta por dois números).

Nesse caso, por se tratar de variáveis categóricas (número de ocorrências), realizamos o teste estatístico do Qui-quadrado ( $\chi^2$ ) de Pearson a fim de mostrar se as distribuições mais recorrentes apresentaram diferença estatisticamente significativa entre elas. Quando isso ocorreu ( $p < 0,05$ ) entre as duas distribuições mais recorrentes, não consideramos necessário realizar o teste inserindo as demais categorias uma vez que, ao inserir valores de ocorrências de distribuições menos frequentes do que as que já estão presentes no teste, encontramos um valor  $p$  cada vez menor, o que continuou indicando diferença significativa entre as distribuições, seguindo metodologia semelhante a Almeida (2017). Ou seja, a distribuição mais recorrente corresponderia a uma forma representativa de sinalização daquele determinado número analisado.



Por outro lado, ao encontrar um valor  $p > 0,05$  entre as duas maiores distribuições significaria que ambas podem ser consideradas formas de sinalização representativas do agrupamento analisado. Nesse caso, inserimos a terceira distribuição mais recorrente e rodamos o teste  $\chi^2$  com as três maiores frequências para verificar se a terceira forma de distribuição também seria representativa como forma de sinalizar o agrupamento numérico.

Os parâmetros duracionais analisados foram referentes à pausa e à duração. Verificamos a presença ou ausência de pausas entre as unidades prosódicas dos agrupamentos numéricos sinalizados. Consideramos como pausa a suspensão do movimento do sinal, conforme Leite (2008). O parâmetro de duração foi analisado não apenas entre os agrupamentos numéricos como um todo, comparando os dados lidos com os espontâneos, mas também observamos valores de duração entre os sinais de números não localizados em posição de fronteira (SNF), os sinais de números localizados em posição pré-fronteira (SFU) e os sinais de números localizados em final de agrupamento (SFA).

Tais estratégias possibilitaram verificar se os dados pessoais, por serem já conhecidos pelo informante e, portanto, com maior frequência de uso, seriam sinalizados de forma mais rápida do que os dados lidos, conforme já verificado em línguas de sinais (BÖRSTELL, HÖRBERG e ÖSTLING, 2016), bem como se também há alongamento do sinal em fronteiras de agrupamentos numéricos, assim como acontece em fronteiras de sentenças não numéricas sinalizadas, como já apresentado por Leite (2008) e Castro (2019) para a Libras.

Com relação aos parâmetros duracionais analisados, todos os valores correspondentes foram extraídos, de forma semiautomática, a partir do *script AnalyseTier* (HIRST, 2012). Portanto, foi necessário exportar a transcrição feita no ELAN como arquivo *textgrid* do Praat e converter o arquivo de vídeo para o formato *wav*, preservando os tempos, de modo a não haver distorções. Esses procedimentos foram necessários, pois o *script* só roda se houver arquivos de áudio e *textgrid*. O fato de o arquivo de áudio apresentar apenas ruídos não foi um problema para a análise, já que estamos investigando parâmetros duracionais.

Por fim, ao considerarmos que se trata de variáveis quantitativas, utilizamos o teste estatístico de análise da variância (ANOVA) de Fisher para verificar se houve diferença significativa entre os níveis de análise. Entretanto, quando analisamos variáveis com mais de dois níveis utilizamos, após o teste ANOVA, o teste estatístico de comparação múltipla de Tukey a fim de comparar os níveis dois a dois e identificar entre quais níveis de análise reside a diferença constatada no teste ANOVA.

Por exemplo, ao analisarmos o alongamento do sinal em sua posição não final (SNF), em final de unidade prosódica (SFU) e em final de agrupamento (SFA), tivemos três níveis de análise e, portanto, o teste de comparação múltipla de Tukey foi utilizado para verificar entre quais níveis a diferença se mostrou estatisticamente significativa.

### 3. RESULTADOS E DISCUSSÃO

Os dados dos agrupamentos numéricos de celular, que compõem o *corpus* deste estudo, foram analisados no que se refere à distribuição numérica e decimal, e também quanto aos parâmetros prosódicos duracionais de pausa e duração nos sinais realizados de forma lida e espontânea.

Participaram da presente pesquisa dezoito informantes, como já mencionado, (nove mulheres e nove homens), dos quais dois não forneceram os dados lidos e verificamos erros de leitura em ambos os estímulos lidos por um desses participantes. Por essa razão, analisamos 18 arquivos de vídeo gerados a partir da sinalização espontânea dos números de celular dos participantes e 30 arquivos resultantes das leituras, de 15 participantes, dos dois estímulos apresentados.

#### 3.1 Distribuição de frequência numérica e decimal

A partir desses dados, realizamos o levantamento das estratégias de leitura e sinalização espontânea utilizadas pelos participantes, tanto no que se refere à distribuição numérica quanto à distribuição decimal, durante a sinalização dos números de celular.

Com relação à distribuição decimal, identificamos que os participantes sinalizaram os números em unidades, em ambas as modalidades, lida e espontânea. Então, procedemos à análise da distribuição de frequência numérica.

##### 3.1.1 Frequência da distribuição das sequências numéricas dos dados lidos

A maioria dos participantes efetuou a leitura dos números de celular, através da distribuição numérica 2-3-2-2, 33,33%, estruturando o agrupamento em quatro unidades prosódicas, sendo a primeira binária, a segunda ternária, a terceira e última binárias.

Assim, os estímulos foram agrupados majoritariamente da seguinte forma: 99-891-11-96 ou 98-201-37-77. Dentre outras formas apresentadas por um total

de 30 ocorrências, essa representou a mais frequente, com 10 ocorrências, conforme apresentado na Tabela 1.

Tabela 1 – Distribuição de frequência numérica de números de celular sinalizados durante a leitura dos estímulos

Distribuição numérica	Ocorrências	%
2 3 2 2	10	33,3
2 2 2 3	7	23,3
1 2 2 2 2	7	23,3
2 3 1 3	2	6,7
1 1 3 1 3	2	6,7
3 2 2 2	1	3,3
2 2 3 2	1	3,3
Total	30	100

Fonte: Dados da pesquisa.

Realizamos o teste estatístico do Qui-quadrado ( $\chi^2$ ) a fim de verificar se a distribuição numérica mais recorrente (2-3-2-2) apresentou frequência significativamente diferente das segundas mais recorrentes (2-2-2-3 e 1-2-2-2-2). O teste mostrou não haver diferença estatisticamente significativa ( $\chi^2 = 0,75$ ,  $p > 0,05$ ,  $df = 2$ ), ou seja, as três formas de distribuição numérica se mostraram produtivas para a sinalização dos estímulos lidos. Ao inserir as terceiras mais frequentes (2-3-1-3 e 1-1-3-1-3), ainda não identificamos diferença significativa entre as estratégias de distribuição numérica na sinalização dos dados lidos ( $\chi^2 = 8,78$ ,  $p > 0,05$ ,  $df = 4$ ).

Tais resultados já nos revelam que parece não ter havido um padrão de estratégia de leitura no que se refere à sinalização apresentada pelos participantes dos números de celular apresentados como estímulos durante a coleta de dados. A seguir, apresentamos os resultados para os dados sinalizados de forma espontânea e realizamos algumas comparações entre eles.

### 3.1.2 Frequência da distribuição das sequências numéricas dos dados espontâneos

Nos dados lidos, a distribuição mais frequente apresentou a primeira unidade binária e a segunda ternária, enquanto nos dados espontâneos, aconteceu o inverso. Observamos que a distribuição mais recorrente apresentou a primeira

unidade prosódica ternária e a segunda binária. A terceira e a quarta unidades prosódicas foram binárias tanto nos dados lidos quanto nos espontâneos.

Conforme demonstrado na Tabela 2, dentre as distribuições apresentadas, que totalizaram 18 ocorrências, a distribuição numérica 3-2-2-2 foi a mais recorrente nos dados espontâneos (50%), estruturando o agrupamento em quatro unidades prosódicas, assim como aconteceu nos dados lidos. Entretanto, a distribuição foi diferente no que se refere às duas primeiras unidades prosódicas.

Tabela 2 – Distribuição de frequência numérica de números de celular sinalizados espontaneamente

Distribuição numérica	Ocorrências	%
3 2 2 2	9	50,0
1 2 2 2 2	4	22,2
2 3 2 2	3	16,7
2 2 2 3	2	11,1
Total	18	100

Fonte: Dados da pesquisa.

O teste estatístico do Qui-quadrado ( $\chi^2$ ) revelou que não houve diferença significativa ( $\chi^2 = 1,92$ ,  $p > 0,05$ ,  $df = 1$ ) entre as estratégias de distribuição numérica mais recorrentes (3-2-2-2 e 1-2-2-2-2). A terceira forma de organização numérica mais frequente (2-3-2-2) também não apresentou diferença significativa em relação às duas mais recorrentes ( $\chi^2 = 3,87$ ,  $p > 0,05$ ,  $df = 2$ ). A distribuição 2-2-2-3, quarta mais recorrente, também poderia representar uma forma de organização numérica de sinalização do número de celular, de acordo com o resultado do teste estatístico ( $\chi^2 = 6,44$ ,  $p > 0,05$ ,  $df = 3$ ). Esses resultados estatísticos nos mostram que, assim como ocorreu com os dados lidos, parece não haver em nosso *corpus* um padrão de organização numérica na sinalização espontânea de números de celular. Entretanto, nossos achados indicam uma tendência de distribuição numérica para a forma 3-2-2-2 nos dados sinalizados espontaneamente, uma vez que tal distribuição representou mais que o dobro do número de ocorrências da segunda distribuição.

Quando comparamos esses achados com os resultantes da leitura de números de celular, identificamos que a forma de distribuição decimal mais recorrente, em ambas as formas de sinalização, foi organizada em unidades. No entanto, a forma como essas unidades foram distribuídas em unidades prosódicas diferiu muito, de modo a dividir preferências com relação à organiza-

ção da distribuição numérica. Por exemplo, a forma de distribuição numérica mais frequente dos informantes ao sinalizarem seus dados pessoais (3-2-2-2) foi sinalizada apenas por 3,3% durante a leitura, o que poderia nos levar a deduzir que parece ser muito mais uma estratégia de sinalização espontânea do que de leitura, embora devamos salientar que tais distribuições numéricas não apresentaram diferença significativa de frequência em relação às demais realizações apresentadas pelos participantes.

## 3.2 Parâmetros prosódicos duracionais

Além da distribuição numérica e decimal, observamos também os parâmetros prosódicos duracionais de pausa e duração dos sinais realizados de forma lida e espontânea durante a sinalização de números de celular.

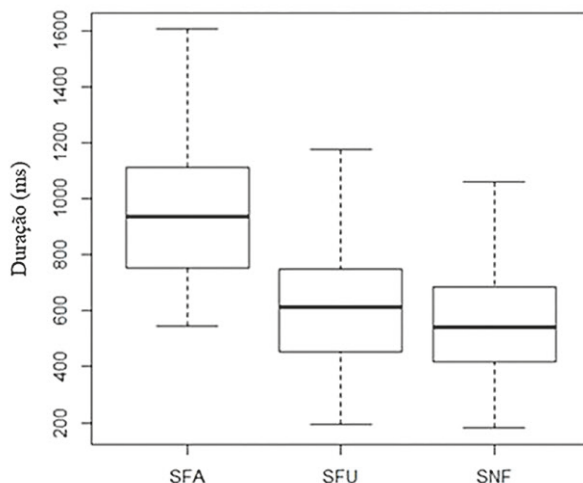
Nos dados espontâneos, não identificamos nenhuma pausa, embora houvesse 58 posições de fronteira entre unidades prosódicas em que a pausa poderia aparecer. Nos dados lidos, havia 99 possibilidades de ocorrência de pausa e identificamos apenas duas pausas no *corpus*, o que representa aproximadamente 2,02% do total de fronteiras prosódicas onde as pausas seriam produtivas. Portanto, a pausa não parece ser pista prosódica utilizada para marcar fronteiras em agrupamentos numéricos de celular sinalizados de forma lida nem espontânea.

Em seguida, analisamos: i) a duração dos agrupamentos numéricos lidos e espontâneos a fim de verificar se os dados pessoais foram sinalizados em uma duração menor do que os dados lidos, conforme pressupomos na metodologia e ii) a duração dos sinais de números não localizados em posição de fronteira (SNF), de números em posição de fronteira (SFU) e de números em posição de final do agrupamento numérico como um todo (SFA) a fim de verificar se, quando em posição de fronteira, há uma marcação com o alongamento do sinal.

### 3.2.1 Duração da sinalização de números nos dados lidos

No que se refere à duração, medida em milissegundos, o Gráfico 1 demonstra que as sinalizações dos números não localizados em posição de fronteira (SNF) apresentaram menor duração do que as dos números que ocupavam a posição final de fronteira (SFU), e que as sinalizações mais longas foram as dos números localizados em final do agrupamento numérico (SFA).

Gráfico 1 – Boxplot da relação entre a posição do número sinalizado e valores de duração nos dados lidos.



SFA: sinal em fim de agrupamento SFU: sinal pré-fronteira SNF: sinal não fronteiro

Fonte: Dados da pesquisa.

A ordem decrescente de duração entre as três condições na leitura do celular foi: SFA > SFU > SNF. Realizamos testes estatísticos para verificar se essa ordem foi estatisticamente significativa.

O teste estatístico ANOVA mostrou que os valores de duração dos sinais de números que estão sendo comparados apresentaram diferença estatisticamente significativa,  $F(2, 246) = 37,63$ ,  $p < 0,05$ . Após o teste ANOVA, realizamos o teste Tukey para comparar os níveis das condições analisadas, dois a dois, conforme Tabela 3.

Tabela 3 – Valores de  $p$ , relativos à duração, resultantes do teste Tukey comparando os níveis das condições analisadas nos dados lidos

Níveis das condições	Valor $p$
SFU – SFA	< 0,0001
SNF – SFA	< 0,0001
SNF – SFU	0,0353303

Fonte: Dados da pesquisa.

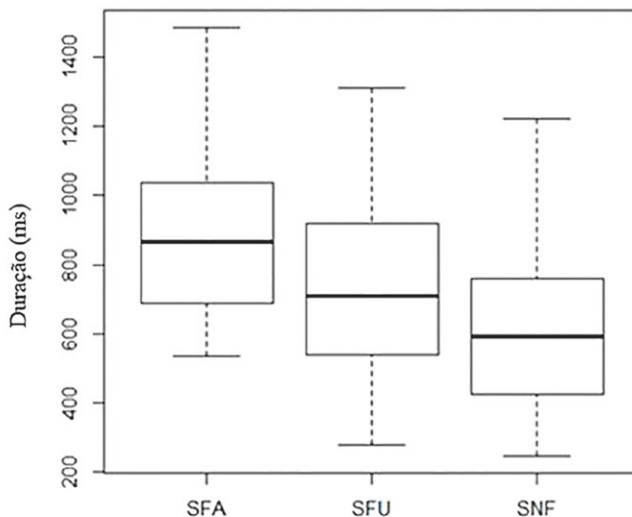
Os resultados do teste Tukey revelaram que, no que se refere à duração, houve diferença estatisticamente significativa entre os níveis das condições

analisadas, ainda quando comparados dois a dois. Então, a duração dos sinais realizados para números em fim de agrupamento numérico foi significativamente maior do que a duração dos sinais de números em posição de fronteira prosódica que é significativamente maior que a duração dos sinais de números em posição não fronteira. Portanto, nossos achados referentes aos dados lidos revelam que o alongamento do sinal foi utilizado pelos Surdos para marcar fronteiras de unidades prosódicas, conforme esperado. Verificamos também se o mesmo ocorreu nos dados pessoais.

### 3.2.2 Duração da sinalização de números nos dados espontâneos

A duração dos sinais em agrupamentos numéricos de celular sinalizados espontaneamente apresentou comportamento similar ao demonstrado nos dados de leitura, abordados na subseção anterior,  $SFA > SFU > SNF$ , conforme demonstrado no Gráfico 2.

Gráfico 2 – Boxplot da relação entre a posição do número sinalizado e valores de duração nos dados espontâneos



SFA: sinal em fim de agrupamento SFU: sinal pré-fronteira SNF: sinal não fronteira

Fonte: Dados da pesquisa.

Os sinais dos números que não estavam em posição de fronteira (SNF) apresentaram duração menor do que os sinais dos números localizados em posição

pré-fronteira (SFU), enquanto os números dispostos no final do agrupamento numérico (SFA) foram sinalizados com a maior duração.

A diferença entre os valores de duração, considerando todos os níveis das condições analisadas, se mostrou estatisticamente significativa,  $F(2, 152) = 9,51$ ,  $p < 0,05$ . Após a realização do teste ANOVA, realizamos o teste Tukey e observamos os resultados apresentados na Tabela 4.

Tabela 4 – Valores de  $p$ , relativos à duração, resultantes do teste Tukey comparando os níveis das condições analisadas nos dados espontâneos

Níveis das condições	Valor $p$
SFU – SFA	0,0707525
SNF – SFA	0,0002754
SNF – SFU	0,0240069

Fonte: Dados da pesquisa.

De acordo com os resultados apresentados na Tabela 4, não houve diferença estatisticamente significativa apenas entre as durações dos sinais dos números em fim do agrupamento (SFA) e dos sinais dos números em posição de final de cada unidade prosódica (SFU). Nesse sentido, identificamos uma pequena diferença entre os dados espontâneos e os lidos, uma vez que estes apresentaram diferença significativa entre os valores de duração de todos os níveis de análise, ainda quando comparados dois a dois. A sinalização dos números que não ocupam posição de fronteira foi a que apresentou menor duração, de forma estatisticamente relevante, assim como ocorreu nos dados numéricos lidos de celular.

Nossos achados, com dados numéricos de celular, corroboram com a literatura prévia a respeito do alongamento do sinal em Libras para marcação de fronteira prosódica em sentenças não numéricas (LEITE, 2008; CASTRO, 2019), uma vez que constatamos, em nossos dados, maior duração dos sinais em posição pré-fronteira (SFU) do que dos sinais que não ocupam posições de fronteira prosódica (SNF).

Leite (2008) aponta para uma universalidade do alongamento final, fato destacado em pesquisas prévias, em outras línguas de sinais como a americana (COULTER, 1993; NICODEMUS e SMITH, 2006; SANDLER, 2012), a britânica, a sueca (DENMARK, CAMPBELL e WOLL, 2008) e a alemã (HANSEN e HESSMANN, 2008). Tais estudos já registraram tal recorrência para sentenças



não numéricas, o que se observa também nas línguas orais, tanto em sentenças não numéricas (SWERTS, 1997; OLIVEIRA JR., 2000; REED, 2011; KENT e READ, 2015), quanto numéricas (ALMEIDA, 2017).

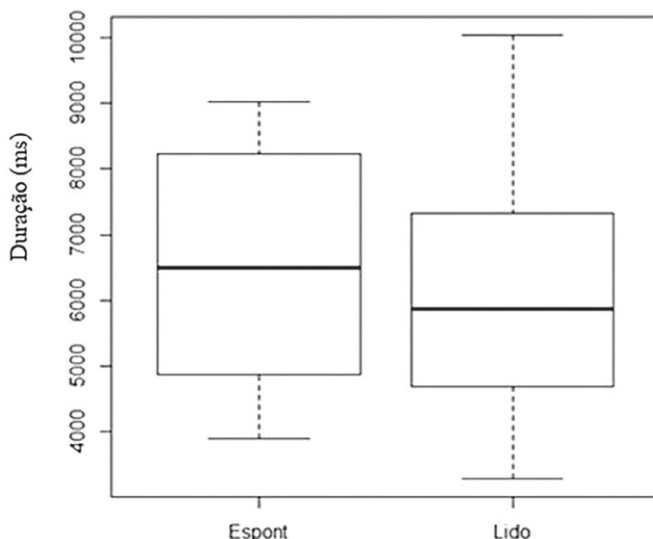
No entanto, ao citar autores como Myers e Hansen (2007), que questionam a motivação para esse alongamento final, Leite (2008) apresenta como hipótese uma antecipação da interrupção do movimento decorrente de uma pausa subsequente (KLATT, 1976; EDWARDS, 1991). Destacamos, porém, que nos dados analisados no presente estudo não identificamos relevância da pausa na marcação de fronteira prosódica, pois na maioria dos casos não houve o registro de pausas que justificasse uma relação entre o alongamento e ocorrências de pausas posteriores, o que confirma o alongamento como uma pista prosódica de marcação de fronteira nos agrupamentos numéricos analisados.

Ao apontar o padrão aceleração-desaceleração das unidades prosódicas como uma relação de duração caracterizada por reduções no início da unidade e alongamentos no final que delimitam cada unidade, Leite (2008) descreve um critério de delimitação de unidades prosódicas que foi recorrente também nos dados dessa pesquisa. Afinal, o alongamento final foi um resultado significativo identificado em quase todos os agrupamentos analisados. Então, em Libras, os sinais realizados em posição final tendem a um alongamento, apresentando uma duração maior, tanto em sentenças não numéricas quanto numéricas.

### **3.2.3 Duração dos agrupamentos numéricos lidos e espontâneos**

Procedemos à análise do parâmetro de duração, em milissegundos, dos agrupamentos numéricos sinalizados, comparando ambas as modalidades de sinalização, lida e espontânea. Conforme já mencionamos, dois participantes não efetuaram a leitura dos estímulos e os dados de leitura de um terceiro participante foram descartados, pois este cometeu equívocos durante a leitura. Portanto, analisamos a duração de 18 agrupamentos numéricos sinalizados de forma espontânea e 30 agrupamentos numéricos sinalizados de forma lida. O Gráfico 3 mostra que os dados lidos foram sinalizados com uma duração menor do que os dados pessoais (espontâneos).

Gráfico 3 – Boxplot da relação entre a modalidade de sinalização e valores de duração dos agrupamentos numéricos



Fonte: Dados da pesquisa.

Embora o gráfico *boxplot* demonstre que os dados pessoais tenham sido sinalizados de forma mais lenta do que os dados lidos, o resultado do teste estatístico ANOVA mostrou que a diferença entre os valores de duração dos agrupamentos numéricos sinalizados não é estatisticamente significativa,  $F(1, 46) = 0,85$ ,  $p > 0,05$ .

Embasados em pesquisas prévias com sentenças não numéricas, em línguas de sinais (BÖRSTELL, HÖRBERG e ÖSTLING, 2016) e em línguas orais (WARNER, 2012), partimos do pressuposto de que os dados espontâneos poderiam ser sinalizados de forma mais rápida do que os dados lidos porque naqueles já há um conhecimento prévio e memorização dos números que sinalizam, pois se trata de seus números pessoais, mas nossos achados revelaram não haver diferença significativa quanto a este parâmetro analisado.

Considerando a quantidade restrita de dados, não podemos generalizar o resultado. Além disso, há a possibilidade de os participantes da pesquisa terem tido dificuldades em fornecer seus dados pessoais por não fazerem isso de forma recorrente, uma vez que, geralmente, tais números são fornecidos por mensagem escrita e não por sinalização. Portanto, ao serem solicitados a sinalizar, pode ter demandado maior esforço cognitivo para o fornecimento desses agrupamentos.

## 4 CONCLUSÕES

Este trabalho objetivou analisar características prosódicas duracionais de agrupamentos numéricos sinalizados em Libras e pretendeu contribuir para a área da prosódia. Os resultados da presente pesquisa mostram que, em Libras, a sinalização dos números de celular parece desenvolver-se através de uma organização decimal em unidades e subdividindo o agrupamento em unidades prosódicas binárias (com dois números) ou até ternárias (com três números), mas não observamos um padrão no ordenamento da organização numérica dessas unidades prosódicas.

Conforme verificamos, nossos dados não revelaram um padrão de distribuição de frequência numérica, nem nos dados lidos, nem nos dados espontâneos, uma vez que não identificamos diferenças estatisticamente significativas entre as frequências de distribuição numérica apresentadas pelos participantes da pesquisa. Talvez um *corpus* mais robusto apresente alguma diferença de modo a mostrar qual seria o padrão de distribuição numérica utilizado pelos Surdos na sinalização de números de celular.

Identificamos, não apenas nos dados de leitura, mas também nos dados pessoais, que os participantes sinalizaram os números localizados em posição pré-fronteira com uma duração significativamente maior do que os números que não estavam localizados em posição de fronteira, o que indica um alongamento do sinal como pista de marcação da fronteira prosódica, conforme havíamos previsto com base na literatura prévia para sentenças não numéricas em línguas de sinais (LEITE, 2008; CASTRO, 2019; SANDLER, 2012; NICODEMUS e SMITH, 2006; COULTER, 1993; HANSEN e HESSMANN, 2008; DENMARK, CAMPBELL e WOLL, 2008). Por outro lado, nossos achados divergem da literatura ao revelarem que a pausa não parece ser parâmetro prosódico relevante para marcar fronteira entre unidades prosódicas na sinalização de agrupamentos numéricos de celular, em Libras.

No que se refere à duração dos agrupamentos numéricos como um todo, ao comparar dados lidos e espontâneos, não constatamos diferença estatisticamente significativa entre tais valores de duração. Ou seja, embora esperássemos que os dados espontâneos fossem sinalizados de forma mais rápida pelos informantes do que os dados lidos, devido a uma maior frequência de uso (BÖRSTELL, HÖRBERG e ÖSTLING, 2016), o que aconteceu foi exatamente o oposto, mas a diferença não se mostrou relevante.

O estudo ampliou o nosso olhar sobre como a estrutura prosódica é empregada em uma língua sinalizada, pois a análise demonstrou que os Surdos fazem

uso de marcadores prosódicos para segmentar as estruturas numéricas fixas como o número de telefone. Observamos que o fato de sinalizarem a maioria dos dados através de unidades pode ser indicativo de uma utilização restrita das representações de dezenas ou centenas para esse fim. Para suposições como essa, precisaríamos de análises mais aprofundadas sobre a produção de dezenas em Libras.

Embora tenhamos a pretensão, com o presente estudo, de minimizar a carência de pesquisas a respeito de características prosódicas dos sinais de agrupamentos numéricos, assim como qualquer investigação, apresenta limitações. Uma das limitações da pesquisa é a quantidade reduzida de informantes, que não nos propiciou a possibilidade de fazer generalizações mais precisas acerca dos nossos achados. Por essa razão, pretendemos ampliar o nosso *corpus*, coletando dados de mais informantes para tentar alcançar resultados que possam nos mostrar efetivamente se há um padrão na distribuição numérica da sinalização de números de celular.

Outro estudo interessante seria a análise de outros parâmetros prosódicos, além dos duracionais, tais como a intensidade e as características entoacionais da Libras na sinalização de agrupamentos numéricos. Dessa forma, seria possível comparar os resultados provenientes da sinalização de agrupamentos numéricos com os achados de pesquisas prévias dessa natureza com sentenças não numéricas.

Esperamos que nossa pesquisa seja uma contribuição para a descrição prosódica da Libras, sobretudo no que se refere à sinalização de agrupamentos numéricos, e que manifeste interesses de novos pesquisadores sobre essa temática a fim de tornar possível a comparação entre dados de diversas comunidades linguísticas surdas do país.

## REFERÊNCIAS

ALMEIDA, Ayane N. S. de. **Análise prosódica de agrupamentos numéricos no português do Brasil**. Tese de Doutorado, Universidade Federal de Alagoas, Maceió, 2017.

AMINO, Kanae; OSANAI, Takashi. Realisation of the prosodic structure of spoken telephone numbers by native and non-native speakers of Japanese. **Proceedings of the 17th International Congress of Phonetic Sciences**. Honk Kong, China, ICPhS, p. 236-239, 2011.

BAUMANN, Stefan; TROUVAIN, Jürgen. On the prosody of German telephone numbers. **Proceedings of the 2th INTERSPEECH and 7th European**

**Conference on Speech Communication and Technology (EUROSPEECH 2001)**, Aalborg, Denmark, p. 557-560, 2001.

BOERSMA, P.; WEENINK, D. **Praat**: doing phonetics by computer (Versão 6.0.36) [Computer program]. Disponível em: <http://www.praat.org/>. Acesso em: 20 out. 2017.

BÖRSTELL, Carl; HÖRBERG, Thomas; ÖSTLING, Robert. Distribution and duration of signs and parts of speech in Swedish Sign Language. **Sign Language and Linguistics**, v. 19, n. 2, p. 143-196, 2016.

CASTRO, Nelson Pimenta. **Prosódia em ASL e Libras**: análise comparativa de aspectos visuais. Tese de Doutorado, Universidade Federal de Santa Catarina, Florianópolis, 2019.

COULTER, Geoffrey R. Phrase-level prosody in ASL: final lengthening and phrasal contours. In: Geoffrey R. Coulter (ed.). **Phonetics and phonology**. Current issues in ASL phonology, v. 3. San Diego: Academic Press, p. 263-272, 1993.

DENMARK, Tanya; CAMPBELL, Ruth; WOLL, Bencie. Seeing sentence boundaries. **Sing Language & Linguistics**, v. 10, n. 2, p. 177-200, 2007.

EDWARDS, Derek. Categories are for talking: On the cognitive and discursive bases of categorization. **Theory & Psychology**, v. 1, n. 4, p. 515-542, 1991.

HANSEN, Martje; HESSMANN, Jens. Matching propositional content and formal markers: Sentence boundaries in a DGS text. **Sing Language & Linguistics**, v. 10, n. 2, p. 145-175, 2007.

HELLWIG, B.; GEERTS, J. **ELAN** - Linguistic Annotator. Versão 5.0. Disponível em: <https://www.mpi.nl/corpus/html/elan/index.html>. Acesso em: 11 set. 2018.

HIRST, Daniel. **Analyse tier PRAAT script**, 2012. Disponível em: [https://uk.groups.yahoo.com/neo/groups/praat-users/files/Daniel\\_Hirst/analyse\\_tier.praat](https://uk.groups.yahoo.com/neo/groups/praat-users/files/Daniel_Hirst/analyse_tier.praat). Acesso em: 15 maio 2013.

KENT, Ray D.; READ, Charles. **Análise acústica da fala**. Tradução Alexandre Rodrigues Meireles. São Paulo: Cortez, 2015.

KLATT, Dennis H. Linguistic uses of segment duration in English: Acoustic and perceptual evidence. **Journal of the Acoustical Society of America**, v. 59. p. 1.208-1.221, 1976.

LEITE, Tarcísio de A. **A segmentação da língua de sinais brasileira (Libras): um estudo linguístico descritivo a partir da conversação espontânea entre surdos.** Tese de Doutorado, Universidade de São Paulo, São Paulo: USP, 2008.

LOPEZ, Eduardo; CAMINERO, Javier; CORTAZAR, Ismael; HERNANDEZ, Luis A. Improvement on connected numbers recognition using prosodic information. **Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP)**, Sydney, Australia, paper 0353, 1998.

MARQUES, Ewerton Carlos M. **Aspectos prosódicos de duração e pausa de agrupamentos numéricos de telefone móvel: um enfoque na Libras.** Dissertação (Mestrado) – Programa de Pós-graduação em Linguística e Literatura. Maceió: Universidade Federal de Alagoas, 2020.

MOKBEL, Chafic; MAUURY, Laurent; KARRAY, Lamia; JOUVET, Denis; MONNÉ, Jean; SIMONIN, Jacques; BARTKOVA, Katarina. Towards improving ASR robustness for PSN and GSM telephone applications, **Speech Communication**, 23, 1-2, p. 141-159, 1997.

MUSILYU, Oyedeji. **Características prosódicas dos números telefônicos no português brasileiro.** Dissertação (Mestrado) – Programa de Pós-graduação em Letras e Linguística. Maceió: Universidade Federal de Alagoas, 2014.

MYERS, Scott; HANSEN, Benjamin B. The origin of vowel length neutralization in final position: Evidence from Finnish Speakers. **Natural Language & Linguistic Theory**, v. 25, p. 157-193, 2007.

NICODEMUS, Brenda; SMITH, Caroline L. Prosody and Utterance Boundaries in American Sign Language Interpretation. **Proceedings of the Annual Meeting of the Berkeley Linguistics Societ**, v. 32, n. 1, p. 275-285, 2006.

OLASZI, Péter. Analysis of Written and Spoken Form of Hungarian Numbers for TTS Applications, **International Journal of Speech Technology**, 3, 3, p. 177-186, 2000.

OLIVEIRA JR., Miguel. **Prosodic features in spontaneous narratives.** Thesis (Doctor of Philosophy). Department of Linguistics, Simon Fraser University, Vancouver, 2000.

PFAU, Roland; QUER, Josep. Nonmanuals: their grammatical and prosodic roles. In Diane Brentari (ed.), **Sign Languages** (Cambridge Language Surveys). Cambridge: Cambridge University Press. 2010, p. 381-402.

R Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2016. Disponível em: <https://www.r-project.org/>. Acesso em: 21 jun. 2016.

RAHIM, Mazin; RICCARDI, Giuseppe; SAUL, Lawrence; WRIGHT, Jerry; BUNTSCHUH, Bruce; GORIN, Allen. Robust numeric recognition in spoken language dialogue. **Speech Communication**, vol. 34, n. 1-2, p. 195-212, 2001.

REED, Beatrice Szczepek. **Analysing Conversation: an introduction to prosody**. Houndmills: Palgrave Macmillan, 2011.

RODRIGUES, Frederico. **Reconhecimento Robusto de Dígitos e Números Naturais**. Tese de doutorado. Universidade Técnica de Lisboa, Portugal, 2001.

SANDLER, Wendy. Visual Prosody. *In: Roland Pfau; Markus Steinbach; Bencie Woll (ed.), Sign Language: an international handbook*. Berlin: Mouton de Gruyter p. 55-76, 2012.

SANTOS, Ricardo F. **A autoria na interpretação de Libras para o Português: aspectos prosódicos e construção de sentidos na perspectiva verbo-visual**. Dissertação (Mestrado) – Programa de Pós-graduação em Linguística Aplicada e Estudos da Linguagem. São Paulo: Pontifícia Universidade Católica de São Paulo, 2018.

SWERTS, Marc. Prosodic features at discourse boundaries of different strength. **Journal of the Acoustical Society of America**, vol. 101, n. 1, p. 514-521, 1997.

WANG, Chao; SENEFF, Stephanie. A study of tones and tempo in continuous Mandarin digit strings and their application in telephone quality speech recognition. **Proceedings of the International Conference on Spoken Language Processing (ICSLP 98)**, Sydney, Australia, p. 695-698, 1998.

WARNER, Natasha. Methods for studying spontaneous speech. *In: COHN, Abigail C.; FOUGERON, Cécile; HUFFMAN, Marie K. (eds.). The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press, 2012, p. 621-633.

WIESE, Heike. **Numbers, Language, and the human mind**. Cambridge: Cambridge University Press, 2003.

## PARADIGMA DO MUNDO VISUAL MÉTODO DE RASTREAMENTO OCULAR

### VISUAL WORLD PARADIGM EYE TRACKING METHOD

*René Alain Santana de Almeida<sup>1</sup>*

*Miguel Oliveira Jr<sup>2</sup>*

*Reinier Cozijn<sup>3</sup>*

**RESUMO:** O presente artigo objetiva apresentar alguns critérios metodológicos a serem observados em experimentos com o método de rastreamento ocular do paradigma do mundo visual (*VWP*). Este paradigma experimental consiste em método *on-line* de processamento linguístico através da apresentação simultânea de estímulos auditivos e visuais a participantes cujas medidas temporais de sacadas e fixações dos movimentos dos olhos são capturadas a partir do rastreador ocular. Para tanto, é indispensável revisitar estudos prévios e atentar para alguns cuidados na elaboração e gravação das sentenças (estímulos auditivos), bem como na confecção das imagens (estímulos visuais). Consideramos que atentar para esses cuidados pode auxiliar o(a) pesquisador(a) a evitar retrabalhos e proporcionar maior confiabilidade aos resultados de suas análises. Entretanto, a decisão sobre qual critério seguir depende dos objetivos e do objeto de estudo de cada investigação.

**Palavras-chave:** Paradigma do mundo visual; Rastreamento ocular; Critérios metodológicos.

---

<sup>1</sup> Universidade Federal de Sergipe (UFS).

<sup>2</sup> Universidade Federal de Alagoas (UFAL).

<sup>3</sup> Tilburg University (UVT).



**ABSTRACT:** This paper aims to present some methodological criteria to be observed in experiments with the eye tracking method of the visual world paradigm (VWP). This experimental paradigm is an online method of linguistic processing through the simultaneous presentation of auditory and visual stimuli to participants whose temporal measurements of saccades and fixations of eye movements are captured from the eye tracker. For that, it is essential to revisit the previous literature and pay attention to some precautions in the elaboration and recording of sentences (auditory stimuli), as well as in the making of images (visual stimuli). We consider that paying attention to these criteria can help the researcher to avoid rework and provide greater reliability to the results of the experiment. However, the decision on which criterion to follow depends on the objectives and the object of study of each investigation.

**Keywords:** Visual world paradigm; Eye tracking; Methodological criteria.

## 1. INTRODUÇÃO

O rastreamento ocular é uma das técnicas mais apuradas de avaliação *on-line* do processamento linguístico (MITCHELL, 2004; KAISER, 2013) e neste capítulo, abordamos este método de análise. Os diversos movimentos dos olhos, durante o processamento de qualquer informação, podem ser utilizados para se inferir a maneira como tal informação é processada, seja lendo (estímulo de leitura) ou olhando alguma imagem (estímulo visual). Vale destacar que, conforme Duchowski (2007), em todo trabalho de rastreamento do olhar há uma suposição tácita muito importante que é normalmente aceita: supõe-se que a atenção está ligada à direção do olhar foveal, mesmo sabendo que pode não ser sempre assim.

Os olhos são os órgãos responsáveis pelo sentido da visão que, de acordo com Teixeira e Soares (2014), é responsável por grande parte da cognição humana. Durante o processamento linguístico, os olhos podem executar movimentos sacádicos (saltos) para frente (progressivos) ou para trás (regressivos), ou ainda podem permanecer fixos (fixações) por um curto período de tempo em uma determinada área, nos intervalos entre as sacadas. Desse modo, a despeito de haver outros movimentos oculares (RAYNER, 1998), os dois tipos mais analisados em experimentos *on-line* são: fixação e sacada.

Em experimentos de leitura (MOXEY et al., 2004; TRAXLER, 2009; MAIA, 2010, por exemplo), os autores costumam analisar ambas as medidas, enquanto que em experimentos baseados no paradigma do mundo visual

(KAMIDE, ALTMANN e HAYWOOD, 2003; COZIIN et al., 2011; ALMEIDA, 2017, dentre outros), os autores costumam analisar as fixações realizadas nas imagens, embora as sacadas sejam observadas a fim de verificar se o processamento acompanhou o que foi enunciado.

As sacadas duram em média de 20 a 50 ms (KAISER, 2013) e podem ser: progressivas, ou seja, os saltos que o olho faz de uma fixação até a fixação à frente, no sentido da leitura; ou regressivas (regressões), que são movimentos sacádicos realizados no sentido oposto da leitura. Para Rayner (1995), cerca de 10 a 15% dos movimentos sacádicos realizados durante a leitura de um texto são regressivos. As regressões normalmente são utilizadas para conferir uma palavra que foi pulada na leitura ou que não foi compreendida (YOKOMIZO et al., 2008).

Yokomizo et al. (2008) definem fixações como sendo breves períodos de tempo durante os quais o olho permanece examinando uma pequena área do estímulo visual (texto escrito ou imagens), com foco na região foveal. A duração média da fixação é de 200 a 250 ms (RAYNER, 1995; KAISER, 2013), período em que os leitores podem adquirir alguma informação útil a partir do texto lido ou imagem visualizada.

Rayner (1998) tece ponderações em torno do termo fixação, uma vez que os olhos nunca estão totalmente parados: i) há um tremor constante dos olhos, que ajuda as células nervosas da retina a se manterem ativas; ii) o controle do sistema motor ocular por parte do sistema nervoso provoca movimentos pequenos e lentos; e iii) quando isso acontece, os olhos se movem muito mais rápido de volta para onde eles estavam – são as micro sacadas, que acontecem “dentro” de uma fixação. Entretanto, esses movimentos são tão pequenos que a maioria dos pesquisadores os consideram “ruídos” e adotam procedimentos que os ignoram.

Portanto, durante a leitura de um texto ou visualização de imagens, deslocamos nossos olhos de modo a colocar caracteres ou imagens na região foveal para enxergarmos com mais nitidez. Durante essa movimentação, realizamos sacadas, período em que os olhos se movem de uma fixação para outra, e fixações, período de tempo entre as sacadas que fornece a indicação de que a informação está sendo obtida a partir do texto (KAISER, 2013). A autora afirma ainda que os paradigmas de pesquisa com rastreamento ocular são o paradigma de leitura (compreensão da escrita) e o paradigma do mundo visual (compreensão da fala), este focado no presente estudo.

Ambas as medidas, sacadas e fixações, são medidas temporais e, portanto, uma atenção metodológica minuciosa deve ser dispensada para que problemas na elaboração dos estímulos (auditivos e visuais) não interfiram no resultado obtido.

Assim, o presente texto objetiva apresentar alguns cuidados metodológicos que devemos ter ao pensar na elaboração de um experimento *on-line* utilizando o paradigma do mundo visual (*visual world paradigm – VWP*) com rastreamento ocular.

Esse método de rastreamento ocular baseia-se em um paradigma (*VWP*) que, em vez de usar estímulos escritos, procura analisar o processamento linguístico durante a produção e/ou compreensão de enunciados orais. No *VWP*, os participantes são expostos a estímulos linguísticos auditivos (em estudos de percepção) ou são estimulados a produzir língua falada (em estudos de produção). Em estudos de compreensão (percepção), Kaiser (2013) afirma que na maioria das pesquisas os estímulos auditivos são acompanhados de um estímulo visual (objetos ou imagens).

Para Huettig, Rommers e Meyer (2011) os estudos de produção, assim como os de compreensão, revelam que a inspeção visual dos participantes ao que é exibido é totalmente coordenada com o seu processamento linguístico, porque, em ambos os casos, os movimentos oculares refletem a direção da atenção visual. Assim, os falantes e ouvintes usam a atenção visual a fim de lidar eficientemente com as respectivas tarefas de produção de enunciados e de mapeamento dos enunciados que eles ouvem a respeito dos estímulos visuais. Ou seja, as pessoas que realizam tarefas linguísticas olham para objetos relevantes não apenas para identificá-los, mas também porque facilita a recuperação de informações sobre esses objetos.

O monitoramento do olhar humano permite não apenas compreender melhor o funcionamento do próprio aparato visual, mas também estabelece indicadores confiáveis dos processos atencionais, sendo, por isso, de grande interesse para a investigação de processos cognitivos, tais como aqueles envolvidos na compreensão da linguagem (MAIA, 2010, p. 26).

Portanto, o rastreamento dos olhos consiste em uma ferramenta essencial para mensurar com precisão os movimentos oculares, detalhando o número e a duração das fixações, bem como o comprimento e a frequência dos movimentos sacádicos. Tal equipamento de captação de movimentos oculares, de acordo com Klein e Bulla (2010), possui uma aplicabilidade extensa, não apenas no campo da Linguística, mas também na sua relação com outras áreas, tais como Psicologia, Medicina, Biologia, Computação e Educação. Os autores salientam que existem diversas possibilidades de estudos linguísticos envolvendo este equipamento, “desde a análise de frequência vocabular, fronteira de palavra e relevância semântica, até o estudo do efeito de palavras cognatas na aprendizagem de uma segunda língua” (KLEIN e BULLA, 2010, p. 10).

## 2. CUIDADOS METODOLÓGICOS PARA EXPERIMENTOS COM O VWP

Em experimentos com o VWP, assim como em outros paradigmas da linguística experimental, é indispensável atenção ao rigor metodológico com os estímulos que serão utilizados. Nas próximas seções, apresentamos alguns cuidados a serem observados a fim de proporcionar análises com resultados mais fiáveis.

### 2.1. Elaboração e seleção de sentenças

As sentenças devem ser elaboradas pensando em todas as etapas da pesquisa, já que as suas gravações serão utilizadas em todas elas, inclusive no experimento *on-line* com o rastreador ocular cujas respostas aos estímulos ocorrem em tempo real. Nesse sentido, tomemos como exemplo a sentença do tipo “o pai abraçou o filho suado” (semelhante às sentenças em MAGALHÃES e MAIA, 2006 e FONSECA, 2008), que satisfaz as condições de ambiguidade sintática estudadas em algumas pesquisas que objetivam verificar estratégias de desambiguação do atributo. Ou seja, identificar quais estratégias fazem com que os participantes interpretem por uma aposição não local (quem estava “suado” era “o pai”, primeiro sintagma nominal) ou por uma aposição local (quem estava “suado” era “o filho”, segundo sintagma nominal).

Embora seja um tipo de estrutura ambígua, precisamos atentar para algumas características da metodologia de rastreamento ocular do paradigma do mundo visual, que considera a análise de medidas temporais em tempo real. Portanto, esse tipo de sentença não seria apropriada para um experimento com o rastreador ocular, devido a algumas razões:

- i) os núcleos dos sintagmas nominais (SN) de uma ou duas sílabas talvez não ofereçam tempo suficiente para que o participante planeje a sua próxima fixação, ou seja, o tempo referente ao período do processamento de um estímulo até o início da resposta motora (movimentação dos olhos), que, de acordo com Altmann e Kamide (2004) leva cerca de 200 ms;
- ii) se queremos estudar o movimento dos olhos durante o processo de desambiguação de uma sentença desse tipo (quem estava suado?), precisamos de um estímulo visual distrator e seu correspondente auditivo, após o segundo protagonista, que faça com que o participante desvie o olhar dos protagonistas e depois retorne a um deles indicando a sua preferência pela aposição local ou não local. Caso contrário, quando o participante mantiver o olhar no segundo sintagma nominal, não será possível saber

se isso ocorreu devido à preferência no processamento da desambiguação ou porque o olhar já estava nele, uma vez que não teria ocorrido nenhum movimento sacádico; afinal, como explica Kaiser (2013), as pessoas não podem fazer movimentos sacádicos para o que já estão olhando;

- iii) é necessário também outro estímulo auditivo após a distração visual e antes do atributo, caso o interesse seja verificar se o participante retorna o olhar a um dos protagonistas antes mesmo do atributo ser enunciado (observação de movimentos antecipatórios). Além disso, caso o pesquisador tenha a intenção de verificar de que forma uma pausa antes do termo ambíguo influencia na desambiguação do enunciado, a presença, por exemplo, de um intensificador do atributo permite que os dois elementos em conjunto (“bastante suado”) apresentem potencial para formar um sintagma fonológico, e até um sintagma entoacional, proporcionando uma fronteira ideal para pausa, conforme o fraseamento prosódico de Nespor e Vogel (2007[1986]). Essa decisão apresenta, de acordo com Fonseca (2010), material fonético suficiente para fazer com que a pausa não seja estranha ao ouvinte.

Por essas razões, na pesquisa de Almeida (2017), por exemplo, inspirada nesse tipo de estrutura ambígua (MAGALHÃES e MAIA, 2006 e FONSECA, 2008), as sentenças experimentais foram elaboradas com a seguinte estrutura: SN1-Verbo-SN2-Advérbio de lugar-Advérbio de intensidade-Atributo (“O sanfoneiro abraçou o pagodeiro no palco bastante saudoso”) e seguindo critérios:

- i. os núcleos dos sintagmas nominais foram substantivos paroxítonas com quatro sílabas, desde que fosse possível estabelecer alguma relação entre eles, para não parecer tão estranho ao participante do experimento, e também que fosse possível distinguir claramente as figuras que os representam (conforme melhor explanado no tópico sobre os cuidados na preparação dos estímulos visuais);
- ii) verbos interpessoais com três sílabas, no passado e neutros em relação aos adjetivos, de modo a evitar qualquer direcionamento do atributo, provocado pelo verbo, a um dos sintagmas nominais, sobretudo ao SN que refere ao sujeito que praticou a ação verbal;
- iii) advérbios de lugar que não se referissem diretamente a um ou outro protagonista ou que se referissem aos dois (para evitar preferências de interpretação provocadas pelo elemento locativo ou ainda criar uma nova

- ambiguidade) e que tivessem como núcleo substantivos com duas sílabas, já que seriam apenas elementos visuais distratores;
- iv) utilizou-se o mesmo advérbio de intensidade (“bastante”) uma vez que, se fossem utilizados outros, não haveria possibilidade de manter o mesmo número de sílabas em todas as sentenças, como geralmente é feito em experimentos dessa natureza para evitar comparações entre enunciados de durações distintas; e
  - v) adjetivos caracteristicamente temporários<sup>4</sup> (PINHO, 2005) e com três sílabas.

Vale destacar a importância de verificar previamente, por meio de um experimento *off-line*, se tais critérios foram realmente atendidos na elaboração/ gravação dos enunciados que serão utilizados no experimento *on-line* com o rastreador ocular. Nesse sentido, Kaiser (2013) destaca que paradigmas experimentais frequentemente combinam medidas *off-line* e *on-line*, a fim de evitar retrabalho para o pesquisador e para produzir evidências sobre o processamento da linguagem que não estariam disponíveis a partir de qualquer um dos métodos isoladamente.

Então, sabendo que algumas sentenças seriam descartadas por qualquer razão de ordem interpretativa em decorrência de critérios não atendidos e verificados no experimento *off-line* prévio, Almeida (2017) elaborou 48 sentenças experimentais. Além de seguir esses critérios, o autor evitou repetir substantivos, verbos, advérbios de lugar e adjetivos, de modo a evitar qualquer possível estranhamento por parte dos participantes nos experimentos que foram realizados (por exemplo, “O baterista recebeu o guitarrista no quarto bastante drogado”). E para verificar se a posição dos sintagmas nominais poderia interferir no processo de desambiguação, inverteu a ordem deles e obteve mais 48 sentenças do tipo SN2-SN1 (“O guitarrista recebeu o baterista no quarto bastante drogado”).

O balanceamento da posição dos referentes na sentença é essencial porque esta estratégia pode identificar se há alguma questão de ordem semântica interferindo nos resultados obtidos em estudos com enunciados ambíguos, por exemplo. Almeida (2017) mostrou também que tal procedimento pode ser verificado em experimento *off-line* prévio ao experimento com o rastreamento ocular, de modo a verificar se a posição do referente no enunciado influencia na desambiguação da sentença enunciada no estímulo auditivo.

---

<sup>4</sup> Adjetivos da Língua Portuguesa que indicam estados ou condições temporárias que o indivíduo apresenta em determinadas situações (emoção, humor e pensamento), ou seja, é a expressão de um determinado momento (PINHO, 2005).

Além de sentenças experimentais ambíguas, geralmente utilizam-se sentenças verificadoras de atenção não ambíguas (ou fortemente inclinadas a uma determinada interpretação, dado as escolhas lexicais), para verificar se os participantes estão atentos durante a realização do teste prévio de julgamento (*off-line*). Então, por exemplo, Almeida (2017) elaborou 22 sentenças para essa finalidade, das quais metade o participante deveria escolher pela aposição local (“O padre encontrou o marginal roubando na sacristia”) e a outra metade pela aposição não local (“O padeiro atendeu o cliente trabalhando na padaria”).

Salientamos que essas sentenças não possuem o caráter de distratoras, uma vez que a intenção, nesses casos, não é distrair os participantes do objetivo do experimento e sim verificar a atenção dos participantes por meio da utilização de enunciados não ambíguos. Por essa razão, não se considera necessária a utilização do dobro do número de sentenças experimentais como sentenças verificadoras, como é normalmente feito quando se utiliza sentenças distratoras. Em seguida, procede-se à gravação das sentenças experimentais e verificadoras de atenção, conforme critérios definidos na próxima seção.

Além dos enunciados gravados das sentenças experimentais que são testadas em experimento prévio, de forma *off-line*, é preciso elaborar sentenças distratoras para o experimento *on-line*, cujas gravações são utilizadas no experimento com o rastreador ocular a fim de distrair os participantes das sentenças experimentais e tentar não evidenciar o objetivo do estudo, seguindo orientações de Schütze e Sprouse (2013), que afirmam ser este um dos propósitos da utilização de sentenças distratoras. Ou seja, reduzir as chances de que os participantes se tornem conscientes de que um tipo de sentença particular está sendo testado, o que poderia desencadear estratégias de resposta conscientes e prejudicar os resultados da pesquisa.

Para isso, segundo Cowart (1997), as pesquisas geralmente utilizam pelo menos o dobro da quantidade de sentenças experimentais ao definir a quantidade de sentenças distratoras. Caso essa proporção seja menor do que o dobro, aumenta-se o risco de os participantes se tornarem cientes do tipo de sentenças que está sendo analisada no experimento. A esse respeito, o autor salienta ainda que o desejável seria três a quatro vezes mais sentenças distratoras do que experimentais, desde que não torne o experimento excessivamente longo.

Entretanto, precisamos destacar que se os estímulos auditivos referentes às sentenças experimentais forem apresentados simultaneamente com um estímulo visual em formato de triângulo isósceles (conforme explicado na seção 2.3), o estímulo visual apresentado simultaneamente aos estímulos auditivos

das sentenças distratoras também deve ser apresentado em formato de triângulo isósceles.

Para melhor entendimento, vamos retomar Almeida (2017) para exemplificar. Quando o estímulo auditivo da sentença experimental “o flautinista sucedeu o trompetista no rádio bastante confuso” foi apresentado ao participante, simultaneamente, um estímulo visual também era apresentado com três figuras: uma representando “o flautinista”, outra representando “o trompetista” e a figura distratora representando o sintagma preposicionado “no rádio”, dispostas em formato de triângulo isósceles com o elemento locativo na região centro-superior do estímulo visual (Figura 7, seção 2.3.3).

Dessa forma, como em seu estudo utilizou 20 sentenças experimentais, selecionadas a partir do questionário (teste *off-line*), e a fim de não tornar o experimento muito cansativo, o autor elaborou 40 sentenças distratoras para o bloco experimental, das quais metade foi do tipo “O advogado e o magistrado discordaram das provas no fórum” e as demais do tipo “O palhaço brincou ao lado do bombeiro no carro da corporação”, de modo que, assim como as sentenças experimentais, também estabelecessem relação com os estímulos visuais contendo dois protagonistas (referentes) e um elemento locativo, em formato de triângulo isósceles.

Embora não seja nosso objetivo, nesse texto, explicar como se elabora um desenho experimental, é importante destacar que antes do bloco experimental costuma-se realizar um bloco de prática, seguindo metodologia semelhante a estudos prévios (ALMEIDA, 2017; COZIJN et al., 2011, por exemplo), com uma quantidade pequena de sentenças experimentais distratoras (que não são utilizadas no bloco experimental), a fim de familiarizar os participantes ao experimento. Além disso, no início do teste, geralmente é realizada uma sessão de nomeação (conforme explicado na seção 2.3.1).

Ainda baseados em Cozijn et al. (2011) e Almeida (2017), salientamos a importância de elaborar declarativas de verificação de atenção (cerca de 25% das sentenças experimentais e distratoras), em pesquisas que utilizam o método de rastreamento ocular do *VWP*. A partir dessas declarativas, é possível obter um retorno por parte do participante (sim/não ou verdadeiro/falso) em relação ao estímulo que acabou de escutar e, assim, verificar o nível de atenção do participante ao experimento.

Após a elaboração das sentenças, experimentais e verificadoras de atenção, (para o teste prévio *off-line*) procede-se à gravação dos estímulos auditivos. Com o resultado do experimento prévio, é possível decidir quantas sentenças



experimentais serão utilizadas no experimento *on-line* com o rastreador ocular, e então elaborar as sentenças distratoras e as declarativas de verificação, para serem gravadas, conforme alguns cuidados apresentados na próxima seção.

## 2.2. Preparação de estímulos auditivos

Para a gravação das sentenças, assim como para a elaboração, também é essencial ter sempre em mente que os estímulos auditivos referentes às sentenças experimentais utilizados no teste de julgamento prévio serão utilizados, juntamente com estímulos visuais, no experimento *on-line*, e, portanto, deve-se evitar uma taxa de elocução muito rápida, uma vez que os participantes precisam de tempo para acompanhar no estímulo visual a sentença que está sendo processada durante o estímulo auditivo, conforme alertam Cozijn et al. (2011).

Nessa etapa da pesquisa, geralmente é solicitado que um locutor profissional, do sexo feminino ou masculino, falante nativo do português brasileiro, realize as gravações das sentenças sob a supervisão e orientação de um linguista, normalmente o(s) próprio(s) pesquisador(es). Para uma realização adequada das gravações, deve-se atentar a critérios técnicos e recomendações de Oliveira Jr. (2014), tais como utilizar um bom microfone tipo headset (DPA Headband 4066, por exemplo) e um bom gravador digital de flash (Marantz PMD661, por exemplo). Além disso, todas as gravações devem ser realizadas em uma cabine acústica, de modo a garantir alta qualidade dos áudios utilizados como estímulos auditivos nos experimentos.

Então, sugerimos a elaboração de listas com as sentenças elaboradas previamente, de modo que as frases experimentais e as verificadoras de atenção (utilizadas no teste de julgamento *off-line*) sejam apresentadas da forma mais nítida (e espaçada) possível, para tornar a gravação menos cansativa ao locutor. Recomendamos também que as sentenças apareçam aleatoriamente, por três vezes (caso seja elaborada apenas duas listas, assim haveria seis opções de gravação para cada sentença), a fim de que o locutor realize a leitura de forma mais natural e “neutra” (ambígua) possível, parando entre as sentenças, para evitar efeito de lista.

O passo seguinte é escolher a melhor gravação para cada sentença “neutra” (ambígua), dentre as opções gravadas, seguindo como critérios: i) a qualidade do som; ii) a naturalidade; e iii) a “neutralidade”, definida como tal a partir da ausência de qualquer pista prosódica que direcione o ouvinte a uma determinada interpretação.

Caso o experimento objetive verificar o papel da prosódia na desambiguação de enunciados, sugerimos que se solicite ao locutor que efetue a leitura das sentenças experimentais de modo a induzir um ouvinte dele a uma determinada interpretação quanto à ambiguidade em investigação. Assim, é possível verificar quais estratégias preponderantes o locutor utiliza para esse fim e, com base nelas, definir quais condições analisar. Ou seja, além da condição “neutra”, outras condições serão apresentadas pelo locutor ao imprimir na sua fala características prosódicas na tentativa de desambiguar o enunciado.

Então, pode surgir a seguinte dúvida: seria melhor utilizar as gravações do locutor tentando induzir seu ouvinte a uma determinada interpretação da ambiguidade ou seria melhor manipular acusticamente as gravações neutras de modo a acrescentar alguma característica prosódica utilizada como estratégia pelo locutor? A esse respeito, recomendamos preferencialmente que se faça a manipulação acústica das gravações das sentenças “neutras” em vez de usar as tentativas do locutor de desambiguar, em virtude de uma análise estatística posterior ser muito complicada, considerando que não haveria um controle geral sobre todos os elementos prosódicos envolvidos (com os dados manipulados, pode-se ter um controle maior).

Caso a opção seja realmente manipular acusticamente as gravações que serão utilizadas como estímulos auditivos no experimento *on-line*, realiza-se a manipulação dessas gravações a fim de alcançar as condições que serão analisadas. Para tanto, geralmente utiliza-se o aplicativo computacional Praat (BOERSMA e WEENINK, 2017) e todas as sentenças são estilizadas em 2 st, inclusive as “neutras”, para evitar comparar estímulos naturais com estímulos acusticamente manipulados, seguindo orientações de Wellmann et al. (2012).

Outro cuidado que devemos atentar para a gravação das sentenças é que as distratoras sejam gravadas seguindo os mesmos critérios utilizados para gravação das frases “neutras” experimentais do teste de julgamento prévio, pelo mesmo locutor, no mesmo ambiente acústico e com o mesmo equipamento, para minimizar ao máximo qualquer possibilidade de distinção entre os grupos de sentenças (experimentais e distratoras). Dessa forma, todos os estímulos auditivos utilizados no experimento *on-line* não podem apresentar nenhum sinal que destoe o grupo de sentenças distratoras das experimentais.

Entretanto, seguindo critérios adotados por Cozijn et al. (2011) e Almeida (2017), geralmente as declarativas de verificação são gravadas por um orador do sexo oposto ao que gravou as sentenças experimentais e distratoras, com o

objetivo de destacar aquelas em relação a essas, conforme os mesmos critérios adotados, no mesmo ambiente acústico e com o mesmo equipamento.

Além da elaboração das sentenças e de suas gravações para serem utilizadas como estímulos auditivos em experimentos com o método de rastreamento ocular do paradigma do mundo visual, o *VWP* também requer estímulos visuais que devem ser elaborados com a mesma cautela metodológica.

### 2.3. Preparação de estímulos visuais

Uma vez que as sentenças foram elaboradas, devidamente gravadas e que tais estímulos auditivos foram testados previamente em um experimento *off-line*, a fim de verificar que atendem ao objetivo proposto e, portanto, poderão ser utilizados no experimento *on-line*, precisamos atentar para a elaboração dos estímulos visuais referentes aos estímulos auditivos aptos a serem utilizados no experimento com o paradigma do mundo visual do rastreamento ocular.

Este paradigma envolve estímulos linguísticos e visuais, portanto, as(os) pesquisadoras(es) precisam prestar muita atenção às propriedades de suas exibições visuais, para evitar a criação de fatores de confusão no desenho experimental. A esse respeito, Kaiser (2013) alerta para as seguintes questões: i) o reconhecimento das imagens; ii) a saliência visual das imagens; iii) as suas posições na tela; e iv) a posição do olhar dos participantes antes do estímulo visual ser apresentado.

Além dessas, Almeida (2017) e Cozijn et al. (2011) afirmam que a repetição das imagens em mais de um estímulo apresentado também pode representar um problema e causar algum estranhamento por parte do participante, dependendo do objetivo do estudo. Aliás, é importante salientar que todos os cuidados metodológicos aqui descritos não são válidos para todas as pesquisas que utilizam o *VWP*, já que tudo dependerá do objetivo de cada investigação. Logo, algum critério levantado aqui pode não se adequar a uma ou outra pesquisa, mas outro certamente será válido.

#### 2.3.1. Reconhecimento das imagens

O reconhecimento das imagens é essencial para qualquer estudo com o paradigma do mundo visual de rastreamento ocular. Por essa razão, comumente é realizada uma sessão de nomeação das imagens com os participantes, antes de se iniciar o experimento, a fim de verificar se tais figuras são facilmente reconhecidas.

Então, se uma imagem em particular é difícil de reconhecer, certamente o participante levará mais tempo do olhar nesta imagem, já que precisará de mais esforço cognitivo para reconhecê-la. Vejamos, por exemplo, a Figura 1 e vamos tentar identificá-la. Qual profissional está representado pela imagem?

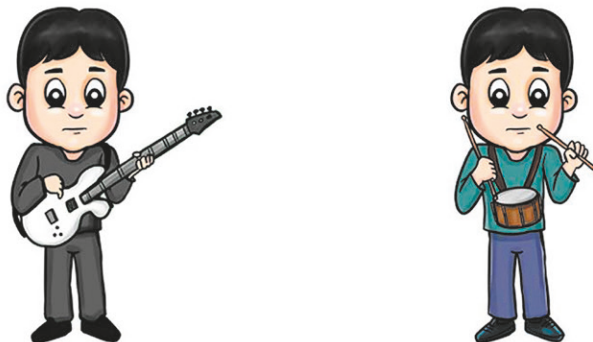
Figura 1 – Exemplo de figura não facilmente reconhecida



Fonte: J. Vanderson Caricatura Estúdio.

A Figura 1 representa uma tentativa de desenho de um baterista. Após sabermos do que se trata, talvez procuremos características na imagem que nos direcione até um baterista, mas não podemos afirmar que tal imagem seja facilmente reconhecida. Então, poderíamos solucionar esse problema inserindo a palavra “baterista” junto à imagem? Dessa forma, provavelmente iríamos gerar outro problema, já que uma figura alvo terá uma figura concorrente no mesmo estímulo visual (Figura 2). Logo, teríamos que orientar o desenhista a refazer o desenho de modo a torná-lo mais rapidamente identificável, sem a necessidade de uma informação escrita.

Figura 2 – Exemplo de imagem com figuras concorrentes de guitarrista e baterista



Fonte: adaptada de ALMEIDA, 2017.

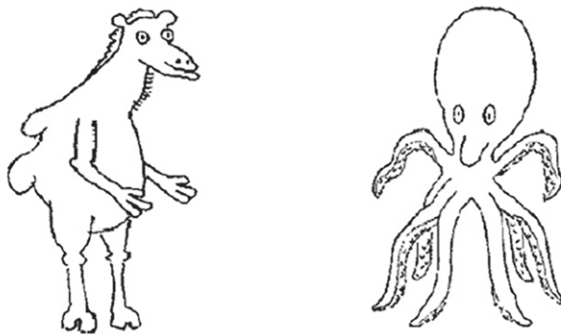
Portanto, se uma determinada imagem precisa de alguma informação adicional para ser reconhecida, que não existe na imagem concorrente, como por exemplo, uma informação escrita, o número de sacadas e fixações bem como as suas durações serão maiores do que na imagem concorrente, já que além de olhar para a imagem o participante ainda iria ler a informação dada no texto escrito. Esse aumento nos valores de tais medidas temporais não resulta do objeto de estudo e sim de um problema metodológico quando da elaboração dos estímulos visuais. Esta é uma das razões pelas quais há dois paradigmas de pesquisa com rastreamento ocular: ou o paradigma de leitura (compreensão da escrita), ou o paradigma do mundo visual (compreensão da fala, por meio de estímulos auditivos e visuais).

### 2.3.2. Saliência visual das imagens

Outro critério metodológico a que devemos atentar é a saliência visual das imagens. Quando uma imagem apresenta-se muito saliente (maior, mais brilhante ou mais colorida do que a sua concorrente, por exemplo), pode atrair uma alta proporção de fixações por esses motivos, enviesando os dados.

Nesse sentido, Cozijn et al. (2011) sugerem que as imagens sejam aproximadamente do mesmo tamanho e em preto e branco (Figura 3) para evitar problemas dessa natureza e Almeida (2017) acrescenta que, caso seja necessária a utilização de cores para facilitar o reconhecimento das imagens (Figura 2), o uso de cores vibrantes deve ser evitado, pois elas farão com que os participantes despendam maior tempo do olhar nessas imagens por serem mais coloridas e não por causa do objeto linguístico que está sendo investigado.

Figura 3 – Exemplo de imagem com figuras concorrentes



Fonte: adaptada de COZIEN et al., 2011.

Outro aspecto relevante relacionado à saliência das imagens é de caráter sociocultural. Bock, Irwin e Davidson (2004) salientam a importância de figuras neutras entre elas e em relação ao foco do estudo. Por isso, os autores orientam a evitar o uso de figuras representativas de sexos opostos, por exemplo, “rei” e “rainha”, uma vez que figuras masculinas podem preceder femininas devido à proeminência tradicional de homens, na maioria das sociedades.

Cozijn et al. (2011) alertam ainda para a possibilidade de o status social dos participantes da pesquisa influenciar no direcionamento do olhar, e sugerem que sejam utilizadas figuras de animais personificados, já que, para eles, os animais não são prontamente associados a um status social.

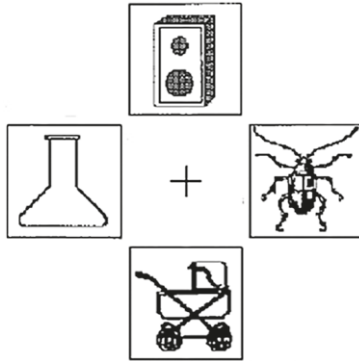
Entretanto, vale destacar que há trabalhos que se utilizam de questões semânticas e de conhecimento de mundo e precisam utilizar tais diferenças (de sexo, idade, ou classe social, por exemplo) em seus experimentos, como o trabalho desenvolvido por Kamide, Altmann e Haywood (2003), que verificou movimentos antecipatórios do olhar, decorrentes desse tipo de pista semântica apresentada pelo sujeito e pelo verbo, conforme sentenças do tipo “O homem vai experimentar...” e “A menina vai experimentar...”, em que os olhares dos participantes anteciparam para as figuras de “uma cerveja” e “um doce”, respectivamente, antes mesmo de tais informações serem enunciadas.

Portanto, as imagens que terão suas fixações posteriormente analisadas devem ser facilmente identificáveis, apresentar aproximadamente o mesmo tamanho e, sempre que possível, uma não deve possuir nada que chame mais a atenção do que a outra imagem concorrente. Além disso, precisamos atentar para as posições das imagens na tela.

### **2.3.3. Posições das imagens na tela**

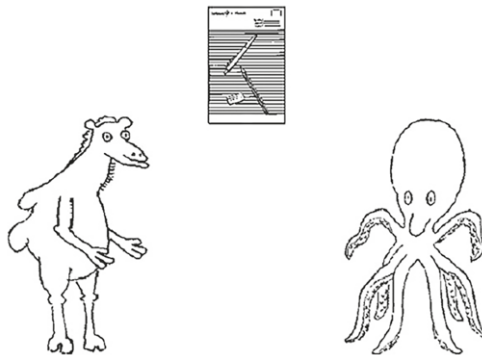
Com relação às posições das imagens, é possível perceber que na maioria dos estudos os estímulos visuais são mostrados na tela de um computador com imagens dispostas: i) em quatro quadrantes do monitor, como em Allopenna, Magnuson e Tanenhaus (1998) (Figura 4); ii) em formato de um triângulo isósceles, como em Almeida (2017), Kaiser (2011) e Cozijn et al. (2011) (Figura 5); ou ainda iii) em formato circular, por exemplo, Paulmann, Titone e Pell (2012) (Figura 6).

Figura 4 – Exemplo de disposição de imagens em quatro quadrantes



Fonte: ALLOPENNA, MAGNUSON e TANENHAUS, 1998.

Figura 5 – Exemplo de disposição de imagens em formato de triângulo isósceles



Fonte: COZIEN et al., 2011.

Figura 6 – Exemplo de disposição de imagens em formato circular



Fonte: PAULMANN, TITONE e PELL, 2012.

Embora haja uma variedade de disposições possíveis para os estímulos visuais, Kaiser (2013) salienta que, em geral, é muito importante balancear os locais dos objetos críticos, ou seja, dos objetos que posteriormente serão analisados no que se refere às medidas *on-line*. Isso é justificado pela tendência que temos de olhar para os estímulos visuais na mesma ordem que utilizamos na leitura e escrita, ou seja, em línguas como o inglês e o português, da esquerda para direita, e em línguas como o árabe e o hebreu, da direita para a esquerda. Portanto, para evitar que essas tendências interfiram nos dados dos movimentos oculares, a maioria dos estudos faz esse balanceamento de posições das imagens, já na etapa de construção do desenho experimental.

Essa é a razão pela qual, em experimentos com rastreador ocular, é comum contrabalançar a posição dos estímulos visuais correspondentes aos objetos alvo (COZIEN et al., 2011; PAULMANN, TITONE e PELL, 2012; NAKAMURA, ARAI e MAZUKA, 2012; ALMEIDA, 2017). Contrabalançar as posições dos protagonistas no monitor elimina esse viés nos dados resultantes dos movimentos dos olhos, além de evitar que a sua posição na tela seja confundida com a sua posição na sentença enunciada. Além disso, Weber, Braun e Crocker (2006) salientam a importância de evitar que os participantes desenvolvam expectativas quanto à posição do protagonista em experimentos dessa natureza.

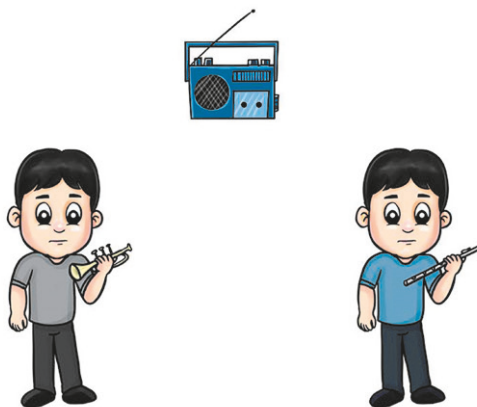
No entanto, esse balanceamento só é necessário entre as imagens que representam os objetos alvo, uma vez que há experimentos em que o estímulo auditivo deve fazer menção também a um objeto (distrator) que está em um local neutro na cena e equidistante dos potenciais concorrentes, de modo a desviar o olhar dos participantes antes que o objeto crítico seja mencionado. Nesse caso, o objeto distrator pode permanecer em um local fixo da cena. Podemos exemplificar com as pesquisas de Kaiser (2011) e Cozijn et al. (2011), que desenvolveram estudos sobre a interpretação de pronomes anafóricos para o inglês e holandês, respectivamente. Então, objetos distratores, que estavam no contexto da sentença ou da narrativa, eram enunciados antes do pronome crítico ocorrer a fim de direcionar os olhos dos participantes para um local neutro que estava equidistante de ambos os potenciais antecedentes para o pronome.

Dessa forma, as imagens concorrentes são mostradas com a mesma frequência, no lado esquerdo e no lado direito do estímulo visual, demonstrando que o experimento está bem equilibrado, seguindo Cozijn et al. (2011) e Almeida (2017). Vale salientar que a posição dos protagonistas no monitor geralmente não é considerada uma variável independente de análise sobre a qual formulamos qualquer tipo de hipótese, mas sim uma variável de controle utilizada justamente para que essa variável não influencie os resultados.



Para compreendermos melhor os cuidados metodológicos na elaboração de estímulos visuais expostos até o momento, tomemos como exemplo um estímulo apresentado em Almeida (2017), conforme Figura 7.

Figura 7 – Exemplo de estímulo visual (SN2-SN1) apresentado no experimento *on-line* durante o enunciado “O flautinista sucedeu o trompetista no rádio bastante confuso”



Fonte: ALMEIDA, 2017.

É possível observar que, além de serem facilmente identificáveis, as imagens que representam os protagonistas (sintagmas nominais) foram neutras (ambíguas) com relação a quem se refere o atributo (“quem estava confuso”). A figura distratora (“rádio”) foi colocada para desviar a atenção visual dos participantes para longe dos protagonistas, para que se tenha uma medida mais clara das mudanças de atenção em direção aos protagonistas nos segmentos críticos das sentenças enunciadas (“bastante”, “atributo” e ao término da sentença).

É importante destacar que os olhares para os protagonistas foram comparados e, por essa razão, os tamanhos das figuras que os representam devem ser (aproximadamente) iguais e apenas eles foram contrabalançados. Na Figura 7, por exemplo, o segundo SN enunciado (“o trompetista”) aparece à esquerda do primeiro SN (“o flautinista”). Outro grupo de participantes visualizou o primeiro SN à esquerda do segundo SN enquanto ouviam o mesmo enunciado. Almeida (2017) afirma que contrabalançar a figura distratora juntamente com os protagonistas iria apenas complicar o desenho experimental do estudo, uma vez que a sua posição fixa não dificulta a medição de atenção aos protagonistas (objetos alvo da análise), seguindo metodologia semelhante a Cozijn et al. (2011) e Kaiser (2011).

Assim, o esperado, em experimentos dessa natureza, é que quando o participante ouvir “o flautinista”, o maior número de fixações seja no protagonista referente a ele; ao escutar “o trompetista”, o maior número de fixações seja na ilustração do trompetista; ao ouvir “no rádio”, o participante desvie o olhar de ambos os protagonistas, olhe para a ilustração do “rádio” e depois retorne a um dos protagonistas, talvez antes mesmo de ouvir o adjetivo “confuso”, demonstrando assim a sua preferência pela aposição local ou não local. Caso o participante não olhe sistematicamente para a ilustração referente ao estímulo auditivo, os dados do participante devem ser descartados da análise, uma vez que isso é requisito para validar o paradigma do mundo visual, de acordo com autores como Altmann e Kamide (2004) e Cozijn et al. (2011).

#### **2.3.4. Posição inicial do olhar dos participantes**

Outra questão levantada por Kaiser (2013) se refere à posição do olhar do participante antes de iniciar os olhares para os objetos alvo, antes dos estímulos auditivo e visual serem apresentados. Nesse sentido, o ideal seria que os participantes olhassem para um lugar neutro equidistante dos objetos de interesse a fim de evitar qualquer tipo de viés. A maioria das pesquisas, de acordo com a autora, utiliza uma cruz (ou ponto) de fixação (Figuras 4 e 6) que direciona os olhos dos participantes para um local central e/ou neutro antes dos estímulos visual e auditivo serem apresentados.

Geralmente, os rastreadores oculares já estão configurados de modo que o participante do experimento só consiga avançar para a apresentação de tais estímulos ao olhar para essa cruz (ou ponto) de fixação, garantindo assim que o olhar de todos os participantes do experimento partiu de um ponto equidistante das imagens que representam os objetos de interesse.

#### **2.3.5. Repetição das imagens**

Outra questão a ser considerada na elaboração do estímulo visual, na verdade, desde a elaboração das sentenças que serão gravadas, conforme Almeida (2017) e Cozijn et al. (2011), é a repetição das imagens em mais de um estímulo apresentado. Essa repetição também pode provocar algum estranhamento por parte do participante do experimento, dependendo do objetivo do estudo. Na pesquisa de Almeida (2017), por exemplo, o autor apresentou em um dos seus estímulos um guitarrista, um baterista e um quarto, caso alguma dessas imagens se repetisse em outro momento do experimento, os participantes da pesquisa poderiam perder tempo processando: i) quem estava no quarto não era o guitarrista

e o baterista, por que agora está no quarto o carateca e o pugilista? (caso fosse repetido o elemento distrator locativo); e ii) o guitarrista/baterista não estava com o baterista/guitarrista, por que agora está aparecendo com o jardineiro? (caso um dos protagonistas fosse repetido em outro estímulo).

Observem que essa repetição, além de poder provocar um prolongamento no tempo de processamento da informação linguística e conseqüentemente distorcer os resultados das medidas de análise, podem ainda distrair o nível de atenção do participante a ponto de ser descartado da análise, uma vez que, para verificar a validação do método do paradigma do mundo visual, é necessário que, ao ouvir um segmento, o participante olhe para o estímulo visual correspondente, e isso pode não ocorrer, pois o estranhamento pode fazer o participante olhar fixamente para apenas uma área do estímulo visual, em qualquer momento do enunciado.

Embora tais cuidados sejam essenciais para não provocar nenhum viés nos resultados, vale destacar que, em estudos de paradigma do mundo visual, normalmente utilizamos movimentos oculares instrumentalmente para revelar qual referente (protagonista) está em foco durante o processamento linguístico, a qualquer momento da sentença enunciada, em tempo real. Portanto, o interesse maior não está no processamento visual das imagens, mas sim no objeto linguístico investigado. Entretanto, tais critérios metodológicos devem ser levados em consideração, uma vez que problemas na elaboração das imagens para o estímulo visual podem comprometer os resultados da pesquisa.

### 3. RECOMENDAÇÕES FINAIS

O presente texto objetivou apresentar alguns critérios metodológicos a serem observados em experimentos com o método de rastreamento ocular do paradigma do mundo visual (*VWP*), sobretudo no que se refere à elaboração dos estímulos auditivos e visuais utilizados em experimentos dessa natureza. Quanto mais atenção é dada a esses cuidados metodológicos, e de forma preventiva em vez de corretiva, menos retrabalho haverá durante a elaboração do experimento e mais fiáveis serão os resultados da análise.

Então, desde a etapa inicial de elaboração das sentenças, e em todas as demais etapas, é necessário pensar, de forma proativa, no experimento *on-line* com o rastreador ocular. Ou seja, as gravações dessas frases serão utilizadas como estímulos auditivos não apenas em um teste de julgamento (*off-line*) prévio, mas também no experimento para análise, em tempo real, do processamento linguístico.

Conforme já salientamos, esse experimento *off-line* prévio é fortemente recomendado, sobretudo em pesquisas cujo objeto de estudo envolva ambiguidade, a fim de testar a ambiguidade das frases elaboradas e selecionar aquelas mais adequadas para o experimento *on-line*. Portanto, a decisão do(s) critério(s) a ser(em) seguido(s) dependerá dos objetivos e do objeto de cada estudo.

## REFERÊNCIAS

ALLOPENNA, Paul D.; MAGNUSON, James S.; TANENHAUS, Michael K. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. **Journal of Memory and Language**, vol. 38, p. 419-439, 1998.

ALMEIDA, René A. S. de. **A prosódia e o processamento *on-line* de sentenças ambíguas do português brasileiro**. Tese de Doutorado, Universidade Federal de Alagoas, Maceió, 2017.

ALTMANN, Gerry T. M.; KAMIDE, Yuki. Now you see it, now you don't: Mediating the mapping between language and the visual world. *In*: HENDERSON, John M.; FERREIRA, Fernanda (eds.), **The interface of language, vision, and action: eye movements and the visual world**. New York: Psychology Press, 2004, p. 347-386.

BOCK, Kathryn; IRWIN, David E.; DAVIDSON, Douglas J. Putting first things first. *In*: HENDERSON, John M.; FERREIRA, Fernanda (eds.), **The interface of language, vision, and action: eye movements and the visual world**. New York: Psychology Press, 2004, p. 249-278.

BOERSMA, Paul; WEENINK, David. **Praat: doing phonetics by computer** (Versão 6.0.36) [Computer program]. 2017. Disponível em: <http://www.praat.org/>.

COWART, Wayne. **Experimental syntax: Applying objective methods to sentence judgments**. Thousand Oaks, California: Sage Publications, 1997.

COZIIN, Reinier; COMMANDEUR, Edwin; VONK, Wietske; NOORDMAN, Leo G. M. The time course of the use of implicit causality information in the processing of pronouns: a visual world paradigm study. **Journal of Memory and Language**, vol. 64, n. 4, p. 381-403, 2011.

DUCHOWSKI, Andrew T. **Eye tracking methodology: theory and practice**. 2.ed. London: Springer-Verlag, 2007.

FONSECA, Aline Alves. **Pistas Prosódicas e o Processamento de sentenças ambíguas do tipo “SN1-V-SN2-Atributo” do Português Brasileiro**, Dissertação de Mestrado, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

FONSECA, Aline Alves. O efeito do peso dos constituintes prosódicos da desambiguação de orações relativas reduzidas. **ReVEL**, vol. 8, n. 15, p. 242-255, 2010.

HUETTIG, Falk; ROMMERS, Joost; MEYER, Antje S. Using the visual world paradigm to study language processing: A review and critical evaluation. **Acta Psychologica**, vol. 137, p. 151-171, 2011.

KAISER, Elsi. Focusing on pronouns: consequences of subjecthood, pronominalisation, and contrastive focus. **Language and Cognitive Processes**, vol. 26, n. 10, p. 1.625-1.666, 2011.

KAISER, Elsi. Experimental paradigms in psycholinguistics. *In*: PODESVA, Robert J.; SHARMA, Devyani. **Research Methods in Linguistics**. Cambridge: Cambridge University Press, 2013, p. 135-168.

KAMIDE, Yuki; ALTMANN, Gerry T. M.; HAYWOOD, Sarah L. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. **Journal of Memory and Language**, vol. 49, n. 1, p. 133-156, 2003.

KLEIN, Ângela I.; BULLA, Julieane P. Eye-tracking e a Linguística: aplicações e interfaces. **Letrônica**. n. 2, v. 3, p. 235-249, 2010.

MAGALHÃES, José Olímpio de; MAIA, Marcus. Pistas prosódicas implícitas na resolução de ambigüidades sintáticas: um caso de adjunção de atributos. **Revista da ABRALIN**, v. 5, n. 1 e 2, p. 143-167, 2006.

MAIA, Marcus. Rastreamento ocular de sintagmas preposicionais ambíguos em português. **Revista da ABRALIN**, vol. 9, n. 2, p. 11-36, 2010.

MITCHELL, Don C. On-line methods in language processing: introduction and historical review. *In*: CARREIRAS, Manuel; CLIFTON JR., Charles (eds.), **The On-line Study of Sentence Comprehension: Eyetracking, ERP and Beyond**. New York: Psychology Press, 2004, p. 15-32.

MOXEY, Linda M.; SANFORD, Anthony J.; STURT, Patrick; MORROW, Lorna I. Constraints on the formation of plural reference objects: the influence of

role, conjunction, and type of description. **Journal of Memory and Language**, vol. 51, n. 3, p. 346-364, 2004.

NAKAMURA, Chie; ARAI, Manabu; MAZUKA, Reiko. Immediate use of prosody and context in predicting a syntactic structure. **Cognition**, vol. 125, n. 2, p. 317-323, 2012.

NESPOR, Marina; VOGEL, Irene. **Prosodic phonology**: with a new foreword. Berlin: Mouton de Gruyter, 2007[1986].

OLIVEIRA JR., Miguel. Aspectos técnicos na coleta de dados linguísticos orais. *In*: FREITAG, Raquel Meister Ko. **Metodologia de Coleta e Manipulação de dados em Sociolinguística**. São Paulo: Blucher, 2014.

PAULMANN, Silke; TITONE, Debra; PELL, Marc D. How emotional prosody guides your way: Evidence from eye movements. **Speech Communication**, n. 54, p. 92-107, 2012.

PINHO, Cristina Coutinho Marques de. **Taxonomia brasileira da personalidade: um estudo dos adjetivos da língua portuguesa**. Tese de doutorado apresentada ao Programa de Pós-graduação em Psicologia da PUC-Campinas. Campinas, 2005.

RAYNER, Keith. Eye movements and cognitive processes in reading, visual search, and scene perception. *In*: FINDLAY, John M.; WALKER, Robin; KENTRIDGE, Robert W. **Eye Movement Research: Mechanisms, Processes and Applications**. Amsterdam: Elsevier, 1995, p. 3-22.

RAYNER, Keith. Eye movements in reading and information processing: 20 years of research. **Psychological Bulletin**, vol. 124, n. 3, p. 372-422, 1998.

SCHÜTZE, Carson T.; SPROUSE, Jon. Judgment data. *In*: PODESVA, Robert J.; SHARMA, Devyani. **Research Methods in Linguistics**. Cambridge: Cambridge University Press, 2013, p. 27-50.

TEIXEIRA, Elisângela N.; SOARES, Maria E. Movimentação ocular no estudo do processamento da referência. *In*: CAVALCANTE, Mônica M.; LIMA, Silvana M. C. de. **Referenciação: teoria e prática**. São Paulo: Cortez, 2014, p. 133-159.

TRAXLER, Matthew J. A Hierarchical Linear Modeling Analysis of Working Memory and Implicit Prosody in the Resolution of Adjunct Attachment Ambiguity. **Journal of Psycholinguistic Research**, vol. 38, n. 5, p. 491-509, 2009.

WEBER, Andrea; BRAUN, Bettina; CROCKER, Matthew W. Finding Referents in Time: Eye-Tracking Evidence for the Role of Contrastive Accents. **Language and Speech**, n. 49, p. 367-392, 2006.

WELLMANN, Caroline; HOLZGREFE, Julia; TRUCKENBRODT, Hubert; WARTENBURGER, Isabell; HÖHLE, Barbara. How each prosodic boundary cue matters: Evidence from German infants. **Frontiers in Psychology/Language Sciences**, vol.3, Art. 580, 2012.

YOKOMIZO, Juliana E.; LUKASOVA, Katerina; FONTELES, Daniel S. R.; MACEDO, Elizeu C. de. Movimentos sacádicos durante leitura de texto em crianças e universitários bons leitores. **O Mundo da Saúde**, vol. 32, n. 2, p. 131-138, 2008.

## SEGMENTAÇÃO AUTOMÁTICA E MANUAL DE NARRATIVAS ESPONTÂNEAS METODOLOGIA DE UM ESTUDO DE PERCEPÇÃO DA PROSÓDIA

*Ebson Wilkerson R. da Silva<sup>1</sup>  
Miguel Oliveira, Jr.<sup>2</sup>*

### 1. INTRODUÇÃO

Este capítulo apresenta os procedimentos metodológicos utilizados em uma pesquisa experimental sobre a percepção de fronteiras prosódicas. O objetivo da pesquisa foi verificar em que medida ouvintes reconhecem essas fronteiras e se as fronteiras reconhecidas coincidem com fronteiras prosódicas detectadas semiautomaticamente através de *scripts* do aplicativo *Praat*.

O capítulo pode ser lido como um guia para experimentos que envolvem percepção de fronteiras prosódicas da fala. Aqui apresentamos os passos que foram seguidos para a coleta dos dados e para a construção da pesquisa. Esses passos podem ser adaptados de acordo com as características dos estudos a serem desenvolvidos.

Organizamos o conteúdo deste capítulo da seguinte maneira: primeiro, apresentamos a descrição dos materiais de pesquisa e do experimento com os participantes, bem como tratamos do *corpus* da pesquisa que serviu como

---

<sup>1</sup> UFAL – Universidade Federal de Alagoas.

<sup>2</sup> UFAL – Universidade Federal de Alagoas.



base para a coleta do experimento de percepção; em seguida, descrevemos o experimento de percepção e o procedimento para seleção dos informantes, assim como o processo de coleta de dados de percepção; na seção seguinte, apresentamos os métodos utilizados para se obter a segmentação semiautomática, baseada em elementos prosódicos, como a duração e grupos acentuais, e também apresentamos conceitos de grupos acentuais e de unidades VV.

## 2. NARRATIVAS ESPONTÂNEAS

Quando os participantes de uma conversa decidem contar uma história durante uma conversa, sem que sejam solicitados a fazê-la, deixam transparecer o julgamento que fazem acerca do que acreditam ser suficientemente interessante para justificar o longo turno conversacional que ocuparão. Por essa razão, narrativas espontâneas constituem um excelente material de estudo da fala em contexto naturalístico, uma vez que há nesse tipo de produção uma motivação pragmática.

De acordo com Wolfson (1979), as pessoas sabem que regras são apropriadas em entrevistas como eventos de fala. Isso explicaria o fato de que algumas narrativas, quando eliciadas, são mais propensas a assumir a forma de resumos, pois são geralmente mais curtas, pontuais e exibem pouquíssimos detalhes, assim como respostas às perguntas em uma entrevista devem ser.

Para o nosso estudo, utilizamos um *corpus* de entrevistas espontâneas (WOLFSON, 1979). Na ocasião, os participantes foram solicitados a falar livremente sobre qualquer tema a partir de uma lista composta por 28 tópicos (Oliveira, 2000). Um total de 17 narrativas apareceram naturalmente nas entrevistas, na maioria das vezes, como uma ilustração de um dado argumento ou tópico.

Desse *corpus*, dez narrativas foram segmentadas, mas apenas seis foram selecionadas aleatoriamente para o estudo: duas foram utilizadas à guisa de ilustração e quatro narrativas foram separadas para o experimento de percepção<sup>3</sup>. As narrativas selecionadas derivam de entrevistas de quatro diferentes sujeitos: um homem e três mulheres, variando de 25 a 37 anos de idade, naturais de diferentes regiões do Brasil. No momento da coleta de dados, todos eram estudantes de pós-graduação.

As narrativas selecionadas tinham uma duração média de 69 segundos. Essas narrativas foram enumeradas (de 1 a 4), transcritas ortograficamente, de uma

---

<sup>3</sup> As narrativas podem ser ouvidas no seguinte repositório: [http://www.archive.org/details/Pesquisa\\_273](http://www.archive.org/details/Pesquisa_273).

maneira linear, sem marca de pontuação e sem indicação de paragrafação, para que a organização da estrutura do texto escrito não influenciasse os participantes na segmentação manual das narrativas. A Figura 1 mostra um exemplo de como as narrativas eram apresentadas aos participantes da pesquisa.

Figura 1 – Exemplo de narrativa espontânea transcrita ortograficamente. Exemplo 1

com a naty já aconteceu de tudo já nesse lance uma vez ela tava descendo na grouse fazendo um ziguezague daí tinha um uma criancinha também tava descendo né pequeninha acho que tinha uns quatro ou cinco anos fazendo ziguezague também daí no bem no meio da montanha os dois se encontraram assim se chocaram de frente assim daí caiu um pra cada lado assim daí uma pessoa levantou chacoalhou tirou a neve e foi embora né a outra ficou chorando então a naty ficou chorando lá a criancinha eu pensei que ela f tinha machucado né que nada levantou tirou a neve e foi embora e a naty ficou deitada lá chorando que pensou que tinha machucado a criança

### 3. MONTAGEM DO EXPERIMENTO DE PERCEPÇÃO

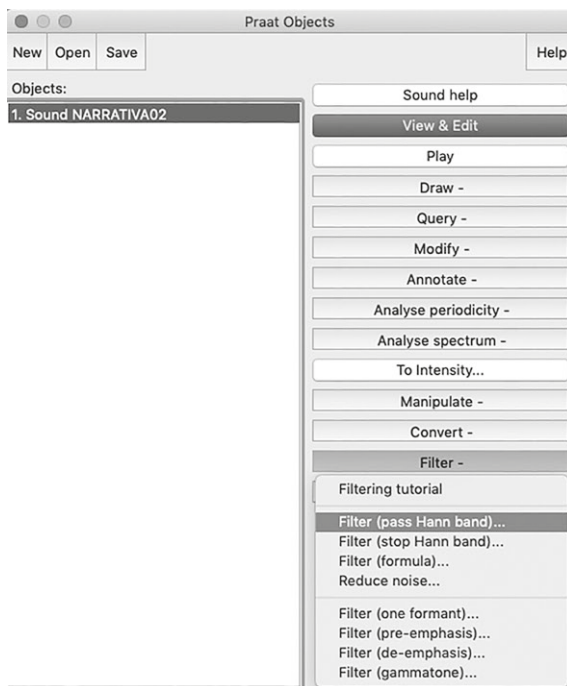
Seguindo com o exemplo do estudo que serviu como modelo para esse artigo, explicaremos os passos para montar o experimento de percepção utilizado na pesquisa.

Para o experimento de percepção, as narrativas foram divididas em quatro diferentes condições de apresentação (C):

- Na Condição 1 (C1), foi apresentada apenas a transcrição ortográfica da narrativa, feita de uma maneira linear, sem marca de pontuação e sem indicação de paragrafação.
- Na condição de apresentação 2 (C2), foram apresentados simultaneamente a transcrição da narrativa e o áudio a que a transcrição se refere.
- Na condição de apresentação 3 (C3), apenas o áudio da narrativa foi apresentado.
- Na condição de apresentação 4 (C4), foi apresentada uma versão deslexicalizada da narrativa.

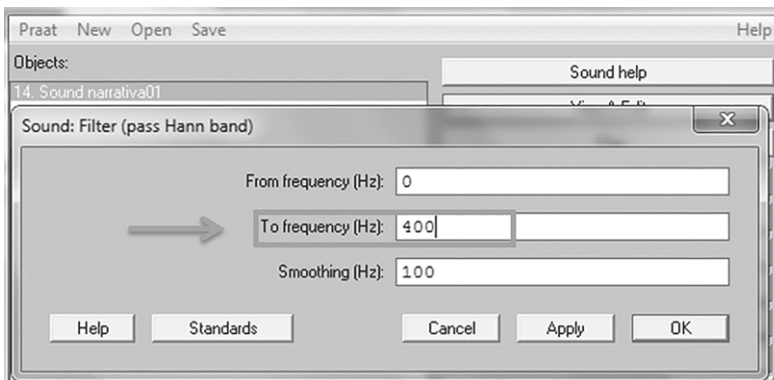
Para produzir a C4, ou seja, a versão deslexicalizada, o arquivo de áudio original foi filtrado com filtro passa-banda, selecionado através da função *filter* no aplicativo *Praat*, como ilustra a Figura 2:

Figura 2 – Janela do Praat: Utilizando filtro passa-banda.



A partir dessa filtragem, toda a informação de frequência superior a 400 Hz foi descartada (cf. Figura 3). Ou seja, a filtragem do áudio resulta em uma fala ininteligível, que preserva, no entanto, a informação prosódica.

Figura 3 – Seleção de valores de frequência, valores acima de 400 Hz são descartados.



As narrativas foram assim divididas para serem apresentadas aos participantes que fizeram a segmentação manual das narrativas. Sendo assim, teremos a segmentação feita da seguinte maneira: i) da leitura das narrativas (C1), que pode ser útil para verificar a atuação da prosódia implícita e ajudar a validar o modelo de intenções; ii) da leitura e audição da narrativa (C2); iii) exclusivamente da audição dos áudios das narrativas (C3); e iv) dos áudios *deslexicalizados*, que podem ser úteis para confirmar se os participantes basearam a sua segmentação por elementos prosódicos e se essa segmentação coincide com a segmentação das narrativas lidas e, além disso, se a segmentação manual coincide com a segmentação automática.

A seguir, explicaremos a sequência utilizada para segmentação manual, segmentação automática e seleção dos participantes.

## 4. SEGMENTAÇÃO MANUAL

Para a segmentação manual, nos baseamos no modelo teórico de Labov (1972). Nesse modelo, a narrativa é considerada um tipo de discurso que pode ser segmentado, tomando-se como princípio a função informacional de suas partes constituintes. De acordo com Labov (Idem), a estrutura de uma narrativa “bem formada” pode conter as seguintes seções: resumo, orientação, seção de complicação, avaliação, resolução e coda.

Entretanto, esse modelo laboviano precisa ser empiricamente validado, se quisermos considerá-lo como exemplo de modelo teórico estrutural de narrativas orais. Para isso, conforme Oliveira Jr. (2000), consideramos o nível suprasegmental do discurso narrativo. Um dos modelos estruturais mais influentes em estudos computacionais é aquele proposto por Passonneau e Litman (1993). Este modelo baseia-se na hipótese de que os falantes segmentam as narrativas que contam com o propósito – ou intenção – de deixar claro para o ouvinte a sua organização estrutural.

Esse modelo de intenções foi testado e validado para o português brasileiro no estudo piloto de Oliveira e Silva (2011). Além disso, foi verificado que, de modo geral, as pessoas reconhecem a localização de fronteiras discursivas e que a concordância tende a ser maior quando considerado o material sonoro das narrativas.

Assim, com o modelo validado, convidamos os participantes para a pesquisa. Vale ressaltar que, no entanto, em pesquisas que envolvem percepção, é necessário estabelecer critérios de inclusão dos participantes. Para que os participantes fizessem o solicitado nas instruções para que não houvesse dificuldades para

realizar a tarefa, utilizamos como critérios para inclusão do participante o grau de escolaridade – terceiro grau completo ou incompleto – e saúde auditiva.

A tarefa solicitada aos participantes foi indicar, nas narrativas, os pontos em que julgassem haver intenção, por parte do falante, de finalizar uma unidade comunicativa. Nenhuma definição para a noção de unidade comunicativa foi apresentada. Os participantes foram instruídos a julgar quais eram as fronteiras dessas unidades em caráter puramente subjetivo, tendo como base, apenas, os exemplos de narrativas segmentadas por intenção, que eram apresentados antes de fazerem as suas próprias segmentações, conforme exemplo apresentado no Quadro 1. O objetivo foi comparar esse julgamento dos informantes com uma segmentação puramente prosódica.

Todos os participantes leram e assinaram um Termo de Consentimento Livre e Esclarecido (TCLE), de acordo com a exigência do Ministério da Saúde e do Comitê de Ética em Pesquisa.<sup>4</sup>

Após a assinatura do TCLE, os participantes receberam instruções gerais impressas acerca do experimento, seguidas de exemplos: duas narrativas adicionais, extraídas do *corpus* supramencionado.

Os exemplos traziam narrativas segmentadas, seguidas de uma análise de suas partes constitutivas, considerando-se as possíveis intenções do falante, conforme o Quadro 1. Desse modo, o participante entenderia o que foi solicitado, quando estivesse diante das condições em que teria acesso às informações da narrativa. As segmentações e análises foram feitas pelos pesquisadores apenas à guisa de ilustração, e não como modelos a serem seguidos. Também, a título de ilustração, uma das narrativas foi apresentada em sua versão filtrada, para que os informantes se familiarizassem com o que ouviriam na C4.

---

<sup>4</sup> Estudo avaliado pelo Comitê de Ética em Pesquisa, aprovado para realização em 24/11/2010 (Processo n. 013526/2010-01).

Quadro 1 – Modelo de narrativa apresentada como exemplo aos participantes, segmentada em unidades comunicativas com base em intenções. Exemplo 2

pai também teve uma experiência assim de ele tava passando por uma rua a gente chama de portão do gelo esse lugar		<i>introduz a narrativa</i>
e agora ele tá bem mudado fizeram muita casa mas antes ele era muito cheio de árvore de mato não de árvore de mato era matagal		<i>descreve o lugar atual vs. passado em que aconteceu a narrativa</i>
e pai vinha por ali de noite e ele começou a ver a sentir ouvir passos atrás dele né e ele olhou pra trás olhou pro lado e não viu ninguém atrás dele aí ele começou a correr assustado e sentindo que a coisa vinha continuando perseguindo ele né vinha atrás dele correndo e ele correndo e a coisa correndo também		<i>apresenta o que aconteceu a seu pai</i>
uma experiência paranormal		<i>conclui dando um tema para a narrativa</i>

Os participantes foram então informados de que receberiam quatro diferentes narrativas em quatro diferentes condições (cf. Seção 3.1 O *corpus*). Foram solicitados a ouvir, ler, ou ler e ouvir cada uma das narrativas com atenção e, em seguida, segmentá-las de acordo com o julgamento que fizessem acerca da intenção do falante de finalizar uma unidade comunicativa. Eles poderiam levar o tempo que julgassem necessário para fazer a segmentação de cada uma das narrativas, podendo também ouvir as versões em áudio quantas vezes julgassem necessário para completar a tarefa.

Nas condições C1 e C2, em que receberam uma transcrição da narrativa, os participantes foram solicitados a fazer a segmentação na transcrição, por meio de barras transversais no papel. Nas condições C3 e C4, em que foi exposta apenas a versão em áudio da narrativa, os participantes foram solicitados a fazer a segmentação simultaneamente à execução do áudio, pressionando a tecla *Enter* de um teclado de computador. As respostas aos estímulos apresentados nessas duas condições (3 e 4) foram registradas no aplicativo computacional *ELAN* (BRUGMAN e RUSSEL, 2004).<sup>5</sup>

Durante todo o experimento, seguiu-se de forma semialeatória a ordem de apresentação das narrativas em diferentes condições aos informantes, ou seja, cada uma das quatro narrativas que compõem o *corpus* do estudo, em diferentes condições experimentais, apareceu três vezes em cada ordem de apresentação

<sup>5</sup> O *ELAN* é um aplicativo computacional desenvolvido pelo Max Planck Institute for Psycholinguistics: <http://tla.mpi.nl/tools/tla-tools/elan/>.

(4 narrativas x 4 condições x 3 ordens = 48 grupos de apresentação únicos). A apresentação semialeatória dos estímulos assegura a generalidade dos resultados, conforme descrito no Quadro 2.

Quadro 2 – Lista randomizada de apresentação das narrativas e suas respectivas condições

PARTICIPANTE	Ordem de apresentação das narrativas			
	NAR_CONDIÇÃO	NAR_CONDIÇÃO	NAR_CONDIÇÃO	NAR_CONDIÇÃO
1	N1_C1	N3_C4	N2_C3	N4_C2
2	N1_C2	N4_C1	N2_C4	N3_C3
3	N1_C3	N4_C2	N3_C1	N2_C4
4	N1_C4	N4_C3	N3_C2	N2_C1
5	N2_C1	N4_C4	N3_C3	N1_C2
6	N2_C2	N1_C1	N3_C4	N4_C3
7	N2_C3	N1_C2	N4_C1	N3_C4
8	N2_C4	N1_C3	N4_C2	N3_C1

O Quadro 2 apresenta a ordem e a condição em que foram apresentadas as narrativas aos participantes. Por exemplo, o informante 1 recebeu a Narrativa 1, na C1 – em que teria acesso apenas à transcrição; depois ouviria a Narrativa 3, na condição 4, em que teria acesso ao áudio deslexicalizado e assim por diante.

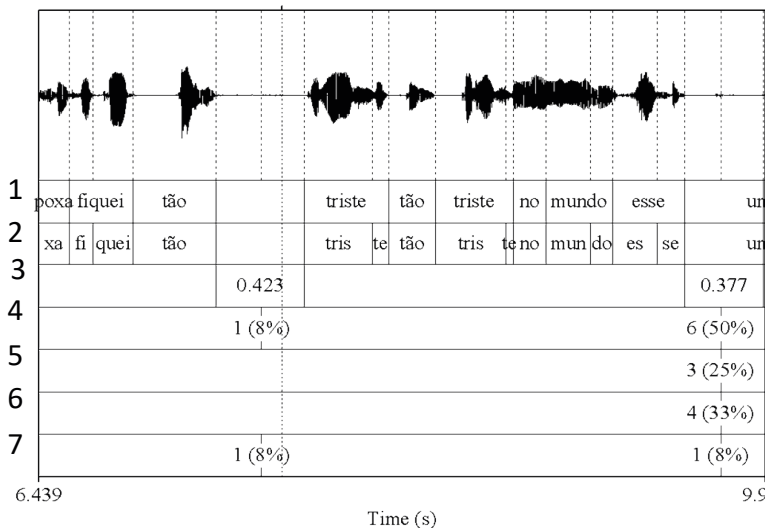
Em seguida, as respostas dos participantes foram salvas, separadas por narrativas e por condições. Cada uma das narrativas foi transcrita ortograficamente em uma tabela do Excel, segmentada por palavras, conforme é possível verificar no Tabela 1, a seguir. Os dados foram, então, computados, isto é, foram marcadas, ao final de cada palavra, a porcentagem de informantes por condição que julgou haver ali uma fronteira de unidade comunicativa. Também foram informadas na tabela a ocorrência e a duração de pausas em cada narrativa.

Tabela 1 – Computação dos dados de fronteiras percebidas pelos informantes

TRANSCRIÇÃO NARRATIVA	OCORRÊNCIA DE PAUSA	DURAÇÃO	FRONTEIRAS N4_01	FRONTEIRAS N4_02	FRONTEIRAS N4_03	FRONTEIRAS N4_04
Acidentes	Não	**	25%	8%	**	**
Aconteceu	Não	**	25%	16%	**	**
Um	Não	**	**	**	**	**
Acidente	Não	**	41%	8%	**	**
Mas	Não	**	**	**	**	**
Eu	Não	**	**	**	**	**
Não	Não	**	**	**	**	**
Lembro	Não	**	50%	25%	8%	**
Eu	Não	**	**	**	**	**
Tinha	Não	**	**	**	**	**
Dois	Não	**	**	**	**	**
Anos	Sim	0.837	41%	75%	100%	91%

A Tabela 1 apresenta um exemplo de segmentação para a narrativa 4, nas 4 condições. A tabela serve apenas como ilustração, a tabela completa foi publicada na seção de anexos em Silva, 2017. Após esta etapa, os dados foram transcritos no *Praat*.

Figura 4 – Tabela de transcrição do *Praat*.





Na Figura 4, temos uma tela do *Praat* que apresenta: na linha 1, a transcrição segmentada em palavras; na linha 2, a segmentação em sílabas; na linha 3 há a indicação de duração de pausas. Nas linhas 4, 5, 6 e 7, estão os valores de concordância (em porcentagem) nas quatro condições experimentais do estudo.

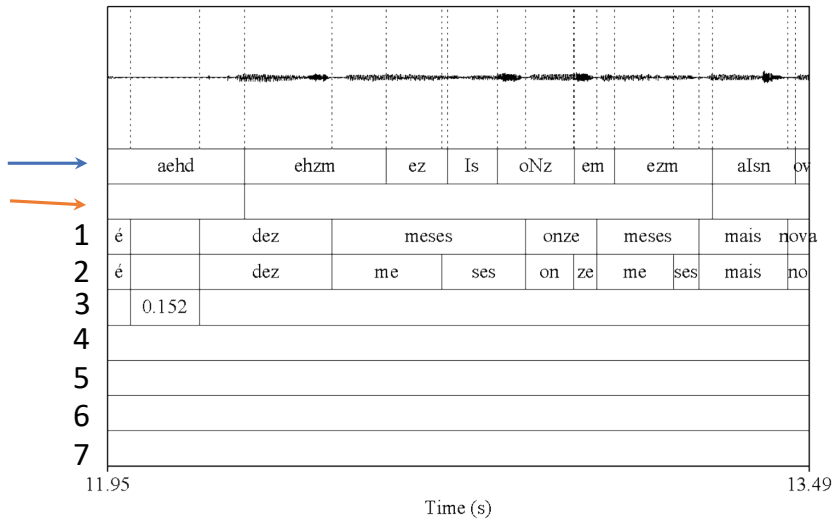
## 5. SEGMENTAÇÃO AUTOMÁTICA

Além da segmentação dos participantes, a qual foi realizada considerando as intenções do falante de terminar uma unidade comunicativa, também foi realizada uma segmentação por *scripts*. O uso dos *scripts* justifica-se por fornecer uma segmentação semiautomática baseada apenas em elementos prosódicos.

Para a segmentação dos grupos acentuais, utilizamos o *SG Detector* (BARBOSA, 2006). O *SG Detector* segmenta a narrativa considerando aspectos prosódicos, logo esta segmentação nos foi útil para comparar com a segmentação dos participantes. Entretanto, para que o *script* funcionasse corretamente, foi necessário fazer uma segmentação em unidades VV, por isso foi utilizado o *script BeatExtractor*, aplicativo implementado por Cummins e Port (1998) e adaptado para o *Praat* por Barbosa (2006). Ao segmentar o áudio em unidades VV, o *BeatExtractor* pode apresentar erros que precisam ser corrigidos manualmente. Assim, a segmentação foi semiautomática. Os conceitos de grupos acentuais e unidade VV e o funcionamento dos *scripts* serão discutidos na próxima seção.

Como é possível observar na Figura 5, a seguir, além das linhas já conhecidas na Figura 4, há também dados de transcrição segmentados em unidades VV (como indicado na linha em que temos a seta azul) e informações de grupos acentuais (*stress groups*) (é possível identificar a linha com a marcação dos grupos pela seta vermelha). A segmentação em unidades VV foi obtida através do *script Beat Extractor* e a segmentação em grupos acentuais através do *SG Detector*.

Figura 5 – Tabela de transcrição do Praat.



## 5.1 Grupos acentuais

Os grupos acentuais ou *stress group* (SG) são definidos em Lucente (2012, p. 96) como “unidades delimitadas por dois acentos frasais consecutivos”. O acento frasal é definido em Barbosa (2006) como uma proeminência do domínio da produção da fala, caracterizada por um pico local de duração ao longo do enunciado.

Barbosa (2006) apresenta três tipos de técnica que podem ser empregadas para delimitar os grupos acentuais a partir da localização dos acentos frasais, dependendo do interesse da pesquisa. A primeira é de base perceptiva e pode ser feita pela escuta do *corpus* de pesquisa por uma bateria de ouvintes, aos quais se solicita marcar as palavras que lhes parecem proeminentes. No entanto, essa técnica possui algumas dificuldades, a principal delas é definir a partir de qual porcentagem é possível considerar que há uma proeminência correspondente a um acento frasal que corresponde a uma fronteira de grupo acentual.

Outra técnica parte dos picos de duração de unidades VV ao longo do enunciado para selecionar, entre as palavras em que ocorreram os picos, as que são reconhecidas por pelo menos 20% dos ouvintes. Essa escolha também é arbitrária, porém é guiada por uma seleção inicial fundamentada em critérios de produção.

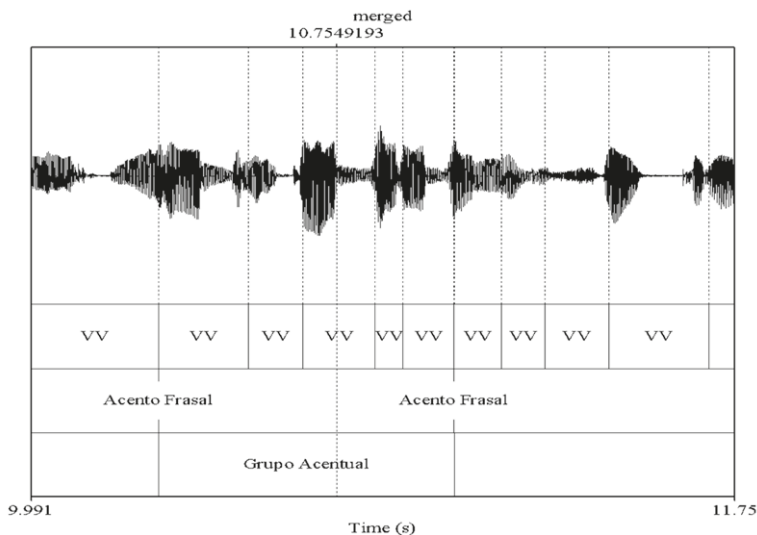
A terceira técnica é fundamentada apenas em dados de duração produzida e, segundo o autor, é perfeitamente adequada para o estudo da produção da organização temporal do enunciado. Essa técnica possui três etapas automáticas,

mas para isso é necessário um procedimento de segmentação e etiquetagem de unidades VV, o que justifica o uso do *BeatExtractor* nesta pesquisa.

As etapas automáticas correspondem a: a) normalização da duração das unidades VV, que consiste em obter um valor de *z-score*;<sup>6</sup> b) suavização, que, de acordo com Barbosa (2006), consiste em atenuar as variações locais de duração advindas da queda de duração em unidades VV pós-tônicas e/ou de durações de fones muito distintos da relação de durações dos fones do português brasileiro; e c) cálculo dos picos de máximo do *z-score* estendido suavizado através da detecção das posições de transição de derivada positiva para negativa (ponto de máximo).

São essas três etapas que foram implementadas no *script SG Detector*, desenvolvido por Barbosa (2006), para segmentar automaticamente o enunciado em grupos acentuais. O *SG Detector* também gera uma camada com a segmentação dos grupos acentuais no *Praat*, conforme a Figura 6:

Figura 6 – Unidades VV, acentos frasais e grupos acentuais.



Descritos os passos tomados para a obtenção dos dados da pesquisa, têm-se, por fim, duas segmentações, uma obtida através do julgamento de informantes

<sup>6</sup> Essa normalização, entretanto, corresponde a uma minimização dos efeitos da duração intrínseca (um segmento pode durar mais porque é simplesmente intrinsecamente mais longo) e da variabilidade da duração (quando um segmento dura mais simplesmente porque é mais variável).

sobre a intenção do falante em terminar uma unidade comunicativa, isto é, uma segmentação baseada em intenções, e outra obtida através do uso dos *scripts* desenvolvidos por Barbosa (2006), o *BeatExtractor* e o *SG Detector*. A segmentação com o uso dos *scripts* é de caráter estritamente fonético, ou seja, apenas elementos fonéticos serão considerados para a segmentação semiautomática. Essa segmentação é de grande relevância para estudos que pretendem analisar, a partir de uma segmentação automática da narrativa, elementos prosódicos associados a essa segmentação.

## 5.2 Unidades VV

Como explicado em seções anteriores, para proceder à segmentação das narrativas em grupos acentuais utilizando o *SG Detector*, fez-se necessário segmentá-las em unidades VV. Portanto, cabe aqui uma breve explicação sobre as unidades VV e o *script* utilizado para fazer a segmentação.

Para Barbosa (2021), a unidade VV (*vowels onset*) é a unidade mínima da programação rítmica da fala, a partir dela se organiza o grupo acentual. Essa unidade é do tamanho da sílaba e contempla a sequência de intervalos a partir do *onset* da vogal até o *onset* da vogal imediatamente seguinte, incluindo todas as consoantes e glides entre as duas. Estudos como os de Pompino-Marschall (1991) e Barbosa (1996) mostram a importância da transição CV para a produção e percepção da fala. Devido à necessidade de manter estável a sequência de transição CV, há uma tendência de preservação da duração das unidades entre as transições (unidades VV). Com o intuito de manter a estrutura rítmica e entoacional do enunciado, essas unidades resistem à variação mais do que a sílaba, mantendo os demais parâmetros constantes. No que diz respeito à percepção, observando o modo como a fala é processada, utilizar a unidade VV seria metodologicamente mais adequado do que a sílaba, de acordo com estes estudos.

Para segmentar as narrativas em unidades VV, é preciso utilizar, como dito anteriormente, o *script BeatExtractor*. Esse *script* foi adaptado por Barbosa (2004) para uso no *Praat* e está disponível na forma de programa aberto, via solicitação ao autor. O *script* detecta automaticamente *onsets* de vogal, através de um algoritmo sugerido por Scott (1993), a partir de quatro etapas que, de acordo com Barbosa (2004, p. 25), são as seguintes:

- filtragem do sinal de fala pelo uso de um filtro passa-faixas com valores default das frequências de corte para um locutor masculino fixadas entre 1000 e 1800 Hz. Os valores default para um locutor feminino são 1150 e 2100 Hz. Essa filtragem preserva a energia na faixa entre o primeiro e segundo formantes. O filtro

default é o filtro Butterworth de segunda ordem, mas a escolha por um filtro Hanning com faixas de transição mais abruptas do que a do Butterworth pode ser escolhida opcionalmente;

- retificação do sinal previamente filtrado;
- filtragem passa-baixas num valor default da frequência de corte de 20 Hz (preserva transições mais bruscas de amplitudes no PB – por exemplo, no caso de “tap” [ɾ] seguido de vogal);
- localização das fronteiras pela recuperação dos pontos que satisfazem duas restrições: (1) ponto de derivada máxima local positiva superior a 15% do valor máximo de derivada do sinal obtido após a aplicação da etapa anterior, e (2) ponto de amplitude superior a 12% da amplitude máxima obtida após a aplicação da etapa anterior.

Após a localização das fronteiras, elas foram gravadas em um arquivo do *Praat* para registro de marcação de fronteiras e rótulos de segmento, na extensão *TextGrid*. Durante a localização das unidades VV, no entanto, devido a ruídos do áudio, comentários do interlocutor etc., o *script* pode apresentar erros, por exemplo, a não marcação do começo ou término de uma unidade, ou pode ocorrer também a marcação equivocada. Estes erros, no caso desta pesquisa, foram corrigidos manualmente.

## 6. CONSIDERAÇÕES FINAIS

Recentemente tem havido um crescente interesse em investigar elementos prosódicos (pausa, duração, tons de fronteira, por exemplo) que são utilizados para delimitar macroestruturas discursivas em narrativas orais de diferentes gêneros. Esses elementos prosódicos contribuem para a identificação semântica das unidades discursivas, assim como explicitam quais intenções do falante para com sua audiência, desse modo, o processo de comunicação é facilitado.

Estudos que envolvem aspectos prosódicos da fala devem considerar não apenas a parte da produção, mas sobretudo a percepção de tal fenômeno. Neste capítulo, apresentamos um guia para pesquisas de percepção da estrutura do discurso narrativo, assim como também para fazer a segmentação de narrativas em unidades VV. Utilizamos a metodologia aqui descrita para coletar dados para nossa pesquisa de mestrado sobre a relação entre produção e percepção de pistas prosódicas na segmentação de narrativas espontâneas apresentada ao Programa de Pós-Graduação em Linguística e Literatura da Universidade Federal de Alagoas, em 2017.

## REFERÊNCIAS

- BARBOSA, P. A. *Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação*. Revista Estudos Linguísticos, Belo Horizonte, v.20, n.1, p.11- 27, jan./jun. 2012.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas, SP: Pontes ed.: FAPESP. 2006.
- BRUGMAN, H.; RUSSEL, A. *Annotating Multi-media / Multi-modal resources with ELAN*. In: *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC 2004)*. Paris: European Language Resources Association. 2004. pp. 2065-2068.
- LABOV, W. *The transformation of experience in narrative syntax*. In: LABOV, W. (ed.) *Language in the inner City*. Philadelphia: University of Pennsylvania Press, 1972. p. 219-247.
- LUCENTE, Luciana. *Aspectos dinâmicos de fala e da entonação no português brasileiro*. 2012. 204f. Tese (Doutorado em Letras). Universidade Estadual de Campinas, Campinas, 2012.
- OLIVEIRA JR., M. *Prosodic Features in Spontaneous Narratives*. 2000. 291 fl. Tese (Doutorado em Linguística) – Simon Fraser University, Vancouver, 2000.
- OLIVEIRA JR., M. *O uso do tom de limite como marca de segmentação da narrativa espontânea*. In *Revista Gelne*, vol.4, N 1, 2002. Disponível em: [http://www.gelne.ufc.br/revista\\_ano4\\_no1\\_13.pdf](http://www.gelne.ufc.br/revista_ano4_no1_13.pdf). Acesso em: 25 jul. 2013.
- OLIVEIRA JR., M. *Percepção dos elementos prosódicos na narrativa espontânea. Projeto de pesquisa*. Faculdade de Letras, Universidade Federal de Alagoas, 2010.
- OLIVEIRA JR., M.; CRUZ, R.; SILVA, E. W. *A relação entre a prosódia e a estrutura de narrativas espontâneas: um estudo perceptual*. Revista Diadorim, Rio de Janeiro, vol. 12, p. 39-50, dez. 2012.
- PASSONNEAU, R. J.; LITMAN, D. J. *Intention-based segmentation: Human reliability and correlation with linguistic cues*. In: *Proceedings of the 31st annual meeting of the association for computational linguistics (acl-93)*, 1993, Columbus, Ohio. Columbus, Ohio, 1993. p. 148-155.

POMPINO-MARSCHALL, B. *The syllable as a prosodic unit and the so-called P-centre effect*. Munique, Alemanha, 1991.

SILVA, E. W.; OLIVEIRA JR., M. *A percepção dos elementos prosódicos como marca de estruturação de narrativas espontâneas*. In: *RESUMOS DO III COLÓQUIO BRASILEIRO DE PROSÓDIA DA FALA*, 2011, Belo Horizonte. Belo Horizonte, 2011. p. 15-17.

SILVA, E. W. *A relação entre produção e percepção de pistas prosódicas na segmentação de narrativas espontâneas*. 2017. 79 f. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Programa de Pós-Graduação em Letras e Linguística, Universidade Federal de Alagoas, Maceió, 2017.

WOLFSON, N. *Speech events and natural speech*. *Language in Society*, Cambridge, n.5, p.189-209, 1979.

## UMA BREVE INTRODUÇÃO AOS ASPECTOS TEÓRICO-METODOLÓGICOS RELACIONADOS À COMPARAÇÃO FORENSE DE LOCUTOR

*Julio Cesar Cavalcanti<sup>1</sup>*

### 1. INTRODUÇÃO

Neste capítulo são abordados aspectos gerais referentes à tarefa de comparação forense de locutor, a partir de uma perspectiva fonético-acústica. Preceitos teórico-metodológicos relevantes para a prática e pesquisa na área são discutidos, de modo a fornecer ao leitor um breve panorama acerca dos diferentes fatores, internos e externos aos locutores, que merecem atenção quando à caracterização da identidade fonético-acústica de um falante se põe em foco.

Diferentes variáveis, de diferentes ordens, a exemplo da influência do canal de transmissão, das degradações de áudio, do estilo de fala e do próprio conceito de espontaneidade na fala são comentadas e critérios teóricos para a seleção de parâmetros fonético-acústicos são apresentados a partir da literatura especializada.

### 2. A FONÉTICA FORENSE

É possível pensar a fonética forense enquanto uma subárea da fonética aplicada inserida na intersecção entre as ciências da fala e as ciências de natureza forense, o que atribui a esse campo uma característica ou um traço híbrido.

---

<sup>1</sup> Universidade Estadual de Campinas (UNICAMP).



Em se tratando de definições, algumas tentativas de descrever o que é esse campo de estudo foram feitas por alguns autores, dentre eles, Jessen (2008), o qual descreve a fonética forense como sendo a aplicação do conhecimento, teorias e métodos da fonética geral para tarefas práticas que surgem em um contexto específico de investigação, bem como no desenvolvimento de saberes e métodos próprios do campo de investigação fonético-forense. Em termos mais abrangentes, a fonética forense poderia ser também descrita como a aplicação do conhecimento fonético geral a problemas de natureza legal (SAN SEGUNDO, 2014).

Para além da “diferenciação” ou “caracterização” forense de um locutor, a fonética forense inclui tarefas como a descrição do perfil de falantes, a partir da análise de características da voz e da fala, adjacientemente à análise de indexadores sociolinguísticos, a exemplo do sexo, idade, região de origem, nível social, interpretados a partir do comportamento linguístico dos locutores. Inclui também a identificação do conteúdo linguístico, processo denominado de “decodificação” do sinal de fala – a fim de determinar o que foi dito quando as gravações são de má qualidade ou apresentam algum tipo de degradação, e a autenticação de gravações, de modo a determinar se um registro foi ou não intencionalmente adulterado (ROSE, 2002; HOLLIEN, 2012). Outras tarefas podem compreender o tratamento do sinal de fala, como na aplicação de filtros de ruído, de modo a tornar a informação acústica inteligível e avaliável do ponto de vista fonético, tarefa que requer um conhecimento bastante aprofundado por parte do perito, de modo a evitar que informações linguísticas relevantes sejam eliminadas junto com o conteúdo indesejável (HOLLIEN, 2012).

Conforme menciona Rose (2002), dentre as atividades desempenhadas por um foneticista forense, a tarefa mais comum é aquela que envolve a comparação de uma ou mais amostras de fala desconhecida, denominada de amostra de “fala questionada”, com uma ou mais amostras de origem conhecida, a amostra de “fala padrão”.

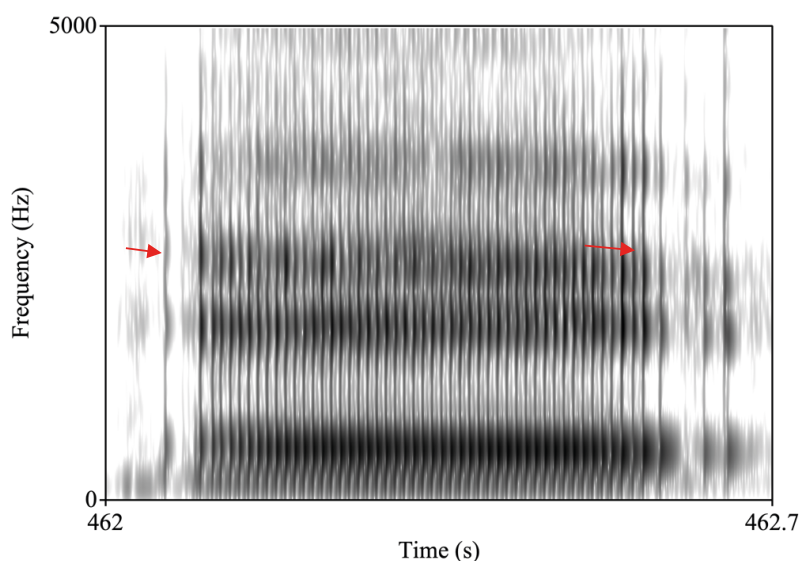
De acordo com Morrison (2009:229), o principal objetivo da comparação entre essas amostras é responder à importante questão: “quão semelhantes são as propriedades observadas nas amostras de fala padrão e questionada, sob a hipótese de que a amostra de fala questionada tenha a mesma origem que a amostra de fala padrão, em contraposição à hipótese de que estas tenham uma origem diferente?”.

A fim de responder a essa pergunta os foneticistas forenses lançam mão de análises baseadas em características distintivas ou idiossincráticas de um locutor, como a análise da qualidade vocal, padrões articulatórios na produção de vogais,

ditongos e consoantes, padrões temporais, a exemplo da duração dos segmentos, taxas de fala, entre outros parâmetros potencialmente discriminatórios.

Conforme menciona Rose (2002, p. 5), as análises comparativas são sempre realizadas em relação às unidades linguísticas, como as citadas anteriormente. Pode-se observar numa análise, por exemplo, que duas vogais idênticas realizadas em um mesmo contexto são produzidas de forma distinta por dois locutores, ou que as consoantes /k/ e /g/ são idiossincraticamente produzidas por um desses indivíduos, em decorrência de algum processo fonológico. Neste caso, o perito ou pesquisador deve ter um conhecimento abrangente sobre o sistema fonológico e inventário fonético da língua em questão, sobre a dinâmica de produção dos gestos articulatórios e as possíveis formas de se descrever as diferentes realizações. Entender a estrutura da língua possibilita a verificação de quando as realizações são ou não previsíveis do ponto de vista linguístico ou de quando essas estão relacionadas ao comportamento particular de um locutor, sendo, portanto, variáveis. Exemplo de um fenômeno prosódico idiossincrático é apresentado na Figura 1.

Figura 1 – Efeitos no espectrograma do evento de laringalização no início e final de uma pausa preenchida (indicado pelas setas vermelhas) com a vogal /ε/ no Português Brasileiro, por parte de um falante do sexo masculino durante a fala espontânea. O fenômeno é considerado como de caráter idiossincrático no PB, não apresentando *status* contrastivo no sistema fonológico da língua, e portanto, com frequência de ocorrência passível de variação entre os locutores. Ver Cavalcanti; Lucente; Barbosa (2018).



Em se tratando da produção de segmentos na fala, um dos fenômenos de grande relevância para a fonética forense é o fenômeno da “coarticulação”. Conforme comenta Hollien (1990:68-69), nenhum gesto articulatório é independente daqueles que o precedem ou sucedem, ao contrário disso, há uma sobreposição entre movimentos adjacentes em qualquer sequência de sons falados. De acordo com o autor, esse fenômeno é resultado de pelo menos dois de muitos fatores relacionados à produção oral humana: a taxa com que falamos e a interação entre as estruturas móveis (músculos) no processo de produção. Essas dimensões têm relação direta com aquilo que denominamos de plasticidade vocal, ou em termos mais amplos, plasticidade fonoarticulatória.

A plasticidade vocal implica a suposição de que um traço orgânico pode ser impresso de forma variável no sinal acústico, no que diz respeito à produção interna de um mesmo indivíduo e entre diferentes locutores, dado que estes normalmente diferem entre si no que diz respeito a características fonoarticulatórias, como tamanho e massa das pregas vocais, dimensões da laringe, língua, lábios etc., o que pode implicar modificações substanciais nas medidas acústicas. Existem, contudo, limites fisiológicos para os graus de variação, assim como tendências e padrões sistematicamente adotados pelos sujeitos (NOLAN, 1983). A análise de tais padrões ou daquilo que seria um comportamento particular de um locutor é um dos principais focos da fonética forense.

Rose (2012) chama a atenção para o fato de que os diferentes tipos de informações transmitidas por meio da voz/fala humana, a exemplo de demarcadores como sexo, idade, nível socioeconômico, escolaridade, não são sinalizadas em canais separados, mas são transmitidas juntamente com o conteúdo linguístico. O conhecimento de como esse processo se dá é fundamentalmente necessário para entender e interpretar a variação na fala.

### **3. QUESTÕES TERMINOLÓGICAS**

A fim de fornecer aos leitores um panorama geral sobre a prática de comparação forense de locutor, alguns comentários acerca da terminologia empregada na área se fazem necessários. Uma distinção clássica que merece atenção diz respeito à diferença entre as nomenclaturas “identificação” e “verificação” de locutor.

Conforme menciona Hollien (2012), a identificação de locutor é um processo em que são feitas tentativas de se identificar um indivíduo por meio da fala e da voz quando a identidade dessa pessoa não é conhecida, dentro de um

conjunto de vozes/falas numa determinada população. Essa tarefa, embora de grande importância, é bastante complexa, uma vez que todos os tipos de distorções e interferências possíveis em uma gravação podem estar presentes.

A verificação de locutor, por sua vez, se dá quando um locutor conhecido e cooperativo deseja ser reconhecido pela sua fala. Assim, amostras da fala/voz desse locutor são coletadas a fim de se construir um banco de referência. A identidade do indivíduo é então verificada quando uma nova amostra de fala é comparada aos dados obtidos previamente e a decisão é tomada: mesmo ou diferentes locutores? Enquanto as técnicas de verificação de falante são bastante úteis e exploradas pela indústria de tecnologia (como em sistemas de reconhecimento de fala, por exemplo), elas são utilizadas apenas ocasionalmente no contexto forense (HOLLIEN, 2012).

Há algum tempo, o emprego do termo “identificação de locutor” tem sido foco de discussões no campo da fonética forense. Atualmente, o termo vem sendo amplamente substituído pela nomenclatura “comparação” forense de locutor, para se referir à análise da autoria de registros de voz e de fala. Tal discussão tem por base razões propriamente teórico-terminológicas.

Para Morrison (2009), numa abordagem probabilística (*likelihood-ratio framework*<sup>2</sup>), o cientista forense não realiza a “identificação” de um locutor, porque esse termo implica determinar uma probabilidade posterior. Segundo o autor, um termo neutro e mais apropriado seria “comparação”. Nesse sentido, a comparação forense entre as amostras padrão e questionada não é uma tarefa que se resume a uma tomada de decisão binária (mesmo locutor vs. locutores diferentes), mas sim uma tarefa que busca determinar a força da evidência em relação a hipóteses de que as gravações têm ou não a mesma origem. Nesse caso, o que o cientista/perito forense tem como resposta é um valor de probabilidade e não a “verdade” expressa em termos de uma resposta categórica.

## 4. FATORES E VARIÁVEIS

A prática de comparação forense de locutor pode apresentar características bastante distintas quando realizada em condições controladas e em contextos sem controle. De acordo com Rose (2002), a falta de condições ideais no mundo

---

<sup>2</sup> *Likelihood-ratio framework* ou a abordagem baseada na razão de verossimilhança é um modelo conceitual que pode ser aplicado a suposições subjetivas baseadas na experiência, com vista à probabilidade de se ter observado uma determinada evidência dadas as hipóteses concorrentes (MORRISON, 2009).

real dificulta as análises forenses em decorrência de dois importantes fatores: a falta de controle sobre a variação e a redução nas dimensões de comparação entre amostras.

Conforme menciona Rose (2002), como geralmente não temos controle sobre as situações em que os dados forenses são coletados, esse se torna um grande desafio na avaliação das diferenças entre as amostras. Na grande maioria das vezes, o perito é levado a confrontar uma gravação de referência coletada num contexto quase ou completamente distinto do contexto de aquisição da amostra padrão, em decorrência da presença e/ou da ausência de variáveis típicas de cada condição.

Mais adiante comentaremos sobre alguns dos principais fatores e variáveis relacionadas à redução da dimensionalidade comparativa entre amostras de fala num contexto de comparação forense de locutor.

#### **4.1 Transmissão e qualidade da gravação**

De acordo com Rose (2002), uma das condições em que a dimensionalidade do contexto real na comparação de falantes é substancialmente reduzida diz respeito ao registro de fala telefônico. Isso se deve comumente a que aspectos da fala que poderiam ser úteis à comparação, características com considerável potencial distintivo, simplesmente não estão disponíveis porque foram filtradas, ou mesmo em decorrência de fatores práticos, como o ruído de fundo, má transmissão, falas sobrepostas, todos aspectos relacionados à degradação do sinal de áudio.

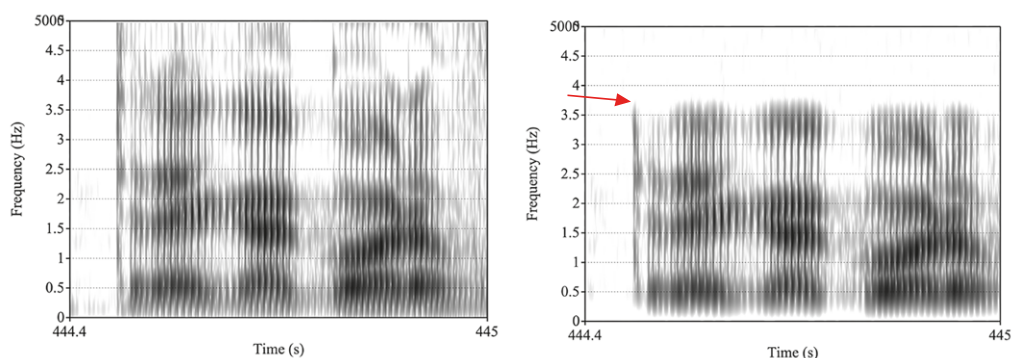
Conforme comenta Hollien (2012), as degradações podem ser resultantes da redução da largura de banda de frequência, ruído aditivo, redução do nível de energia acústica, distorção espectral ou harmônica, transmissão inadequada, transdutores de captação inadequados (telefones, gravadores), fatores ambientais, tais como o vento, atrito da roupa, interferências de outros falantes, música, entre outros.

Como mencionado anteriormente, esses problemas ocorrem basicamente porque, em geral, não podemos ter controle das situações nas quais as amostras de fala são registradas, o que se torna inevitavelmente um grande problema para os foneticistas forenses: tentar comparar amostras coletadas em condições distintas e que de algum modo incorporam marcas e efeitos diferentes no sinal acústico (ROSE, 2002).

De acordo com Künzel (2001), embora alguns dos distúrbios típicos, como altos níveis de ruído de fundo ou distorções em decorrência do uso de microfones

de carbono tenham se tornado um menor problema com o advento da telefonia digital moderna, outros fatores persistiram, como por exemplo os efeitos de passa-banda do canal de transmissão telefônico (350-3.400 Hz). A Figura 2 ilustra o efeito isolado do filtro de telefone sobre o espectrograma da fala. Na imagem é possível constatar a perda da informação acústica em torno da frequência de 3.500 Hz, como indicado pela seta.

Figura 2 – Efeito artificial do filtro de telefone na fala encadeada durante a frase: “(en)trega agora”, por um falante do sexo masculino. Espectrograma manipulado no *software* Praat (BOERSMA; WEENINK, 2021).



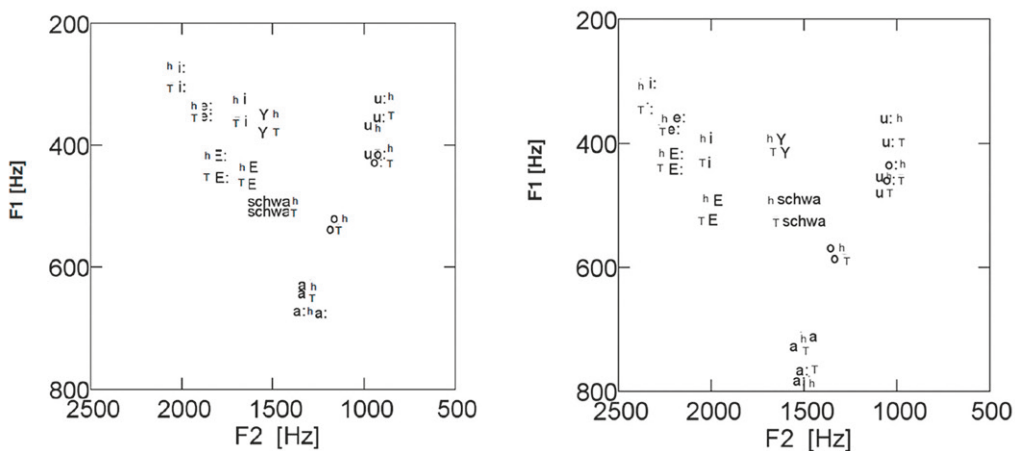
Um exemplo prático dos efeitos de telefones móveis no sinal de fala pode ser constatado no trabalho de Passetti (2015). Em um estudo com vogais orais do Português Brasileiro, a autora buscou determinar o grau de modificação fonético-acústica causado pelo filtro de banda do canal telefônico. Nesse estudo foram analisados dez locutores do sexo masculino, em gravações simultâneas nas condições de fala via celular e via direta. Os resultados desse trabalho revelaram que houve alterações nas frequências do primeiro (F1) e do terceiro formante (F3), em torno de 14% entre as duas condições testadas. No que diz respeito às frequências do segundo formante (F2), a transmissão telefônica influenciou de forma a aumentar as frequências de vogais com baixos valores de F2 e a diminuir as frequências de vogais com altos valores de F2, desencadeando o abaixamento global do espaço vocálico na gravação telefônica, pelo aumento nas frequências de F1, e o estreitamento do mesmo pela diminuição dos valores de F2. Apenas a *baseline*<sup>3</sup> e a duração interpícos de F0 não apresentaram diferenças significativas

<sup>3</sup> Caracterizado pelo valor de F0 mais típico de um locutor, correspondente aproximadamente ao sétimo quantil dos valores F0 da amostra (ARANTES; ERIKSSON, 2014).

no estudo. De acordo com a autora, as modificações nas disposições das vogais têm implicações perceptuais, uma vez que o abaixamento e redução do espaço vocálico fizeram com que as vogais se situassem proximamente a regiões centrais, podendo soar como mais abertas no telefone celular.

Numa investigação com vogais do alemão, a partir de gravações diretas e via telefone com dez indivíduos do sexo masculino e dez do sexo feminino, Künzel (2001), por sua vez, verificou que a medida de F1 foi a mais afetada dentre os parâmetros acústicos testados. As medidas centrais de F1 de cada vogal apresentaram valores mais altos nos dados transmitidos por telefone em comparação com os dados registrados de forma direta, corroborando a hipótese inicial do autor de que a atenuação dos componentes inferiores de um formante aumentaria o peso relativo dos componentes superiores, provocando um deslocamento artificial ascendente das frequências centrais. Nesse estudo, o efeito previsto em F1 ocorreu para todas as vogais testadas, exceto para a vogal /a/, para a qual, conforme menciona o autor, a medida apresenta um valor muito alto para ser afetado pelo efeito de passa-banda. Ver Figura 3.

Figura 3 – Espaço vocálico para falantes do sexo masculino (esquerda) e feminino (direita) para as condições de gravação direta (h) e telefônica (T).



Fonte: Künzel (2001:91)

Outra variável de considerável relevância e que geralmente acompanha a degradação causada pelo filtro telefônico é a presença de ruído. O ruído pode ser de diferentes naturezas, a exemplo do ruído ambiental, nos quais se enquadram os ruídos de vento, automóvel, trânsito, motores, equipamentos elétricos.

Eles podem ainda ser de natureza “linguageira”, quando há a sobreposição de falas, como uma conversa no fundo de uma gravação ou ligação telefônica. Nesse último caso, o impacto acústico é substancialmente elevado, uma vez que o ruído de fala pode representar modificações significativas nos parâmetros fonético-acústicos do sujeito foco da análise. O ruído pode ainda ser proveniente de mecanismos internos do sistema de captação de áudio, como o ruído interno de microfones, computadores, gravadores e celulares.

O trabalho de Constantini e Barbosa (2013) a partir da análise de diferentes parâmetros prosódicos em diferentes relação sinal-ruído, evidenciou que alguns parâmetros acústicos apresentaram mudanças em suas medidas após a adição de ruído gaussiano, dos quais: a ênfase espectral, mediana de frequência fundamental, taxa de saliência duracional, média, desvio-padrão e assimetria de z-score de duração de unidade VV. Destes, a ênfase espectral e a mediana da frequência fundamental foram os parâmetros mais afetados, tendo sido observado uma diferença mínima para os demais parâmetros. De acordo com os autores, a presença do ruído aditivo torna a extração dos parâmetros mais difícil por parte do software de análise (*praat*), repercutindo em erro nas medidas. Ver Tabela 1.

Tabela 1 – Médias de ênfase espectral e de mediana de frequência fundamental e magnitudes de mudança após adição de ruído gaussiano. Dados do trabalho de Constantini e Barbosa (2013:3)

	Gravação original	Ruído 0,01	Magnitude de mudança	Ruído 0,02	Magnitude de mudança
<b>M<sub>mediana f0</sub></b>	125 Hz	126 Hz	0,8%	128 Hz	2,4%
<b>M<sub>ênfase espectral</sub></b>	1.18 dB	1.8 dB	55%	3 dB	154%

De acordo com os autores, a análise da medida de ênfase espectral em gravações ruidosas é desencorajada, uma vez que a medida é considerada bastante sensível ao componente de ruído. A análise da estrutura rítmica do enunciado, no entanto, pode ser considerada mais consistente em situações em que se é preciso comparar amostras de fala em diferentes relações sinal-ruído. Isso se deve ao fato de que, por apresentarem maior energia acústica em relação aos demais segmentos da fala, os segmentos vocálicos tendem a ser preservados no espectrograma.



## 4.2 Estilo de fala

É um fato inegável o de que os falantes tendem a modificar os seus comportamentos linguísticos de acordo com os diferentes contextos comunicativos ou em função do interlocutor, o que reforça a natureza dinâmica da comunicação humana. Esses ajustes podem ocorrer tanto no nível lexical, a partir da seleção cuidadosa de palavras ou de um vocabulário específico, no nível sintático, refletindo uma maior ou menor produção de concordâncias, e até mesmo no plano fonético, a partir da modificação de parâmetros fonético-acústicos, a depender do estilo de fala adotado.

Um dos contextos que podem implicar modificações no comportamento fonético-acústico de um locutor diz respeito à condição de fala por telefone. O trabalho de Passetti (2018) sobre o estilo de fala telefônico com falantes do PB, verificou que, entre locutores do sexo masculino, o estilo de fala telefônico apresentou-se caracterizado pela modificação de descritores da frequência fundamental e de parâmetros temporais, a exemplo da distribuição dos picos de duração normalizada de sílabas fonéticas<sup>4</sup>, duração dos intervalos entre duas pausas silenciosas consecutivas e a duração de grupos acentuais, sendo essas modificações potencializadas em decorrência da presença de ruídos ambientais, com exceção da duração dos intervalos entre pausas silenciosas.

De acordo com a autora, em geral, o estilo de fala telefônico está associado a valores globais da frequência fundamental mais elevados, a uma maior variação entoacional, um maior grau de saliência duracional e uma maior quantidade de pausas silenciosas. Os intervalos entre pausas silenciosas tendem, por sua vez, a serem menores, sinalizando trechos de fala mais curtos na interação por telefone.

Outros fatores importantes dizem respeito ao grau de monitoramento dos falantes e aos efeitos das emoções na fala. Ao falar em público (exemplo de um registro de fala mais monitorado), articulações mais precisas tendem a ser produzidas em contraste com a fala espontânea. É comum que quando o falante se sente mais à vontade e confortável em uma conversa, o estilo de fala se torne mais casual, com palavras e frases simplificadas, ou mesmo utilizando um outro vocabulário. É nesse tipo de contexto de fala que reduções, assimilações e coarticulações são mais frequentemente e claramente produzidas. A taxa de fala tende a ser aumentada em comparação com a fala clara e, portanto, a quantidade de reduções e apagamentos também tende a aumentar (BJURSÄTER, 2004).

---

<sup>4</sup> Unidade constituída por uma vogal e todos os segmentos consonantais que a seguem, independentemente da fronteira silábica, até o *onset* da vogal seguinte, o qual determina o início da próxima unidade VV (BARBOSA, 2006).

Emoções podem resultar no aumento ou diminuição dos valores dos parâmetros acústicos, como na expressão de raiva, por exemplo, frequentemente relacionada a maiores taxas de fala, maiores valores de F0 e de intensidade, uma maior gama de variação de F0 e uma articulação mais tensa (MURRAY & ARNOTT, 1993). É importante estar atento, portanto, a essas possíveis modificações na fala relacionadas ao estado emocional dos sujeitos, especialmente em situações com altos níveis de estresse.

### 4.3 Sobre a espontaneidade na fala

Uma questão que interessa a linguistas, foneticistas forenses e foneticistas de um modo geral diz respeito ao impacto do grau de controle do experimentador no comportamento de fala dos sujeitos. Esse tópico também é tratado com bastante atenção por pesquisas em sociolinguística, as quais buscam elicitar o comportamento linguístico do sujeito o mais próximo daquilo que configura o uso real, de modo a reduzir ao máximo a intervenção do pesquisador sobre os dados obtidos – efeito classicamente referido no campo da sociolinguística como o “paradoxo do observador” (ver LABOV, 1972).

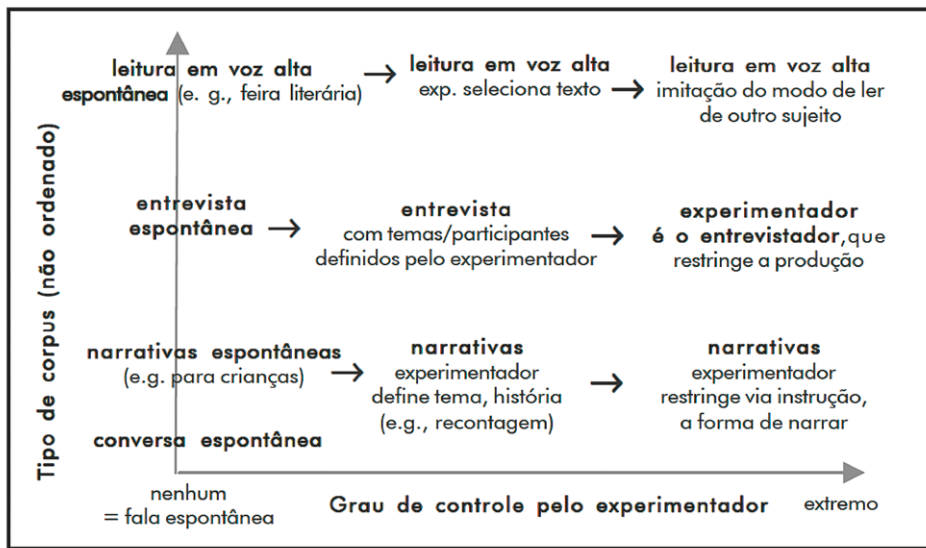
A ideia bastante recorrente de que a amostra de fala padrão, aquela coletada em um contexto de gravação controlado, apresenta uma redução significativa na dimensão de espontaneidade em relação à amostra de fala questionada, pode e deve ser discutida. Para tanto, alguns comentários acerca daquilo que se entende por fala espontânea, em detrimento de uma instância de fala dita não espontânea, devem ser tecidos.

De acordo com Barbosa e Madureira (2015), se considerarmos a questão da espontaneidade da fala, podemos pensar em um contínuo que vai da fala espontânea, sem intervenção ou controle por parte do pesquisador, à fala de laboratório, contexto no qual o pesquisador exerce algum tipo de controle sobre aquilo que deve ser produzido pelos sujeitos, como a seleção do gênero e do conteúdo a ser falado.

A Figura 4 representa uma proposta de classificação dos *corpora* de fala desenvolvida por Barbosa (2012), a partir de duas dimensões: o grau de controle do experimentador e o gênero (não ordenado). Neste modelo, é possível observar que conversas espontâneas, sem tópicos predefinidos, configuram o contexto com menor ou nenhum controle do experimentador correspondente ao eixo de espontaneidade, enquanto uma tarefa de imitação representa o oposto, maior nível de controle experimental por parte do pesquisador.

O gênero do tipo “entrevista”, com temas definidos, seria o gênero conversacional correspondente àquele de um interrogatório em um contexto forense, somado, é claro, a características particulares de uma situação investigativa, com uma possível acentuação do fator “estresse”.

Figura 4 – Proposta de classificação de corpora segundo eixo “grau de controle do experimentador” e “gênero ou tipo de corpus”.



Fonte: Barbosa (2012:16).

Como menciona Xu (2010), existe uma visão generalizada de que somente a partir da análise da fala espontânea é que podemos entender a natureza da fala cotidiana. Acompanhando essa visão está a ideia, cada vez mais difundida, de que a chamada fala de laboratório<sup>5</sup> é inadequada e não expressa ou esclarece a riqueza da fala espontânea, sendo, portanto, um registro pouco representativo do falante. De acordo com o autor, essa crença é baseada numa consideração limitada da literatura da área e uma apreciação inadequada de princípios científicos básicos. Em defesa da naturalidade da fala de laboratório o autor argumenta (XU, 2010:3-4):

This may be one of the most readily conceived characteristics of lab speech, because it

<sup>5</sup> Em uma definição ampla, a partir de Xu (2010), a fala de laboratório refere-se à fala gravada em laboratório, geralmente obedecendo um roteiro definido previamente. Contudo, o termo “fala de laboratório” é frequentemente usado para se referir a um discurso estereotipado

seems to contain an element of truth, i.e., scripted speech, by definition, is non-spontaneous. And non-spontaneous seems to be the opposite of natural. But it is important to first determine what is meant by “natural”. If naturalness is taken to mean reflecting human capabilities, all speech must be natural by definition. Indeed, even the most stereotyped lab speech is produced by real speakers. Regardless of what the speakers are asked to do in an experiment, their performance is based on their naturally acquired ability to speak, and is therefore a reflection of what they do everyday. It is not the case, for example, that they learn from the experimenter how to produce a vowel or a consonant or a tone, or they learn from the experimenter how to make an emphasis or ask questions. They already know how to do those things, and that’s why they are invited into the laboratory in the first place.

Conforme menciona o autor, se naturalidade significa refletir as capacidades humanas, todo registro de fala é por essência natural, uma vez que mesmo um registro de fala controlado é produzido por falantes reais e os seus desempenhos são resultado da capacidade de falar adquirida naturalmente. São por essas razões, inclusive, que os falantes são convidados para participar em pesquisas de laboratório.

De acordo com Barbosa (2012: 14), é preciso desvincular completamente os conceitos de fala espontânea e conversa livre, pois, segundo o autor, uma leitura, uma narrativa ou uma entrevista são de igual maneira espontâneas se ocorrerem em uma situação de comunicação dita “natural”, cabendo ao pesquisador apenas a seleção do material.

A redução nas dimensionalidades entre as amostras de fala padrão e questionada num contexto forense pode estar, muito mais provavelmente, relacionada ao quanto essas amostras são ou não representativas do sujeito em questão, ou seja, o quanto elas refletem do uso linguístico do falante, levando-se em conta fatores como os tamanhos e durações das amostras, os fatores de degradação mencionados anteriormente, a familiaridade entre os interlocutores, e os estilos de fala comparados, normalmente distintos no contexto de fala “padrão” e “questionada”.

Nesse sentido, a representatividade das amostras pode apresentar um impacto considerável na tarefa de comparação de locutor e na extração de medidas acústicas repercutindo na comparabilidade dos materiais coletados, como comentado a seguir.

---

como aquele em que o locutor apenas realiza a repetição de frases.

## 4.4 Sobre a representatividade dos dados

Por representatividade, pode-se entender o quão representativas são as observações reais do registro de voz e de fala em questão, levando-se em conta a sua origem. Nesse sentido, quanto mais representativos forem os dados, mais consistentes serão as estimativas, a favor ou contra a hipótese de uma origem comum (ROSE, 2012).

Num contexto forense real, por exemplo, é bastante comum que a amostra questionada apresente discrepâncias em termos de duração e extensão em comparação à amostra padrão, assim como diferenças no que diz respeito ao número de registros disponíveis, sendo essa última em geral mais extensa e numerosa, dada a possibilidade da realização de múltiplas coletas. Esse fator pode apresentar repercussões tanto na representatividade quanto na comparabilidade entre amostras ou mesmo limitar o potencial explicativo da análise. Nesse sentido, a depender do grau de assimetria entre as amostras, maiores ou menores são as chances de se realizar generalizações e de se chegar a conclusões gerais a respeito do perfil de um locutor.

Uma das principais questões relacionadas à tarefa de comparação de locutor diz respeito à quantidade de material de fala necessário e à duração mínima dos registros, de modo que esses incorporem dados suficientes de um falante e expressem uma gama de variabilidade significativa de parâmetros fonético-acústicos relacionados ao perfil locucional.

Algumas pesquisas têm se voltado para essa questão, como é o caso do trabalho de Arantes e Eriksson (2014). Os autores buscaram determinar, a partir de um método estatístico denominado “*change point analysis*”, o ponto de estabilização da variabilidade de medidas da frequência fundamental numa análise com diversas línguas. A razão para uma análise dessa natureza, com um considerável número de sistemas prosódicos (26 línguas, tonais e não tonais), tem como objetivo evitar que os dados obtidos estejam fortemente condicionados à influência de um conjunto de línguas em particular.

Os resultados desse experimento revelaram que a média e a mediana da frequência fundamental tendem a atingir a estabilidade com cerca de dez segundos de duração, enquanto o valor de base (*baseline*) leva em torno de cinco segundos, um intervalo de tempo ainda menor do que sugerido por trabalhos anteriores.

No que diz respeito a taxas de fala, a exemplo da taxa de elocução e taxa de articulação, os experimentos realizados por Arantes, Eriksson e Lima (2018) evidenciaram que, em geral para essas medidas, o tempo médio de estabilidade

é de 12,1 segundos, com a maior parte da distribuição entre 7,9 e 16,2 s. Os autores verificaram, no entanto, que houve um efeito significativo do tipo de taxa no tempo de estabilização, com a taxa de articulação apresentando um tempo de estabilização mais rápido. De acordo com Arantes, Eriksson e Lima (2018), é possível que os intervalos de estabilização ligeiramente mais longos na condição de taxa de elocução devam-se à inclusão das durações das pausas silenciosas, o que repercute no aumento da variabilidade de duração dos segmentos. Variáveis como o nível das taxas de fala (se rápido, normal ou lento) e o critério de ponto de estabilização (se levando em conta ou não o limiar perceptivo), também demonstraram influenciar significativamente o tempo de estabilização das medidas.

O tamanho e extensão das amostras de fala também podem apresentar impacto na capacidade de discriminação por parte dos ouvintes, como, por exemplo, em tarefas de reconhecimento de locutor, nas quais o ouvinte é solicitado a identificar o registro de um falante específico a partir de uma sequência de vozes (*vocal line-up*).

Os resultados obtidos por Decoster et al. (2001), por exemplo, revelam que há uma relação direta entre o tipo e duração do estímulo empregado e o desempenho dos ouvintes na tarefa de discriminação. Nesse caso, a performance dos sujeitos tende a ser tanto melhor quanto mais informações acústicas e prosódicas sobre o locutor estiverem contidas no estímulo, com um melhor desempenho para unidades linguísticas maiores, como sentenças, em detrimento de unidades linguísticas menores em extensão e duração, a exemplo de vogais. Tanto experimentos fonéticos gerais (KREIMAN; SIDTIS, 2011) quanto estudos que tratam de dados de fala/voz de locutores geneticamente relacionados (SAN SEGUNDO, 2014) confirmam a variação na performance dos ouvintes em decorrência desses fatores.

Ainda em se tratando da representatividade dos dados, é preciso levar em consideração a variável “contemporaneidade” das amostras, que diz respeito à janela de tempo que separa os registros das falas padrão e questionada. A não contemporaneidade entre amostras pode implicar diferenças nos parâmetros fonético-acústicos em função da variação intralocutor. Nesse sentido, é importante saber o que esperar em termos de modificação nos padrões de voz e de fala, especialmente se as amostras foram coletadas em janelas de tempo consideravelmente grandes.

Mudanças adotadas ou sofridas pelos locutores, como migrar para uma outra região (com diferente dialeto/língua) ou mesmo hábitos adquiridos ao longo da vida, como o tabagismo por exemplo – capaz de afetar clinicamente a qualidade

vocal, resultando numa diminuição da frequência fundamental e aumento dos valores da relação ruído-harmônico (NHR), *jitter* e *shimmer* (PINHO; CRESPO; MOURÃO, 2014) – são exemplos de fatores que devem ser considerados numa análise comparativa, além de mudanças na qualidade da voz em decorrência do envelhecimento (presbifonia).

## 5. CRITÉRIOS PARA A SELEÇÃO DE PARÂMETROS FONÉTICO-ACÚSTICOS

Tendo em vista todos os fatores comentados anteriormente e a suas implicações para a prática de comparação forense de locutor, faz-se necessário tecer alguns comentários sobre os critérios de seleção de parâmetros fonético-acústicos para a análise do perfil de falantes. Existem alguns requisitos teóricos básicos que precisam ser respeitados, de modo que a seleção de um determinado conjunto de medidas seja tanto coerente do ponto de vista da análise quanto sensível à variação entre sujeitos.

De acordo com Nolan (1983), um parâmetro candidato ao uso forense deve idealmente satisfazer seis premissas básicas, conforme apresentadas na Tabela 2. Os critérios propostos pelo autor representam uma tentativa de se estabelecer quais informações contidas no sinal acústico são altamente dependentes e representativas de um indivíduo, de modo a restringir aquelas de baixo poder discriminatório.

É preciso considerar no entanto que, na realidade prática a seleção de parâmetros é um processo altamente dependente das técnicas de análise disponíveis, o que não implica que critérios teóricos e práticos não possam ser propostos e discutidos. Além do fato de que é preciso ter em mente que os parâmetros elencados como discriminatórios para uma dupla de locutores podem não ser contrastivos para outra. Nesse sentido, a relevância de cada parâmetro precisa ser estudada caso a caso, conforme menciona San Segundo, (2014).

Tabela 2 – Critério para a seleção de parâmetros para a tarefa de comparação forense de locutores (CFL) de acordo com Nolan (1983:11)

<i>Alta variabilidade entre locutores</i>	O parâmetro deve apresentar um alto grau de variação entre locutores.
<i>Baixa variabilidade intralocutor</i>	Deve ser consistente em relação às produções de um mesmo indivíduo e preferencialmente resistente ao seu estado de saúde, condição emocional e ao contexto comunicativo.
<i>Resiste à tentativa de disfarce ou imitação</i>	Deve ser resistente às tentativas de disfarce ou imitação, seja por se tratar de uma característica fisiológica difícil de ser modificada ou por refletir um aspecto da fala que esteja abaixo do nível de percepção do falante.
<i>Resistente ao canal de transmissão</i>	A aplicabilidade de um parâmetro será limitada se a sua informação for passível de ser perdida ou reduzida em decorrência do efeito do canal de transmissão.
<i>Disponibilidade</i>	O parâmetro ou fenômeno deve apresentar uma frequência de ocorrência considerável, de forma a não requerer uma grande quantidade de material de fala, nem sempre disponível num contexto forense.
<i>Mensurabilidade</i>	A extração do parâmetro ou análise do fenômeno não devem ser demasiadamente complexas, de modo que inviabilize a mensuração.

Encontrar parâmetros que satisfaçam a esses critérios é um objetivo e um grande desafio para as pesquisas em fonética forense. Contudo, algumas estratégias podem ser adotadas no contexto experimental de modo a se avaliar o grau de consistência dos parâmetros fonético-acústicos para a tarefa de comparação de locutor e as suas resistências às variáveis presentes no contexto prático. Dentre algumas das estratégias mais comuns podemos citar:

- Gravações não contemporâneas: as gravações são realizadas respeitando-se um determinado intervalo de tempo entre si e posteriormente submetidas à comparação. Na prática forense a amostra de fala padrão é quase sempre coletada algum tempo após o incidente, sendo neste caso, posterior à amostra de fala padrão. Nesse sentido, pesquisas com coletas de caráter transversal são encorajadas.
- Análise de diferentes estilos de fala: diferentes estilos de fala são comparados, como no caso de uma gravação no contexto de um diálogo e um material de fala coletado a partir de entrevista. Em contextos forenses, o material de fala questionado é quase sempre confrontado com um material padrão coletado num estilo diferente, como no contexto de um interrogatório, por exemplo.



- Análise de diferentes estados/emoções: o estado psicológico de um falante pode implicar modificações nos parâmetros acústicos analisados. Sujeitos em situações espontâneas podem apresentar um diferente comportamento linguístico em comparação a situações em que estão sob efeito de um fator de estresse.
- Análise do efeito de diferentes tipos de degradação: a presença do ruído de fundo ambiental, um dos tipos mais comuns e frequentes de degradação, pode implicar erros no cálculo de medidas acústicas, principalmente os ruídos provenientes de outras falas (falas sobrepostas). O disfarce vocal, por sua vez, outro tipo de degradação, pode causar modificações substanciais nos parâmetros acústicos, como na qualidade vocal, no fluxo discursivo, no padrão entoacional, nas taxas e na intensidade dinâmica da fala, além de variações nos valores de frequência fundamental da voz (MATHUR, CHOUDHARY E VYAS, 2016). Entender o impacto desses tipos de degradações é fundamental para a tarefa de comparação de locutores num contexto de investigação. Ver Passetti; Barbosa (2018). Sobre os efeitos de outros fatores, a exemplo da compressão de áudio e do microfone nos parâmetros acústicos da voz, ver Cavalcanti et al., (2021).
- Gravações com diferentes canais de transmissão/captação: Nesse caso, diferentes tipos de gravação são realizadas, como a comparação de gravações via telefone, WhatsApp, a partir de gravadores e microfones não profissionais, sempre analisadas em relação à gravação em contextos ideais. Aqui, busca-se entender o impacto das diferentes formas de captação e transmissão do registro acústico na tarefa de comparação de locutor.
- Gravações com sujeitos relacionados: gravações realizadas com indivíduos que compartilham material genético em comum (gêmeos monozigóticos, gêmeos heterozigóticos, irmãos) ou compartilham o mesmo ambiente linguístico é uma das estratégias que pode ser adotada num contexto experimental para a análise do grau de variação entre locutores semelhantes e de seus possíveis limites. Outro fator importante a ser considerado diz respeito ao grau de familiaridade entre os locutores, que é quase sempre reduzido num contexto de interrogatório. Sobre estudos com locutores geneticamente relacionados ver Cavalcanti (2020), Cavalcanti, Eriksson, Barbosa (2021).

Algumas das estratégias supracitadas ajudam a aproximar o contexto experimental do contexto real em que se dão as análises forenses, minimizando possíveis vieses comuns e inerentes a essa primeira condição. Tais estratégias

podem ainda serem ou não combinadas, a depender do objetivo da investigação e do nível de proximidade ou distanciamento que se busca do contexto real em que se dão as comparações forenses de locutor.

## REFERÊNCIAS

ARANTES, P., ERIKSSON, A., LIMA, V. Minimum Sample Length for the Estimation of Long-term Speaking Rate. Proc. 9th International Conference on Speech Prosody, p. 661-665, 2018.

ARANTES, Pablo; ERIKSSON, Anders. Temporal stability of long-term measures of fundamental frequency. Dublin: ISCA, 2014. p. 1.149-1.152, 2014.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. Revista de Estudos da Linguagem, v. 20, n. 1, p. 11-27, junho, 2012.

BARBOSA, P.; MADUREIRA, S. Manual de Fonética Acústica Experimental: aplicações a dados do Português. São Paulo: Cortez, 2015.

BJURSÄTER, U. Speaking styles and Phonetic variation. Term paper autumn, Department of Linguistics, Stockholm University, 2004.

BOERSMA, PAUL & WEENINK, DAVID. Praat: doing phonetics by computer [Computer program], 2021. Version 6.1.39, retrieved 8 February 2021 from <http://www.praat.org/>.

CAVALCANTI, J. C.; ERIKSSON, A.; BARBOSA, P. A. Acoustic analysis of vowel formant frequencies in genetically-related and non-genetically related speakers with implications for forensic speaker comparison. Plos one, v. 16, n. 2, p. e0246645, 2021.

CAVALCANTI, J. C., ENGLERT, M., OLIVEIRA JR., M., & CONSTANTINI, A. C. Microphone and Audio Compression Effects on Acoustic Voice Analysis: A Pilot Study. Journal of Voice, 2021.

CAVALCANTI, J. C. Análise fonético-acústica em gêmeos idênticos: os limites da variação entre locutores. (151-172). In: BARBOSA *et al.*, Análise Fonético-Forense em tarefa de Comparação de Locutor. Campinas, SP: Millennium Editora, 2020.

CAVALCANTI, J. C. LUCENTE, L. BARBOSA, P. Laryngealization, Gender and Speakers' Distinctiveness in Brazilian Portuguese. Proceedings of the 9th International Conference on Speech Prosody, Poznań, Poland, 2018.

CONSTANTINI, A. C. BARBOSA, P. A. Análise de parâmetros prosódicos em diferentes relações sinal-ruído. Anais do Colóquio Brasileiro de Prosódia da Fala, v. 2, Brasília, 2013.

DECOSTER, W. VAN GYSEL A. VERCAMMEN W.D, DEBRUYNE J, F. Voice similarity in identical twins. Acta OtoRhinoLaryngologica Belgica, 2001.

HOLLIEN, H. About forensic phonetics. Linguistica, Vol 52 No 1: Spoken discourse, 27-53, 2012.

HOLLIEN, H. The acoustics of crime. New York: Plenum Press, 1990.

JESSEN, M. Forensic phonetics. Language and Linguistics Compass, 2(4), 671-711, 2008.

KREIMAN, J. SIDTIS, D. Voices and listeners: Toward a model of voice perception. Acoustics Today, p.7-15, 2011.

KÜNZEL, H. J. Beware of the 'telephone effect': the influence of telephone transmissions on the measurement of formant frequencies, Forensic Linguistics, 8 (1), 80-99, 2001.

LABOV, William. Padrões sociolinguísticos. São Paulo: Parábola, [1972] 2008.

MATHUR S, CHOUDHARY BSK, VYAS CJM. Effect of Disguise on Fundamental Frequency of Voice. J Forensic Res 7:327. 2016.

MORRISON, G. S. Forensic voice comparison and the paradigm shift. Science & Justice, Volume 49, Issue 4, p. 298-308, 2009.

MURRAY, I.R. & ARNOTT, J.L. Toward the simulation of emotion in synthetic speech: A review of the literature on vocal emotion. Journal of the Acoustical Society of America 93 (2), 1097-1107, 1993.

NOLAN, F. The phonetic bases of speaker recognition. Cambridge: Cambridge University Press, 1983.

PASSETTI, R. R.; BARBOSA, P. A. Análise fonético-acústica de disfarces vocais com aplicações para a fonética forense. Cadernos de Estudos Linguísticos, v. 60, n. 3, p. 567-583, 22 nov. 2018.

PASSETTI, R. R. “O efeito do telefone celular no sinal da fala: uma análise fonético-acústica com implicações para a verificação de locutor em português brasileiro”. Dissertação de mestrado. Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas SP, 2015.

PASSETTI, R. R. Estudo acústico-perceptual do estilo de fala telefônico com implicações para a verificação de locutor em português brasileiro. Tese de doutorado, Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP, 2018.

PINTO, Aline Gomes Lustosa; CRESPO, Agrício Nubiato; MOURAO, Lucia Figueiredo. Influência do tabagismo isolado e associado a aspectos multifatoriais nos parâmetros acústicos vocais. *Braz. j. otorhinolaryngol.*, São Paulo, v. 80, n. 1, p. 60-67, Feb. 2014.

ROSE, P. *Forensic speaker identification*. London: Taylor & Francis, 2002.

SAN SEGUNDO, E. *Forensic speaker comparison of Spanish twins and non-twin siblings*, PhD dissertation, Menéndez Pelayo International University & CSIC, 2014.

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



# RELAÇÕES DE SENTIDO EM CONSTRUÇÕES GRAMATICAIS HOMOMÍMIA, POLISSEMIA E SINONÍMIA EM *TIRAS* DA MAFALDA

*Eliane Barbosa da Silva*<sup>1</sup>

### RESUMO

Através deste estudo, analisamos as relações semânticas em construções gramaticais da língua espanhola, como língua estrangeira (ELE), observando os efeitos de sentidos originados no contexto de produção da língua em questão. Buscamos fazer uma análise dessas relações de sentidos em textos de *tiras* de Mafalda, os quais constituíram o *corpus* da pesquisa, a fim de descrever e explicar os usos das noções de sinonímia, polissemia e homonímia, além de outras relações semânticas e pragmático-discursivas, como os implícitos, as implicaturas, a pressuposição, presentes nas construções gramaticais de tais textos, visto serem fenômenos tão usuais e corriqueiros tanto na escrita como na oralidade nas línguas, embora, ao que parece, sejam pouco estudadas no processo de ensino-aprendizagem da língua estrangeira ou mesmo no ensino da língua materna, fato que se observa na pouca atenção ou ausência de tratamento aos casos nos próprios manuais didáticos de ensino de línguas. Para isso, teremos como base teórica e conceitual a semântica, visando especialmente uma discussão voltada ao estudo do sentido das construções gramaticais, dentro da perspectiva da

---

<sup>1</sup> Universidade Federal de Alagoas.

semântica formal e da semântica linguística, com base em Ilari (2001; 2002; 2006), Cançado (2008; 2012), Henriques (2011), Hjelmslev (1966), como também da pragmática, com base em Grice (1957; 1975), Ducrot (2005), além de outros teóricos que tratam dessa temática.

**Palavras-chave:** Semântica; Relações de sentido; Língua espanhola. Tiras.

## 1. INTRODUÇÃO

Nos últimos anos, tem sido significativa a atenção que vem sendo dada aos questionamentos sobre ensino-aprendizagem de língua estrangeira (LE) e língua materna (LM). Diante dessa realidade, ressurgiu também uma discussão ainda mais peculiar: o(s) problema(s) que envolve(m) o processo de ensino-aprendizagem de línguas próximas em relação a aspectos semântico-linguísticos. Observa-se, contudo, que ainda é bastante incipiente o tratamento dado às questões semânticas, especialmente na sua aplicabilidade metodológica e pedagógica nos manuais didáticos, visando uma abordagem mais produtiva em sala de aula, problema que precisa ser enfrentado quando se trata do ensino da LM, e também quando lidamos com o ensino da LE.

Diante desse fato, propomo-nos a refletir sobre aspectos linguísticos, voltando-nos basicamente para a observação de aspectos semânticos, os quais instigam questionamentos sobre as relações de sentido que podem estar presentes em textos do gênero *tiras cômicas* de Mafalda, e analisar as relações semânticas de sinonímia, polissemia e homonímia, como também as relações pragmático-discursivas, como os implícitos, as implicaturas, a pressuposição nas construções gramaticais desses textos, a fim de observar o sentido que as palavras e os textos promovem em determinados contextos de uso.

A bibliografia disponível sobre essa temática no campo teórico é vasta e, no entanto, parece ser ainda um tanto escassa do ponto de vista de sua aplicabilidade pedagógica, pois os manuais didáticos, como as gramáticas, livros didáticos e dicionários, quando tratam do assunto apresentam, via de regra, descrições sumárias e exemplos isolados, preocupando-se em fornecer apenas conceitos básicos, glossários, ou frases feitas, deixando à margem, nesse caso, questões relacionadas à descrição e análise, à frequência e às condições de uso e o contexto em que são empregados. Nosso objetivo, portanto, é fazer uma análise descritiva e explicativa a respeito dessas relações de sentido em construções gramaticais presentes nas *tiras* de Mafalda, descrevendo essas noções e exemplificando os

seus usos, procurando, assim, contribuir para que essas noções tenham aplicabilidade no ensino através da leitura e compreensão de textos.

Para realizarmos a pesquisa, traçamos como objetivo estudar as relações de sentido em construções gramaticais do espanhol (ELE), presentes em textos do gênero *tiras* da Mafalda, a partir de princípios teórico-metodológicos da Semântica Linguística e da Semântica Formal, que buscam a descrição e a análise de tais fenômenos. Para tanto, selecionamos textos em *tiras* de Mafalda nos quais estejam presentes construções gramaticais com as seguintes relações de sentido: sinonímia, polissemia e homonímia e outras relações pragmático-discursivas; analisamos tais relações de sentido, descrevendo a sua ocorrência e uso em determinados contextos nessa língua; também refletimos sobre a importância do estudo dessas relações de sentido na formação do professor-aprendiz de ELE, discutindo os aspectos teóricos e práticos a esse respeito.

Quanto à metodologia, buscamos identificar tais relações semânticas e pragmático-discursivas nas construções gramaticais em espanhol por sabermos que estas ocorrem com certa frequência em diversos tipos de textos orais e escritos nas línguas. Além disso, observamos que os textos de materiais didáticos, assim como outros tipos e gêneros textuais, como as tiras, também apresentam essas relações de sentidos por serem amostras (quase) reais<sup>2</sup> da língua. Tais fenômenos semântico-linguísticos e pragmático-discursivos nos motivam a observar, identificar e analisar a sua ocorrência na língua espanhola, e também contribuir para a sua aplicabilidade no ensino, especialmente no Curso de Letras – Espanhol, área em que atuamos, na Universidade Federal de Alagoas/UFAL.

Os dados foram selecionados e apresentados graficamente por FIGURAS enumeradas, em formato PDF, conforme disponíveis na internet.<sup>3</sup> Os textos que compõem o *corpus* para análise são compostos por tiras de Mafalda escritas a partir da década de 1960. Constituímos um *corpus* de 30 (trinta) sequências de tiras de Mafalda, nas quais sempre é possível observar essas relações de sentido, visto constituírem amostras (quase) reais da língua e também por estarem, de certa forma, acessíveis aos estudantes de espanhol como língua estrangeira e aos próprios falantes da língua. Das trinta sequências de tiras de Mafalda, 9 (nove) delas são selecionadas para serem analisadas no bojo da discussão teórica e as demais 21 (vinte e uma) fazem parte de um glossário do trabalho final, que se configura como um aporte didático para o ensino da língua espanhola no tocante a essa temática.

---

<sup>2</sup> Não se trata de amostras reais de língua porque são textos construídos pelo autor. Mas podem ser ditos «quase reais» porque são diálogos que se aproximam muito da fala.

<sup>3</sup> As fontes ou referências das tiras selecionadas são citadas logo abaixo de cada figura.



A coleta de dados foi feita a partir da identificação dos fenômenos nas construções gramaticais dos textos pré-selecionados, as amostras identificadas no parágrafo anterior. Posteriormente, tais sentenças foram classificadas como exemplificação de acordo com as relações de sentido que estabelecem, seguido das análises das relações de sinonímia, polissemia, homonímia, implícitos, implicaturas e pressuposição, de acordo com a fundamentação teórica sobre a qual nos debruçamos para descrever e explicar tais fenômenos.

## 2. AS TIRAS DE MAFALDA COMO OBJETO DE ESTUDO E APRENDIZAGEM

As muitas palavras análogas, assim como algumas construções gramaticais, por exemplo, podem apresentar muitas semelhanças, como também grandes divergências quanto ao sentido na língua. Isto é, inúmeras palavras, expressões e diversas sentenças da língua, no caso da língua espanhola, são iguais ou semelhantes tanto no plano da expressão quanto no plano do conteúdo. Muitas outras, porém, são iguais ou semelhantes no plano da expressão (substância da expressão) e diferentes no plano do conteúdo (forma do conteúdo).

O espanhol, como as demais línguas naturais, apresenta, em suas construções gramaticais, determinadas relações de sentido no contexto de produção em que elas ocorrem, tais como as relações de sinonímia, polissemia e homonímia, implícitos, implicaturas, pressuposição. Essas construções nos mostram a capacidade que todo falante tem de produzir, de compreender e/ou de refletir sobre determinados aspectos semânticos, pragmáticos e discursivos em sentenças da sua própria língua materna e/ou de uma determinada língua estrangeira.

A escolha pelo tema partiu de nossa reflexão, enquanto professora de língua estrangeira e de língua materna, sobre determinadas relações de sentido em sentenças, empiricamente observáveis na fala de sujeitos falantes de português (LM) e de espanhol (LE), como também nas construções gramaticais presentes em textos do gênero *tiras* de Mafalda. Essas produções revelam o conhecimento que o falante ou aprendiz tem sobre o significado das construções gramaticais da sua língua ou da língua alvo e como ele as compreende no que tange às relações de sentido que se estabelecem em determinadas sentenças. Nesse sentido, observamos também se tais relações de sentido podem interferir, de certa forma, na aprendizagem da língua estrangeira, no caso, o espanhol, ou seja, se há transferência de sentido da língua materna para língua estrangeira, visto serem línguas próximas, analisando, portanto, se os sentidos ou mesmo a expressão são iguais ou diferentes entre as línguas.

É comum observar essas construções em livros teóricos, principalmente, ou em manuais didáticos, como casos exemplares e modelos formais apresentados para a explicação e descrição de tais relações. No entanto, esses exemplos se apresentam como clichês e modelos pré-elaborados e adotados para simplesmente exemplificar ou caracterizar essas relações de sentido, em alguns casos, sem uma reflexão ou análise sobre o seu uso em textos ou nas reais condições de produção do falante, por exemplo. Nossa intenção, portanto, é analisar esses dados em *tiras* de Mafalda, a fim de descrever e explicar tais fenômenos a partir da sua presença em contextos reais de produção e, assim, contribuir para que o estudo dessas relações de sentidos (homonímia, sinonímia e polissemia), assim como os implícitos, as implicaturas e a pressuposição sejam estudadas e aplicadas em aula de línguas materna ou estrangeira.

Diante disso, observamos que os textos selecionados das *tiras* de Mafalda revelam muito mais do que as relações de sentido, como sinonímia, polissemia e homonímia. É necessário ir além das questões semântico-linguísticas para compreender as relações discursivas e pragmáticas que se apresentam nesse gênero textual e em seu contexto de uso. Partimos, portanto, de uma análise de aspectos mais especificamente semânticos, como a relação do sentido nas construções gramaticais, porém observamos que tais relações não se esgotam em si mesmas, pois os textos revelam elementos para além do estritamente linguístico, direcionando-nos aos aspectos pragmáticos e discursivos marcantes nesse gênero textual.

### 3. OBSERVANDO ALGUMAS TIRAS DE MAFALDA

Iniciamos essa discussão com base na concepção de signo e contexto. Qual a significação ou a função do signo “sopa” no contexto da tirinha da Figura 1? Podemos analisar questões de significação a partir do signo ou do signo ao contexto e enunciado?

Falando sobre a palavra signo, Hjelmslev (2013, p. 49) afirma que ‘signo’ se define por uma função. Para ele, um ‘signo’ funciona, designa, significa, é portador, portanto, de uma significação. Assim, ele assume, visto que a proposta inicial é de Saussure, que a linguagem é um sistema de ‘signos’. O autor ainda afirma que grandezas como as frases, as proposições e as palavras parecem atender a essa condição: elas veiculam uma significação e, portanto, são ‘signos’. No entanto, ele diz que é interessante aprofundar essa análise o quanto possível para assegurar uma descrição exaustiva e simples. Toda grandeza, segundo Hjelmslev (Idem, p. 50), assim como todo signo, se define de modo

relativo e não absoluto, ou seja, essa definição se dá unicamente pelo lugar que ela ocupa no contexto. Não existem significações reconhecíveis outras que não as significações contextuais, pois:

Considerados isoladamente, signo algum tem significação. Toda significação de signo nasce de um contexto, quer entendamos por isso um contexto de situação ou um contexto explícito, o que vem a dar no mesmo; com efeito, num texto ilimitado ou produtivo (uma língua viva, por exemplo), um contexto situacional pode sempre ser tornado explícito (HJELMSLEV, 2013, p. 50).

Assim, é possível entender que determinado signo (“sopa”, por exemplo), isoladamente, tem uma significação relativa, limitada, ou sem significação, como afirma Hjelmslev (2013), enquanto que analisado no contexto situacional (da tirinha, no nosso caso) ele passa a ter outra função, outra significação. Observa-se que a espécie de sinonímia criada pelo sujeito que utiliza a linguagem no contexto situacional do texto a seguir traz um jogo com a linguagem figurada, tornando esse contexto explícito.

De que forma essa questão da significação, do contexto, do sentido figurado pode contribuir para o ensino-aprendizagem de uma segunda língua ou da língua estrangeira? Veremos isso *pari passo*, à medida que se analisem as tirinhas em língua espanhola, embora já se observe a relevância da temática para a compreensão de textos, e de fenômenos como a sinonímia, polissemia e homonímia para o ensino-aprendizagem tanto da língua materna como da língua estrangeira.

Figura 1



Fonte: <http://www.todohistorietas.com.ar/tiras2.htm> Consulta: 24/05/2018.

Na tirinha (Figura 1), observa-se como Mafalda faz um jogo de sentidos da palavra “sopa”, ou seja, o enunciador (Mafalda) cria uma espécie de sinonímia na linguagem ao atribuir outros sentidos a determinado signo. Inicialmente, ela faz

um elogio a sua mãe pela comida gostosa sem fazer menção ou sem saber que se tratava de sopa, comida de que ela não gosta. Em seguida, quando a mãe diz que é sopa, ela repreende com um “chst” e diz que não se fala “palavrões” na mesa. Nesse contexto, entendemos que ela cria uma sinonímia de sopa com palavrões (sopa = palavrões) na enunciação.

Em seguida, a mãe responde dizendo-lhe que sopa não é palavrão. Após observar o prato, a pequena Mafalda reforça a sua repulsa pelo prato de sopa, dizendo no último balão que também não se dizem “mentiras” na mesa. Nesse contexto, é importante observar que o discurso de Mafalda vai além do sentido das palavras nas sentenças, ela dá sentidos novos às coisas, aspecto que pode ser melhor compreendido se analisarmos tais discursos através das implicaturas. Como diz Ilari & GERALDI (2006, p. 75), “O sentido que a expressão assume então no contexto de fala pouco ou nada tem a ver com o sentido que se poderia esperar para a expressão a partir das palavras que a compõem”. É exatamente o que ocorre nessa tira, pois Mafalda usa de alguns artifícios do discurso para reforçar a sua negação ou repulsa ao prato de sopa, reforçando que não se diz palavrões nem mentiras na mesa. Como é possível entender o que significam os enunciados do discurso de Mafalda?

Exemplos como esse podem ser analisados a partir dos fenômenos conversacionais propostos pelo filósofo Paul Grice (apud ILARI & GERALDI, 2006, p. 76), que diz que “a conversação obedece a uma lógica própria, expressa por ‘regras conversacionais’ com as quais os falantes se comprometem ao engajar-se numa conversação”. Essas regras, tais como: “seja breve”, “seja relevante”, “seja tão informativo quanto exigido no momento da conversação” (idem, p. 76), podem ser exploradas pelos falantes sem que as cumpram, ou seja, violando-as, justamente para produzir sentidos diferentes. Assim, a relação de sentido se estabelece através de uma implicatura.

O mesmo autor (Idem 2006, p. 76-77), ao destacar a diferença entre implicaturas e pressuposições, diz que no processo para compreendermos um conteúdo pressuposto à estrutura linguística nos dá todos os elementos que nos permitem derivá-lo. Ao contrário, quando “derivamos implicaturas do discurso do nosso interlocutor ou as impingimos a quem nos ouve, esse suporte na construção linguística é muito menos óbvio” (Idem, 2006, p. 77). Assim, somos convidados a reinterpretar o enunciado, a observar que a interpretação da frase é inapropriada para a situação, como ocorre no discurso de Mafalda na tira da Figura 1, ao invés de procurar interpretar o enunciado através das expressões linguísticas particulares na frase.

No processo de ensino-aprendizagem de uma língua estrangeira, é importante o professor e o aprendiz estarem atentos a esses recursos da linguagem, a esses outros sentidos, no caso explicitado anteriormente, as implicaturas, pois tal fenômeno contribui para a compreensão de nuances não explicitadas no signo observado isoladamente, ou nas expressões linguísticas das frases, e só no contexto situacional do texto, ou melhor, na enunciação, evidencia-se como o enunciador, no caso a personagem Mafalda, cria sentidos diversos através de implicaturas. O significado pode ser literal – sopa, mesa –, mas os sentidos é que são novos no discurso de Mafalda, como palavrão, mentira, lugar sagrado.

Nessa tira, portanto, entendemos que Mafalda contrapõe o discurso da mãe de que “sopa não é palavrão” simplesmente para manter o seu argumento de rejeição ou negação ao prato de sopa, quebrando assim o princípio de relevância na enunciação. Ela poderia apenas dizer: não gosto de sopa, portanto, não vou tomar a sopa. Para ela, dizer que não se fala palavrões ou não se diz mentiras na mesa implica em rejeitar aquele tipo de prato, inclusive resgata o sentido de que mesa é um lugar puro, sagrado, por isso um lugar onde não se diz tais coisas, como palavrões e mentiras, convencionadas como impróprias, inadequadas.

Se por um lado entendemos que nesse contexto de enunciação há implicatura no discurso de Mafalda, entendemos também que ela, de certo modo, cria uma relação entre “sopa, palavrões, mentira”, a qual podemos entender como uma relação de sinonímia nesse texto, e que em outros discursos, o enunciador poderia dar à sopa o sentido de “coisa suja, feia, nojenta” e, nesse caso, não poderíamos chamar essa relação de sinonímia textual? Assim, resta ao interlocutor descobrir que na fala do enunciador há um outro sentido e não o literal, no nosso entender, há um outro sentido no texto. (ILARI & GERALDI, 2006, p. 76).

Outra discussão trazida por Hjelmslev é sobre as noções de expressão e conteúdo. Ele afirma que deixará de falar, por ora, em signos para falar em *função semiótica*, situada entre duas grandezas: *expressão* e *conteúdo* (Idem, p. 53). Segundo Hjelmslev, 2013, p. 54:

A função semiótica é, em si mesma, uma solidariedade: expressão e conteúdo são solidários e um pressupõe necessariamente o outro. Uma expressão só é expressão porque é a expressão de um conteúdo, e um conteúdo só é conteúdo porque é conteúdo de uma expressão. Do mesmo modo, é impossível existir (a menos que sejam isolados artificialmente) um conteúdo sem expressão e uma expressão sem conteúdo.

Para o autor (2013, p. 54), é evidente que “não se deve confundir ausência de conteúdo com ausência de sentido: o conteúdo de uma expressão pode

perfeitamente ser caracterizado como desprovido de sentido de um ponto de vista qualquer (...) sem com isso deixar de ser um conteúdo”.

É importante ressaltar que, de acordo com a teoria tradicional, o signo é a *expressão* de um *conteúdo* exterior ao próprio signo. No entanto, na teoria moderna formulada por Ferdinand de Saussure e seguida por Leo Weisberger, concebe-se o signo como um todo formado por uma expressão e um conteúdo (Idem, p. 53).

A partir de então, Hjelmslev (2013, p. 53) passa a falar, por ora, em *função semiótica* e não em signos, pois, para ele, a função semiótica está situada entre essas grandezas: a expressão e o conteúdo. O autor adota tais termos (expressão e conteúdo) para designar os functivos que contraem a função semiótica; para ele, esta concepção é puramente operacional e formal e, nesse sentido, nenhum outro significado é dado aos termos *expressão* e *conteúdo*.

Hjelmslev afirma ainda (2013, p. 54-55) que se se deixasse de considerar a função semiótica na análise do texto, não se poderia delimitar os signos, e não se poderia proceder de forma alguma a uma descrição do texto.

O sentido, para o autor, é o fator comum a todas as línguas, mas cuja execução é diferente em cada uma delas. Esse fator comum é uma grandeza que só se define pela função que a une ao princípio da estrutura da língua e a todos os fatores que fazem com que as línguas se distingam umas das outras. Para exemplificar, ele cita as sequências (cadeias) em diferentes línguas:

Jeg véd det ikke (dinamarquês)

I do not know (inglês)

Je ne sais pas (francês)

En tiedä (finlandês)

Naluvava (esquimó)

Em tais sequências, apesar das diferenças, o fator comum é o sentido, o mesmo pensamento que, considerado dessa forma, apresenta-se provisoriamente como uma massa amorfa, uma grandeza não analisada, definida somente por suas funções externas, ou seja, pela função absorvida com cada uma das proposições citadas.

Vale ressaltar que o sentido aqui apresentado pelo autor parece ser o mesmo que significado, mas sabemos que existe uma diferença entre significado e sentido na literatura, tanto em relação à visão de linguistas e semanticistas, como

na visão de teóricos do discurso, da pragmática e da enunciação, discussão que pode ser ampliada no decorrer do trabalho.

Para o autor (p. 56), pode-se analisar o sentido a partir de vários pontos de vista, de um ponto de vista lógico, ou psicológico quaisquer, no entanto, percebe-se que ele deve ser analisado de um modo particular em cada uma dessas línguas, pois o sentido só pode ser compreendido porque ele é ordenado, articulado, formado de modo diferente de acordo com as diferentes línguas, como se pode observar ao analisar-se as sentenças das línguas citadas, assim como também veremos na tira de Mafalda abaixo. Nesse aspecto, ressaltamos a validade de o professor e o aprendiz reconhecerem essas nuances no âmbito de um discurso ou de uma enunciação, através de noções como as implicaturas em determinada língua, no nosso caso o espanhol, como vemos através das tiras de Mafalda, pois nosso interesse não é fazer a comparação dessas nuances de sentidos entre línguas.

Figura 2



Fonte: <http://www.todohistorietas.com.ar/tiras2.htm> Consulta: 24/05/2018.

Na tira (Figura 2), de Mafalda, aparecem algumas questões bastante curiosas, uma relacionada à palavra “postre” (sobremesa), outra a “soborno” (suborno), e outra relacionada a “desertar” (abandonar). Quanto à palavra “postre”, observamos que ela surge como um tipo de chantagem caso Mafalda tome a sopa, ou seja, o enunciador deixa implícito que se ela não tomar a sopa, ela não come panquecas. Na sequência da tira, Mafalda parece querer resistir à chantagem da mãe, quando diz de forma imperativa “não a tomo!”.

Com relação à palavra “soborno”, ao que parece, observamos haver uma espécie de retomadas anafóricas, estabelecida pela relação semântica de sinonímia entre *postre* e *soborno*, quando a personagem, no caso Mafalda, não mais se refere ao significado semântico literal de *postre*, mas faz uma relação no texto

do segundo balão com “soborno”. Ela parece sugerir no segundo enunciado que a sua mãe a suborna com o “postre” quando ela toma a sopa, o que se confirma no terceiro quadro da sequência da tira quando a mãe de Mafalda oferece panquecas, e no último enunciado Mafalda aparece tomando a sopa e diz que às vezes tem nojo dela mesma. Nesse contexto, fica subentendido que a máxima da qualidade foi violada por Mafalda, quando ela defende seus princípios, mas, em seguida, nega-os, ainda que através de um enunciado que apenas lançando-se mão do recurso semântico das implicaturas pode ser compreendido.

Grice (1975, 1978, *apud* CANÇADO, 2008, p. 132) afirma que as implicaturas conversacionais podem ser previstas por um princípio de cooperação entre os falantes. A realização linguística desse princípio é apresentada por Grice em uma série de normas ou máximas, dentre elas a *máxima da qualidade*, descrita como segue: “Tente fazer da sua contribuição uma verdade, ou seja, não diga o que você acredita que seja falso, ou não diga nada de que você não tenha evidências adequadas” (idem, p. 132). Outra máxima que parece ser violada por Mafalda nesse enunciado é a *máxima de modo*, na qual se espera que o falante ou enunciador seja claro, evite ambiguidades, evite obscuridades, seja breve e ordenado.

Na tira da Figura 2, Mafalda cria ambiguidades no enunciado, e não é clara quando diz que seria repugnante se houvesse suborno, no entanto, no final da sua enunciação ela se submete ao suborno, tomando a sopa, ou seja, além de não ser clara, parece não falar a verdade. No discurso de Mafalda, parece estar implícito também que a mãe deve sempre suborná-la para tomar a sopa. No primeiro enunciado, a mãe tenta convencê-la (parece ser uma chantagem ou a primeira tentativa de suborná-la) de que se Mafalda não tomar a sopa, não come sobre-mesa. Mafalda parece resistir respondendo à mãe que “não a toma” e que não aceita suborno (por isso, cremos que seja um costume da mãe), mas quando a mãe oferece panquecas, Mafalda não resiste, pois adora comer panquecas. Como afirma Cançado (2008, p. 134),

[...] fica claro que esses princípios cooperativos divergem dos princípios linguísticos, no sentido de que eles podem ser e são violados frequentemente: muitas mentiras são ditas, as conversações são desviadas subitamente do seu curso por respostas desconexas e quem nunca conversou com alguém que dá muito mais informações do que as necessárias? O que ocorre é que essas normas podem ser violadas de forma deliberada, de modo que o falante sabe e reconhece que a máxima foi desconsiderada de uma maneira intencional.

Entendemos, portanto, que há também uma relação sinonímica entre as duas sentenças com relação ao sentido das palavras “postre” e “soborno”, não mais no



sentido literal ou isolado das palavras, mas no contexto em que ocorrem. Nesse caso, observamos que o fator comum entre ambas as palavras é o sentido empregado nesse contexto situacional e não em relação ao sentido específico ou determinado de cada signo isoladamente, como diz Hjelmslev, ou seja, esse tipo de relação semântica extrapola o nível linguístico (da palavra) e se forma no discurso.

No caso de “desertar”, o sentido empregado nesse contexto pode ser analisado em relação a sua construção sintática. Observamos que no texto o verbo “desertar” é seguido da preposição “de”, cuja forma intransitiva usada nesse contexto e cujo sentido implica em entender que o suborno seria capaz de fazê-la abandonar de algo que lhe é peculiar, próprio, diferentemente de se a personagem dissesse desertar algo, como se isso não lhe fosse peculiar ou que não lhe fosse obrigado (submetido) a fazer. Vemos nesse contexto que as construções sintáticas também são importantes na compreensão de determinados sentidos do texto na língua, o que é relevante na língua estrangeira, principalmente, pois o aprendiz deve desenvolver determinada competência linguística para reconstruir ou interpretar essas criações sintáticas visto que elas implicam ou podem implicar sentidos diferentes na(s) língua(s).

Assim, como afirma Hjelmslev (2013, p. 57), “cada uma dessas línguas estabelece suas fronteiras na ‘massa amorfa do pensamento’ ao enfatizar valores diferentes numa ordem diferente, [...] é o mesmo sentido que se forma ou se estrutura diferentemente em diferentes línguas”.

Segundo o autor, portanto, constata-se “no conteúdo linguístico, em seu processo, uma *forma* específica, a *forma do conteúdo*, que é independente do *sentido* com o qual ela se mantém numa relação arbitrária e que ela transforma em *substância do conteúdo*”. (p. 57), como vemos no caso da Figura 2 e na Figura 3, a seguir.

Figura 3



Fonte: <http://www.todohistorietas.com.ar/tiras2.htm> Consulta: 24/05/2018.

Aparece também outra situação de comunicação em que a palavra “solteirão” é enfatizada como sinônimo de “solitário” pelos personagens. A equivalência semântica entre as duas palavras nas sentenças da tirinha surge quando, no diálogo entre Felipe e Susanita, esta diz que solteirão e solitário é a mesma coisa, ou seja, todo solteirão é um solitário (solteirão = solitário, nesse contexto). As duas palavras do ponto de vista lexical não possuem identidade de significados e, no entanto, nessa sentença específica, a personagem traz essa relação entre elas nesse contexto de uso. Segundo a história da personagem, ela tem interesse romântico por Felipe, e, talvez por isso, use desse argumento linguístico buscando comparar o sentido de solteirão com solitário, para chamar a atenção do seu pretendente. Por isso, o uso desse artifício da linguagem torna as palavras sinônimas.

Esse caso também reforça a tese anterior, como o exemplo de sopa e surbro, de que a relação de sinonímia pode ocorrer entre palavras como também se estabelece no discurso, pois, sem considerar o contexto em que se inserem, poderiam não ser consideradas como sinônimos, no entanto, o texto evidencia essa relação de sinonímia entre solteiro e solitário, embora reconheçamos que existe um traço semântico (ou cultural?) que as aproxima. Apontamos, dessa forma, para o caso de que tais relações semânticas preexistentes, permitidas pelo linguístico, podem também ser vistas no plano do discurso, da enunciação, analisando a questão das implicaturas, implícitos, pressuposições, a serem retomadas mais adiante em Grice (1982), no texto “Lógica e conversação”.

#### 4. ALGUMAS CONSIDERAÇÕES FINAIS

É fascinante nos debruçarmos sobre o estudo de diferentes aspectos e diferentes níveis de análise da língua. O aspecto semântico da língua nos chama atenção em particular, pois perpassa, de certo modo, os demais aspectos linguísticos.

Nossa proposta inicial, visava observar aspectos semântico-linguísticos, como as relações de sinonímia, polissemia e homonímia nas tiras de Mafalda e em textos de livros didáticos. No entanto, desde a seleção dos dados, tanto das tiras como dos textos, começamos a observar que as tiras de Mafalda poderiam provocar um viés teórico distinto ou para além daquele que delimitamos no projeto inicial. Da mesma forma, percebemos que os textos selecionados nos manuais didáticos apontados no projeto não seriam suficientes para observar os fenômenos semânticos a serem analisados no trabalho, fato confirmado quando iniciamos as análises propriamente ditas.

Por isso, a partir de um novo recorte, tanto nos dados quanto no olhar teórico sobre os fenômenos que surgiam a cada tira de Mafalda selecionada e que analisamos, pudemos alcançar os objetivos propostos na pesquisa. Isso nos pareceu, inicialmente, um novo desafio na condução, desenvolvimento e conclusão do trabalho, mas nos permitiu fazer outras leituras e tecer observações sobre novos vieses dos fenômenos semânticos da língua através dos textos de Mafalda.

Buscamos, portanto, fazer um percurso através do aspecto estritamente linguístico-semântico e percorrer as nuances de sentidos que permeiam os textos através do discurso e da enunciação nas falas das personagens em Mafalda, um caminho que se tornou ainda mais gratificante, pois observamos que os textos, independentemente do aspecto espaço-temporal, como o das tiras de Mafalda, que são de um período político-social complexo e conflituoso da sociedade argentina da década de 1960 (do século XX), mas também de outros países da América Latina e do mundo, revelam através da linguagem quadrinística e cômica fenômenos linguísticos nos textos e nas atitudes cotidianas de qualquer falante de determinada língua.

## REFERÊNCIAS

CANÇADO, Márcia. **Manual de semântica: noções básicas e exercícios**. São Paulo: Contexto, 2012.

CANÇADO, Márcia. **Manual de semântica: noções básicas e exercícios**. Belo Horizonte: Editora UFMG, 2008.

HENRIQUES, Claudio Cezar. **Léxico e semântica: estudos produtivos sobre palavra e significação**. Rio de Janeiro: Elsevier, 2011.

HJELMSLEV, Louis. **Prolégomènes à une théorie du langage**. Paris: Editions de Minuit, 1966 [1943].

HJELMSLEV, Louis. **Prolegômenos a uma teoria da linguagem**. Trad. J. Ferreira Coelho Neto. São Paulo: Perspectiva, 2013.

ILARI, Rodolfo & GERALDI, João Wanderley. **Semântica**. 11.ed. São Paulo: Ática, 2006.

ILARI, Rodolfo & GERALDI, João Wanderley. **Semântica**. 3.ed. São Paulo: Ática, 1987.

ILARI, Rodolfo. **Introdução à semântica**: brincando com a gramática. 2.ed. São Paulo: Contexto, 2001.

OLIVEIRA, Roberta Pires de. **Semântica formal**: uma breve introdução. Campinas: Mercado de Letras, 2001.

SAUSSURE, Ferdinand de. **Curso de lingüística geral**. São Paulo: Cultrix, 1922.

SEÑAS. **Diccionario** para la enseñanza de la lengua española para brasileños. São Paulo: Martins Fontes, 2001.



# A RELAÇÃO DAS MUDANÇAS LINGUÍSTICAS HISTÓRICAS COM OS ERROS FONÉTICO-FONOLÓGICOS DO SISTEMA VOCÁLICO NO ENSINO-APRENDIZAGEM DE ESPANHOL/LE POR BRASILEIROS

*Aline Vieira Bezerra Higino de Oliveira<sup>1</sup>*

## 1. INTRODUÇÃO

A elaboração de materiais e recursos utilizados no ensino-aprendizagem de línguas estrangeiras, como a preparação de livros, editoração de manuais didáticos e produção de subsídios pedagógicos demonstram a escassez de desenvolvimento da competência oral no processo educativo, ratificando assim, a desvalorização da área da fonética e da fonologia no estudos escolares. Para ensinar uma língua estrangeira (LE), todos os aspectos são importantes e necessários, pois a interação é resultado de conhecimentos linguísticos compartilhados pelo coletivo, que interagem conhecendo os códigos, símbolos e sons. Por isso, ao lecionar, o professor deve nutrir o aluno de todas as informações essenciais para o desenvolvimento das conversações e várias outras modalidades de fala. Dentre elas, está a produção oral, que traz consigo a realização adequada para cada fonema para que se possa fazer a interpretação correta da mensagem enunciada. Segundo Fernández (2007), deve-se reiteradamente estar atento para esse aspecto, pois os esforços empreendidos na aprendizagem de uma língua estrangeira podem ser desvalorizados em virtude da pronún-

---

<sup>1</sup> Universidade Federal de Alagoas.

cia imprópria de alguns fonemas. Muitas das inadequações ocasionadas nas interlocuções são causadas pelas interferências da língua materna (LM), que se adentram na fala do aprendiz de LE para suprir alguma carência e/ou desconhecimento da língua pretendida.

No ensino-aprendizagem da língua espanhola para brasileiros percebemos diversas interferências oriundas da proximidade entre as línguas espanhola e portuguesa, já que ambas são derivadas do Latim vulgar, originado das línguas do ramo indo-europeu, as línguas romances. Essa semelhança provê a inclusão de desenvolvimentos sonoros análogos e/ou desiguais na aprendizagem das línguas irmãs, por elas partilharem da mesma origem.

Podemos considerar os percalços dessa proximidade e analisar suas ocorrências a fim de analisar e/ou dirimir os possíveis ‘erros’ originados das interferências da LM no ensino-aprendizagem da LE. Segundo Brisolara & Semino (2016), o processo de aprendizagem de espanhol/Língua Estrangeira (LE) induz o aprendiz brasileiro a utilizar regras fonético-fonológicas pertencentes ao sistema linguístico de sua língua materna (LM), causando as interferências que transferem marcas da LM na aprendizagem do espanhol e são processos que devem ser desinstalados a fim de aprimorar o desempenho do estudante.

Vários estudos linguísticos – Beux (2014), Santos (2016), Sandes (2010), Oliveira (2009), Barbosa (2014), Brisolara & Semino (2014), Brandão (2003), Carvalho (2004), Costa (2013), Dias (2011), Machry da Silva (2014), Mignorini (1999), Silva (2007) – explanaram sobre as dificuldades fonético-fonológicas da aprendizagem de espanhol por brasileiros. Dos fenômenos observados, analisando-se, contrastivamente, os quadros fonológicos das línguas reveladas, notou-se que as interferências da LM aconteciam no sistema consonantal e vocálico do espanhol, sendo percebido na produção oral e escrita e na compreensão oral e escrita.

Nesta pesquisa, temos retomado a análise dos erros no ensino-aprendizagem de espanhol LE por brasileiros, conforme explanado em OLIVEIRA (2009 e 2020) e nas pesquisas supracitadas, de acordo com diferentes linhas teóricas que buscam explicar o fenômeno. Por isso, questionamos quais são as causas dos erros fonético-fonológicos do espanhol que ocorrem na aprendizagem de espanhol/LE que geram dificuldades no ensino-aprendizagem do espanhol/LE para alunos brasileiros. Temos como hipótese que o conhecimento histórico e teórico do sistema fonético-fonológico do espanhol e do português pode contribuir para a identificação, descrição e explicação dos erros de alunos brasileiros, visando a contribuições no ensino-aprendizagem do espanhol/LE. Esta explanação tem como finalidade descrever e analisar historicamente os erros fonético-fonológicos do sistema

vocálico de alunos brasileiros de espanhol/LE através de estudo contrastivo entre o português e o espanhol a fim de verificar a relação histórica dos fenômenos linguísticos com as causas das dificuldades apresentadas por estudantes brasileiros na aprendizagem do Espanhol/LE. Nesse sentido, estamos retomando resultados das pesquisas já realizadas, inserindo tanto o conhecimento histórico quanto os fatos sincrônicos das duas línguas, de modo a delinear uma possível convergência entre esses conhecimentos e os erros já determinados pelos autores.

## **2. ESTUDOS REALIZADOS**

Pesquisas comparativas e contrastivas vêm sendo feitas sobre o ensino-aprendizagem de espanhol/LE sob a ótica da análise das estruturas e identificação de pontos que podem motivar discordâncias e dificuldades de pronúncia. Apresentamos estudos que vêm abordando esse assunto contrastivamente objetivando confirmar a presença das dificuldades fonéticas e fonológicas na aprendizagem desses idiomas próximos e evidenciar a existência de dados já detectados.

Os estudos de Hoyos-Andrade (1978) contrastaram os sistemas vocálicos do espanhol e do português, assinalando já as tendências do vocalismo português que podem causar ‘erros’ na aprendizagem do espanhol, sugerindo o controle constante do professor e dos alunos para os eventuais deslizos, além de fazer várias sugestões para impedir a pronúncia incorreta.

Quilis (1979) realizou uma descrição mais ampla do espanhol e do português, usando fontes de países com o idioma espanhol e submergiu nas regiões do Brasil para deter-se sobre os sistemas linguísticos, empregando já demonstrações acústicas das diferenças dos idiomas.

Brandão (2003) organizou um estudo comparativo considerando algumas das variações dialetais de ambos os idiomas, para posteriormente identificar os erros fonéticos constantes na produção oral dos alunos com a finalidade de caracterizar e interpretar as metodologias de correção ou de tratamento desse tipo de erro empregado.

Masip (2003), apresentou as origens do português e do espanhol sob uma visão sintética, indutiva e fragmentária, de forma contrastiva, com um viés histórico, exibindo os fonemas e prosodemas portugueses e espanhóis e suas respectivas grafias, formas e funções sincrônicas.

Milani (2006), tratou das semelhanças e diferenças entre os dois sistemas, explanando quais recursos linguísticos os alunos, aprendizes brasileiros de espanhol, têm de adaptar para pronunciar os fonemas da língua espanhola.



Oliveira (2009), afirmou que os brasileiros têm problemas pontuais na aprendizagem de espanhol, em função das interferências da língua materna, tanto no sistema consonantal quanto no sistema vocálico, incidindo em erros que causam dificuldades na compreensão oral e escrita e na produção oral e escrita.

Sandes (2010) utilizou os estudos comparativos, identificando as dificuldades dos estudantes brasileiros, privilegiando a análise acústico-articulatória das aproximantes e das nasais da língua espanhola, objetivando encontrar estratégias de correção no âmbito da produção de sons em E/LE.

Masip (2014) ilustrou as dificuldades fonéticas, fonológicas e ortográficas específicas da aprendizagem de espanhol por brasileiros, minudenciando a forma de evitar a pronúncia incorreta e a interferência do português, e igualmente explanou sobre em (2010), com a finalidade de explicar as dificuldades existentes na pronúncia entre duas línguas próximas.

Brisolara & Semino (2016), demonstraram a existência de interferências específicas no sistema vocálico e no sistema consonantal. No sistema vocálico apresentaram como interferências: a abertura das vogais médias, a harmonia vocálica, a neutralização das vogais átonas em posição postônica final, a nasalização das vogais e o alongamento vocálico.

Objetivamos demonstrar as interferências supracitadas pelos autores particularizando através dos modos de articulação os possíveis erros que os aprendizes podem ter durante o processo de ensino-aprendizagem e possivelmente, levando à fossilização dessas inaptações em estágios finais. Discutimos os erros apontados segundo Hoyos-Andrade (1978), Quilis (1979), Brandão (2003), Milani (2006), Oliveira (2009 e 2020), Sandes (2010), Masip (2003, 2010 e 2014) e Brisolara & Semino (2016).

A exposição dos erros dar-se-á pelo modo de articulação, explicando como surge o fenômeno e em que ambiente linguístico, explanando através dos traços distintivos a produção oral adequada da língua espanhola/LE e a produção do estudante brasileiro aprendiz de espanhol, exemplificando o erro cometido. Após os quadros demonstrativos retirados de Oliveira (2020), serão expostas análises históricas dos fenômenos em português e em espanhol, objetivando esclarecer as possíveis causas dos ‘erros’ por meio de uma ponderação cronológica e linguística.

### 3. INTERFERÊNCIAS VOCÁLICAS

Inicialmente, faremos a descrição do sistema vocálico das línguas e em seguida pontuaremos as interferências da língua materna na aprendizagem de espanhol por brasileiros através da análise histórica e seus processos.

Segundo Mattoso (2014), o sistema vocálico do português (PB) possui 7 fonemas vocálicos orais (/i, e, ε, a, o, ɔ, u/) em posição tônica, 5 vogais em posição pretônica (/i, e, a, o, u/), observando-se somente em alguns dialetos /ε, ɔ/; em posição postônica não final, abrevia-se o sistema vocálico pela neutralização da oposição entre /o/ e /u/, enquanto em posição pós-tônica final reduz-se ainda mais para as vogais /a, i, u/. Podemos e devemos considerar as nasalizações existentes no sistema vocálico do português em ambientes e palavras específicas.

A língua espanhola possui 5 fonemas vocálicos /i, e, a, o, u/. De acordo com Martínez Celdrán & Fernández Planas (2007), considera-se a existência de vogais médias baixas em espanhol sem função fonológica, em virtude da percepção fonética de realizações de maior ou menor grau de abertura vocálica.

Apresentamos, na próxima seção, as seguintes interferências ocasionadas na aprendizagem de espanhol por brasileiros: abertura das vogais médias, harmonia vocálica, neutralização das vogais átonas em posição final, nasalização das vogais e alongamento vocálico.

#### 3.1. Abertura das vogais médias

Na aprendizagem do espanhol por brasileiros, percebemos a pronúncia de vogais médias altas como médias baixas em posição tônica. Essa realização é característica da língua portuguesa e comumente evidenciada na pronúncia da língua espanhola quando realizada por estudantes brasileiros.

Abertura das vogais médias		
REPRESENTAÇÃO ORTOGRÁFICA	PRODUÇÃO DE LE	PRODUÇÃO DO ESTUDANTE DE LE
	[e]	[e] → [ɛ] / _\$
	[o]	[o] → [ɔ] / _\$
MÉDICO	[ˈmediko]	[ˈmediko]
MODA	[ˈmoda]	[ˈmoda]

Para compreensão desse fenômeno, devemos demonstrar seu desenvolvimento do latim para a língua portuguesa e para a língua espanhola. No latim

imperial, existiu a perda das oposições de quantidade, permanecendo as oposições de timbre. Essa modificação dos cinco timbres vocálicos, breves e longos, ao todo de dez fonemas para sete fonemas, trouxe para a língua portuguesa, a vogal /ε/, nascida da redução do /ě/ > /ẹ/ > /ε/ e a vogal /ɔ/, nascida do /ǒ/ > /o/ > /ɔ/.

As vogais do latim evoluíram de forma distinta na língua espanhola. As mudanças do latim clássico para o latim falado, e posteriormente o espanhol, ocorreram das vogais breves para ditongos. Podemos esclarecer dessa forma: do /ě/ > /ẹ/ > /ie/; do /ǒ/ > /o/ > /ue/. Portanto, podemos compreender que no português as vogais abertas permaneceram na oralidade, enquanto em espanhol, as vogais abertas evoluíram para ditongos.

A explicação exposta demonstra o que ocorre na aprendizagem de espanhol por brasileiros, que têm dificuldade em substituir o uso das vogais abertas tônicas pela ditongação em ambiente específico.

Podemos afirmar que as vogais médias em português têm na abertura sua característica distintiva e, portanto, fonológica. Segundo Mattoso Câmara (2014), os fonemas vocálicos não sofrem nenhum processo e só ocorrem distintivamente em sílaba acentuada, podendo ocorrer em sílaba não acentuada e também em sílaba postônica ou pretônica.

A existência de abertura das vogais em espanhol /ε/ e /ɔ/ é uma particularidade alofônica que ocorre em ambientes específicos, ou seja, são alofones posicionais. O fonema /ε/ acontece em contato com vibrante múltipla /r/, antecedida ou precedida à vogal, com exclusão dos casos nos quais a vogal ocorre em sílaba travada pelas consoantes <d, m, n, s, x, z>; diante de /x/; nos ditongo aberto ‘ei’ e ‘oi’; em sílaba travada, com exceção das consoantes <m, n, s, d, x, z>.

Tomas (1989) alega que a vogal [ɔ] pode ser percebida nas seguintes situações: a) em contato com uma vibrante, involuntariamente, antecedendo-a ou precedendo-a; b) diante de <j> e <g> com som de /x/; c) no ditongo <oi> u <ou>; d) em sílaba travada, independentemente da consoante que se encontra na referida posição; e) em posição tônica, quando a vogal [ɔ] tenha como antecedente a vogal <a> e precedente um <r> ou <l>.

Para o português, a abertura de vogais médias e para o espanhol, a ditongação das vogais, é, de fato, uma dificuldade para a aprendizagem do espanhol por brasileiros. Os estudantes não intuem a diferença da abertura e também pouco identificam o lugar da inclusão dos ditongos. Esses fenômenos pesquisados por Santos (2010: 149) demonstram nos resultados da análise dos dados sobre “A percepção e produção das vogais médias do espanhol por falantes do português

brasileiro”,<sup>2</sup> que na produção das vogais [e] e [o] do espanhol por falantes do PB mais aberta não provoca ao falante de E/L2 dificuldades de entendimento entre ele e nativos de E/L1, contudo, origina a identificação do acento estrangeiro, especialmente quanto aos falantes de L2 em nível avançado de proficiência. Como também, a interferência da L1 pode ocorrer além da oralidade, com a pronúncia das vogais médias-baixas do PB frente a das médias-altas do espanhol podendo originar erros na ortografia, como acentuação gráfica indevida.

Nos estudos perceptuais de Silva (2014, p. 220), analisou-se nas decorrências que os nativos falantes do espanhol não percebem com clareza a distinção quanto à abertura das vogais médias tônicas e, por isso, também têm problemas em realizá-las, em oposição à pesquisa anterior, pois não veem necessidade da pronúncia correta para comunicação.<sup>3</sup> Sobre a percepção do contraste entre vogais médias tônicas averiguou-se que falantes nativos do espanhol são moderadamente sensíveis à realização fonética das vogais /e/ e /o/ do português, entretanto, tendem a expor dificuldade para distinguir os contrastes fonológicos /e/ – /ɛ/ e /o/ – /ɔ/ do português. Dessa forma, entende-se que as categorias fonológicas da L2 não estão perceptualmente bem apuradas para os estudantes e que as vogais /e/ e /ɔ/, não próprias ao sistema vocálico do espanhol, inclinam-se a ser compreendidas com base nas categorias existentes na língua nativa. Além disso, o contraste fonológico entre vogais tônicas aponta que os estudantes têm dificuldade em realizar as diferenças entre os pares de vogais /e/ – /ɛ/ e /o/ – /ɔ/ do português. Assim sendo, as duas categorias de vogais /e/ – /ɛ/ e /o/ – /ɔ/ na fala de não nativos são no aprendizado justapostas uma à outra. Os resultados de percepção apontam que as categorias fonológicas da L2, referente ao sistema vocálicoônico, não apresentam clareza para os aprendizes também em sua realização.

O galego-português, o espanhol, vulgo castelhano, e o português europeu e o brasileiro, tiveram percursos diferentes que geraram vogais abertas no português e ditongação no espanhol. Essas propriedades originárias e a diferenciação de ambientes linguísticos de realização das vogais causam na atualidade dificuldades de aprendizagem do espanhol para brasileiros. A interferência na aprendizagem de espanhol por brasileiros é originária desde a transformação do

---

<sup>2</sup> Textos que tratam da interferência das vogais médias na aprendizagem do espanhol por brasileiros, BAZZAN, M. A. T. (2005), MIGNONI, R. P. L. (1999).

<sup>3</sup> A tese “Aprendizagem fonológica e alofônica em L2: percepção e produção das vogais médias do português por falantes nativos do espanhol” de Silva (2014) trata de um estudo com colaboradores nativos de espanhol, tendo nessa análise uma perspectiva inversa ao tratado nesse texto.

latim clássico em latim vulgar e da separação do galego-português em português e castelhano.

Processos fonético-fonológicos da abertura das vogais médias		
Espanhol	/ɛ/ > /e/ > /ie/	- Perda das oposições de quantidade
	/ɔ/ > /o/ > /ue/	- Evolução das vogais breves para ditongos.
Português	/ɛ/ > /e/	- Perda das oposições de quantidade
	/ɔ/ > /o/	- Evolução das vogais breves para vogais médias abertas.

### 3.2. Harmonia vocálica

A harmonia vocálica, tipo de apropriação de propriedades vocálicas, na quais as vogais de uma palavra tornam-se foneticamente semelhantes a outra vogal da mesma palavra (comumente a tônica, mas não obrigatoriamente), com a mudança de alguns dos seus traços, é um fenômeno que acontece no português do Brasil de forma constante, sendo analisado em vários estudos científicos, até mesmo já tendo sido introduzido em Mattoso Câmara (2014:45).

Harmonia vocálica		
REPRESENTAÇÃO ORTOGRÁFICA	PRODUÇÃO DE LE	PRODUÇÃO DO ESTUDANTE DE LE
	[e]	[e] → [i] / _\$
	[o]	[o] → [u] / _\$
Escuela	[es'kwela]	[is'kwela]
Mantenimiento	[manteni'miento]	[mantini'miento]

De acordo com Cagliari (2002:104), “A harmonia vocálica é um tipo especial de assimilação que faz que vogais se tornem mais semelhantes entre si, em geral, por alguma razão morfológica (regra morfofonológica)”. Segundo Bisol (2015:188), “A harmonia ou harmonização vocálica consiste na substituição da vogal média /e, o/ pela vogal alta /i, u/ respectivamente, quando a média precede uma sílaba com vogal alta, a exemplo de pepino ~ pipino, coruja ~ curuja, bonito ~ bunito”.<sup>4</sup>

Como explica Bisol (2015), que aborda a harmonização vocálica entendendo-a como uma regra variável e uma indicação de desestabilizações históricas, decompondo o português em dois dialetos, português europeu e português bra-

<sup>4</sup> Sobre harmonia vocálica sugerimos BISOL, L. (1981). Harmonização vocálica: uma regra variável. Rio de Janeiro, UFRJ. Tese de doutorado.

sileiro, com início no final de XVIII, sendo assim vislumbrado de forma transparente no início do século XIX. Para Bisol (2015:203), o português brasileiro ergue-se historicamente a partir do século XVI, com o sistema átono de cinco vogais e harmonização vocálica na pretônica. O português europeu traz a centralização das vogais átonas [-post], seguindo outra deriva, em função das sequelas de três processos: centralização, harmonização e neutralização. Apesar disso, PEC e PBC<sup>5</sup> são dialetos de uma língua só (BISOL, 2015:203).

Segundo Bisol (2015:189), apontamentos escritos em latim-português surgiram nos séculos IX a XI. Analisou-se que vocábulos portugueses apareceram permutados com latim. Nesses esboços, ficou demonstrada já uma disposição à harmonização vocálica nos dados descobertos, manifestando a alteração da vogal média ante uma sílaba com vogal alta e alçamento da vogal média sem estímulo manifesto, ou seja, sem a presença de uma sílaba subsequente com vogal alta. Determinadas palavras evidenciam esse fenômeno como uma herança do latim vulgar, entendendo o início desse processo ao evidenciar a alteração da vogal média pretônica. Ex.: *cumtestamus* < *contestamos*; *mulinos* < *molina* *contuversia* < *controversiam*; *obturigare* < *auctoricare*.

Fenômeno peculiar da língua portuguesa falada no Brasil, a harmonia vocálica, não pode ser empregada na produção oral do espanhol, pois o espanhol mantém as vogais médias em todas as posições, não consentindo que exista essa transferência, alçamento e/ou alteamento das vogais com a intenção da simplificação articulatória.

A simplificação da movimentação no trato vocal, igualando a altura das vogais e facilitando a articulação dos sons vocálicos, não ocorre na língua espanhola, pois ela não faz uso desse procedimento articulatório. Observa-se que a ausência desse fenômeno, não estigmatizado no português, provoca um procedimento igual na aprendizagem do espanhol, pois, além da ausência das vogais abertas, existe a obrigação de realizar a pronúncia das vogais sem harmonizá-las. “La armonía vocálica no es un fenómeno característico del español el cual mantiene sus vocales medias en cualquier posición de las palabras”. (BRISOLARA, L. & SEMINO, M. J. 2016: 47).<sup>6</sup>

Entre os séculos IX e XI, esse fenômeno que foi observado na transição do latim vulgar para a língua portuguesa podendo já ser visualizado nitidamen-

---

<sup>5</sup> Acreditamos que as siglas PEC e PBC, citadas pela autora, sejam, respectivamente, português europeu contemporâneo e português brasileiro contemporâneo.

<sup>6</sup> Tradução nossa: [...] “A harmonia vocálica não é um fenômeno característico do espanhol que mantém suas vogais médias em qualquer posição das palavras”.

te no início do final de XVIII, não permitiu sinais no espanhol, mantendo seu sistema vocálico até os dias atuais. É importante saber que na aprendizagem de português por falantes de espanhol devemos observar esse fenômeno, haja vista que a compreensão oral e as demais competências do estudante são afetadas por essas realizações de harmonização vocálica, segundo os autores Machry da Silva (2014) e Guimarães (2011), que perceberam a inexistência de subsídios didáticos que versassem sobre as dificuldades do aprendiz falante de espanhol na aprendizagem da harmonia vocálica presente no sistema vocálico do português.

Processos fonético-fonológicos da harmonia vocálica		
Espanhol		- Não ocorre na língua espanhola
Português	/e, o/ por /i, u/ /i, u/ por /e, o/	- Alçamento e/ou alteamento das vogais

### 3.3. Neutralização das vogais átonas em posição final

A neutralização das vogais átonas em posição final é um fenômeno circular no português do Brasil, pronúncia reconhecida e não estigmatizada pelos brasileiros. Nesse processo, as vogais médias /e, o/ em posição final se transformam em [i, u], amortizando o sistema vocal do português brasileiro a três vogais /i, u, a/ ou, melhor representado, [ɪ, ʊ, a].

Neutralização das vogais átonas em posição final		
REPRESENTAÇÃO ORTOGRÁFICA	PRODUÇÃO DE LE	PRODUÇÃO DO ESTUDANTE DE LE
	[e]	[e] → [i] / _ \$
	[o]	[o] → [ʊ] / _ \$
Pato	[ˈpato]	[ˈpatɔ]
Caliente	[kaˈlyente]	[kaˈlyenti]

Estudiosos da biografia da língua portuguesa retroagiram quanto à periodização do fenômeno da pronúncia do /o/ e /e/ como [u] e [i] em posição átona final, que datava da primeira metade do século XVIII, para uma época anterior, no mínimo até o século XVI, segundo Teyssier (2007, p. 73).

Percebemos que no galego-português havia um [o] e um [e] fechados que a deriva da língua ia desviando para um grau de fechamento e brevidade cada vez maior, pois havia certa grafia em -u no galego-português medieval. A pronúncia dessas vogais era [o] e [e], como ainda hoje permanece em galego, desde

o português do século XVI. Existia nessa posição apenas três fonemas e sendo pouco provável a transformação de [o] a [u] e /e/ a /i/. Os lusófonos passaram muito tempo sem ter a consciência dessa deriva, sendo imperioso esperar pela primeira metade do século XVIII para confirmarem a pronúncia de [u] e [i]. Após essa confirmação, a realização [u] para a vogal escrita -o foi crucial. Todavia, a realização [i] para a escrita de -e será em Portugal, em linguagem popular, uma simples fase intermediária para o [ẽ] atual, que deve ter surgido na segunda metade do século XVIII.

Teyssier (2007, pp. 100-101) afirma sobre a pronúncia das vogais átonas na língua portuguesa falada no Brasil, “Em posição átona final o falante brasileiro (excetuando-se os do extremo sul) pronuncia [u] o -o gráfico, tal como o português de hoje, mas diz [i] por -e, ignorando completamente a realização [ẽ]; ex.: passo pronunciado [páso], passe pronunciado [pási]”. Como também, afirmou que “Conservou-se, logo, no ‘brasileiro’ a situação que era a do português europeu da primeira metade do século XVIII, a que Verney descrevia em 1746. Acrescente-se que o -a átono final, embora mais breve que em posição tônica, permanece muito aberto”.

Mesmo de forma estilística em espanhol, é necessário revelar a presença de estudos que apresentam elementos sobre esse contexto, demonstrando algumas realizações já existentes do fenômeno em língua espanhola. Em Ramírez (1998), que trata da neutralização das vogais átonas em posição final percebidas em México e Porto Rico, temos que “el proceso se cumple en las hablas populares y campesinas de algunas regiones hispano-americanas, favorecido por la presencia de consonante palatal anterior”.<sup>7</sup> Essa revelação exposta pelo autor não foi observada como um aspecto distintivo, evidenciando apenas uma tendência a ser assimilada pelos falantes. Herrero (2004:10) também aclara sobre essa tendência: “En cuanto a la presencia de las vocales <-e>, <-o> finales cerradas, es un fenómeno general en el español propiciado por la natural perdida de fuerza articulatoria al final de palabra, favorecida en hablas populares por la presencia de palatal anterior”.<sup>8</sup> Como essa pesquisa ainda é incipiente, observada em algumas regiões específicas, vamos tratá-los como variações geográficas específicas, não podendo assim expandir sua abrangência.

---

<sup>7</sup> Tradução nossa: [...] “o processo se realiza nas falas populares e camponesas de algumas regiões hispano-americanas, favorecidas pela presença da consoante palatal anterior”.

<sup>8</sup> Tradução nossa: [...] Quanto à presença das vogais <-e>, <-o> finais fechadas, é um fenômeno geral em espanhol causado pela perda natural da força articulatoria ao final da palavra, favorecida na fala popular pela presença do palatal anterior.



Diante do exposto, Brandão (2003) exhibe dois motivos pelos quais os aprendizes brasileiros de espanhol tendem a cometer inadequações nas suas realizações orais no caso da neutralização vocálica, que são: reprodução do dialeto falado pelo professor, caso conste esse fenômeno ou pela transferência dos padrões do Português ao Espanhol.

Processos fonético-fonológicos da neutralização das vogais átonas em posição final		
Espanhol		- Ocorrência branda pouco registrada na língua espanhola.
Português	/e, o/ por [ɪ, ʊ]	- Deriva da língua indutiva para a neutralização das vogais átonas em posição final.

Desde o século XVI esse fenômeno já encontrado no Brasil, não é característico do espanhol. Entretanto, podemos ponderar que a pronúncia real dos falantes de espanhol tem um relaxamento articulatório final na posição átona, evidenciando uma redução da intensidade na emissão das expressões. Na fala mais distraída ou em situações familiares de maior intimidade, pode acontecer uma flexibilização, como também em falares dos camponeses ou populares pode surgir essa realização acrescentada à presença da consoante palatal anterior, mas em falas cultas ou menos descuidadas torna-se clara a ausência desse fenômeno (BRISOLARA e SEMINO, 2016, p. 47).

### 3.4. Nasalização das vogais

A nasalização das vogais procede do encontro de uma vogal com uma consoante nasal, independente de essa consoante pertencer à mesma sílaba, que se estabelece a regra geral para que haja o fenômeno de nasalização da vogal em português, de acordo com Câmara Jr. (2014). Essa norma pondera sobre dois tipos de nasalização: a nasalização contrastiva ou fonêmica, por ter função distintiva, e a nasalização automática ou fonética, sem distinguir significados. Evidencia-se na aprendizagem de espanhol por falantes do português a disposição a nasalizar toda vogal seguida de consoante nasal, sendo na mesma sílaba ou em outra, segundo Hoyos-Andrade (1978).

Nasalização das vogais		
REPRESENTAÇÃO ORTOGRÁFICA	PRODUÇÃO DE LE	PRODUÇÃO DO ESTUDANTE DE LE
	[a,e,i,o,u]	[a,e,i,o,u] → [ã,ê,ĩ,õ,ũ-] / _ \$
Cama	[ˈkama]	[ˈkãma]
Mente	[ˈmẽnte]	[ˈmẽnti]

Ainda sobre a nasalidade, Hoyos-Andrade (1978) trata da afinidade entre o galego e o português, sobretudo da evolução do comportamento das nasais e da nasalidade. Inclusive no século XIII, essas línguas apresentaram evoluções análogas e conjuntas e nesse momento histórico principiaram a sofrer alterações, como podemos confirmar por meio da perda da nasal latina intervocálica -n-, através de uma realização velar com nasalização da vogal anterior, como em *luna > lua, manu > mano > mão*. Essas mutações nas consoantes nasais deram início à diferenciação entre as línguas: inicialmente, a perda da nasalidade, em seguida, a recuperação da consoante nasal, (port. vinho, gal. viño), conservação das vogais nasais em português (*pão*) e, de forma distinta, o galego já não expõe essa nasalidade vocálica (*pan*). Ainda hoje, as vogais são expostas diferentemente: o português exhibe vogais com nasalidade contrastiva, enquanto o galego apresenta uma nasal intervocálica em alguns termos sem possuir distinção fonológica, sendo apresentada como uma ocorrência intermediária entre português e espanhol.

Ao confrontar os sistemas fonológicos do espanhol e do português a fim de ressaltar seus elementos com finalidades didáticas, Quilis (1979) trata sobre a nasalidade dos dois idiomas, explanando com exemplos suas características. Quilis (1979), ainda sobre o sistema vocálico referente à nasalidade, fala que o espanhol não apresenta dificuldades nessa questão porque o traço de nasalidade não é pertinente ao sistema vocálico e seus cinco fonemas possuem alofones em distribuição complementar, no que concerne ao fator nasal. Dessa forma, temos [ĩ, ê, ã, õ, ù] quando /i, e, a, o, u/ se encontram entre pausa e consoante nasal ou entre duas consoantes nasais. Nas outras proximidades, se realizam como [i, e, a, o, u]. “Ejemplos: [ẽ̃ntre] entre, [mãno] mano, en posición tónica; o [ẽ̃ntre] entré, [mãnirróto] manirroto, en posición átona. Por otro lado, [kámpo] campo, [tẽ̃nte] tente, etc. Su status fonológico es el siguiente: /ẽ̃Ntre/, /mãno/, /eNtré/, /manirróto/, /káNpo/, /téNte/”<sup>9</sup> (QUILIS, 1979: 05).

Em português, é necessário esclarecer que sobre a nasalidade são apontadas posições de estudiosos que afirmam ou negam a existência de fonemas vocálicos nasais. Alguns afirmam a ocorrência de vogais nasais frente a vogais orais em português, outros contrariam a existência de fonemas vocálicos nasais

---

<sup>9</sup> Tradução nossa: Exemplos: [ẽ̃ntre] entre, [mãno] mano, em posição tónica, ou [ẽ̃ntre] entré. [mãnirróto] manirroto, em posição átona. Por outro lado, [kámpo] campo, [tẽ̃nte] tente, etc. Seu status fonológico é o seguinte: /ẽ̃Ntre/, /mãno/, /eNtré/, /manirróto/, /káNpo/, /téNte/”. A tradução preserva os exemplos em língua espanhola, pois sua tradução altera o sentido do texto.

no Português. Sobre a não existência de vogais nasais na língua portuguesa, são citados Mattoso Câmara (1977 e 1979), proporcionando as seguintes motivações: a presença de uma consoante nasal impede a oposição de uma vogal nasal pura a uma vogal nasal com consoante nasal pós-vocálica; a nasalidade pode estar presente nas vogais seguidas de nasais e mesmo assim permanecer em sílabas diferentes, com um grau maior ou menor; a vogal nasal se comporta como sílaba travada por consoante; sílaba com nasal em coda iniciando a sílaba posterior por vibrante interfere na realização desse som, tornando-a simples ou múltipla.

Sugere-se, finalmente, que as vogais nasais são vogais orais seguidas do arquifonema nasal. Ante as afirmativas se compara o português com o espanhol, assinalando que a exclusiva semelhança fonológica descoberta é que nos dois sistemas existem somente vogais orais, tendo a dessemelhança centrada nas realizações fonéticas dos timbres das nasais entre ambas.

Em Battisti & Vieira (2005), sobre o debate das vogais nasais ou nasalizadas no português, abeira-se a uma conclusão com base nos estudos já defendidos antes, como os de Câmara Jr. (1976 e 1979), que na subjacência, a vogal nasal é uma sequência de dois segmentos: VN. Considera-se uma vogal nasal pura quando despontada na superfície. Isso quer dizer que a nasalidade vocálica não intervém na composição fonológica das vogais do português. Para Mendonça (2019), diversos estudos sobre nasalização ainda são elementares, desse modo faz-se indispensável a ampliação de pesquisas sobre nasalização fonética sob o viés da sociolinguística e de análises acústicas a fim de constatar com mais propriedade esse fenômeno. Cristófaró Silva (2005) afirma que todas as vogais orais, com exceção das médias baixas [ɛ, ɔ], podem realizar-se nasalizadas.<sup>10</sup>

Em língua espanhola acontece a nasalização das vogais orais quando estas se concretizam antes de pausa e consoante nasal ou entre duas consoantes nasais. Portanto, em outros ambientes elas se realizarão orais. É salutar ressaltar que o grau de nasalização dos falantes realiza-se em menor grau, sendo para os brasileiros pouco reconhecível.

Podemos também adicionar as observações de Barbosa (2014), que realizou uma pesquisa auditiva e acústica para examinar a situação dessas vogais na atualidade para falantes da Europa e da América do Sul. Corroboraram nos resultados que a despeito das diferenças fonéticas entre o espanhol e o português brasileiro, a existência de vogais nasalizadas diante de consoante nasal na língua espanhola é mais habitual do que se avaliava. “Finalmente, foi mostrado, através

---

<sup>10</sup> Cristófaró Silva (2005) afirma que as vogais que não se nasalizam são exatamente as vogais abertas.

de nossas análises, que ocorre nasalização vocálica em espanhol também em contextos não contemplados na literatura tradicional” (BARBOSA, 2014:152).

Concluimos que havia nas duas línguas a nasalidade vocálica como resíduo das consoantes nasais existentes do latim clássico e vulgar, que se transportaram para o espanhol e para o português nas suas vogais. Na língua portuguesa, algumas palavras sustentam o registro da nasalidade em vogais, independente da sua representação estar atualizada na superfície. Em língua espanhola, a nasalidade aparece nas vogais com a proximidade de nasais. Esses fenômenos expressados são representações das origens latinas, que ainda estão presentes na oralidade das duas línguas. Esse esclarecimento deve auxiliar professores de línguas a instruir mais claramente como funciona a fala da língua espanhola para o falante de português, e vice-versa.

Processos fonético-fonológicos da nasalização das vogais		
Espanhol	[a, e, i, o, u] → [ã, ê, î, ô, û]	- Perda da nasal latina intervocálica -n-. - Nasalização das vogais entre consoantes nasais, depois de pausa e antes de consoante nasal.
Português	[a, e, i, o, u] → [ã, ê, î, ô, û]	- Perda da nasal latina intervocálica -n-. - Nasalização das vogais, independente do registro na superfície.

### 3.5. Alongamento vocálico

Esse fenômeno consiste em se pronunciar as vogais com maior duração, sendo alongadas, quando estão em posição tônica. É importante salientar que a realização das vogais tônicas em espanhol tem menor duração, por esse motivo os aprendizes brasileiros de espanhol devem observar a duração da vogal tônica em espanhol, tentando evitar ou transferir o alongamento vocálico do português.

Alongamento vocálico		
REPRESENTAÇÃO ORTOGRÁFICA	PRODUÇÃO DE LE	PRODUÇÃO DO ESTUDANTE DE LE
	[a, e, i, o, u]	[a, e, i, o, u] → [a: e: i: o: u:]
Cantar	[kan'tar]	[kan'ta:r]
Beber	[be'ber]	[be'be:r]

Temos o conhecimento histórico que a duração das vogais era um dos elementos distintivos das vogais no latim clássico, que na passagem para o

latim vulgar sofreu a perda da quantidade e mudou para a distinção das vogais através da tonicidade.<sup>11</sup> “El latín clásico contaba con dos subsistemas vocálicos – el de las vocales largas y el de las breves -, y, por lo tanto, con diez fonemas de este tipo. Esta diferenciación cuantitativa se perdió en el paso del latín al romance” (ARIZA, 2016: 22)<sup>12</sup>. Observa-se que no latim as vogais tiveram a duração como componente fonológico, não aparecendo hoje o alongamento vocálico como um componente inovador, mas como um processo prosódico que emite sentidos.

Sobre o alongamento vocálico, podemos citar a comprovação de Brisolara e Semino (2016), que conferiram a tendência à maior duração das vogais em posição tônica tanto em português quanto em espanhol. No entanto, entendemos que o alongamento vocálico sucedido no português tem uma duração maior que o espanhol. Ao estudar a língua espanhola, os aprendizes brasileiros habituam-se a sofrer a interferência da língua materna ao realizar a transferência da duração da vogal tônica para a língua estudada, conforme explicam Brisolara & Semino (2014).

Brandão (2003, p. 177), ao tratar do alongamento vocálico cotejando a língua portuguesa à língua espanhola, ratifica que as vogais em sílaba tônica em português têm uma maior duração. Compreendemos que a duração é uma característica do português, que pode também ser justificada pelo ritmo da língua, considerando-a como uma língua acentual, na qual se aplica a maior intensidade e duração na sílaba tônica, enquanto em espanhol, o ritmo silábico, característico do espanhol, divide a duração das sílabas igualmente. De acordo com Ramírez (1998:22), “El alargamiento vocálico, presente en las vocales tónicas de muchas regiones, y de acuerdo con el sistema de la lengua española, no tiene valor fonológico, sino expresivo o enfático”.<sup>13</sup> O autor também afirma sobre alongamento vocálico estar presente nas Antilhas, Chile e México, como traço estilístico.

---

<sup>11</sup> Explicações sobre a perda de quantidade vocálica latina podem ser verificadas em Ariza (2016: 23-24).

<sup>12</sup> Tradução nossa: “O latim clássico contava com dois subsistemas vocálicos – o das vogais longas e o das breves - e, portanto, com dez fonemas deste tipo. Esta diferenciação quantitativa se perdeu na passagem do latim ao romance”.

<sup>13</sup> Tradução nossa: “O alongamento vocálico, presente nas vogais tônicas em muitas regiões, e de acordo com o sistema da língua espanhola, não tem valor fonológico, mas sim expressivo ou enfático”.

<b>Processos fonético-fonológicos do alongamento vocálico</b>		
Espanhol		- Perda da quantidade das vogais latinas. - Ocorrência branda na língua espanhola.
Português	[a, e, i, o, u] → [a: e: i: o: u:] > ā, ē, ī, ō, ū.	- Perda da quantidade das vogais latinas. - Deriva da língua indutiva para a neutralização das vogais átonas em posição final.

O alongamento vocálico no português significa uma interferência, sendo uma realização inadequada para a fala espanhola, tornando manifesta a detecção de traços de inabilidade linguística por parte do aprendiz. A finalidade não é suprimir traços identitários do estudante, mas aprovisionar mais dados sobre a língua almejada.

## 4. CONCLUSÕES

Os erros fonético-fonológicos atuais percebidos em trabalhos relacionados a essa área de estudos, demonstram as decorrências dos processos linguísticos em um panorama cronológico de movimentação constante com a finalidade de definição de uma representação fonológica que suportasse as alterações. Portanto, podemos assegurar que as modificações fonéticas que ocorrem em determinadas palavras durante a expansão do latim até a formação das línguas românicas são ainda nascentes de movimentação subjacente e superficial da língua, que foram observadas nesse estudo diacrônico e sincrônico e, além disso, perduram na língua usada pelos seus falantes, como a perda de quantidade das vogais que decompueram o sistema acentual das línguas neolatinas e aspectos que, apesar de estarem submersos, conservam-se na língua, como, por exemplo, a nasalização de ditongos.

Há toda uma estrutura amparada pela tradição fonética e fonológica, principalmente binária, que fixa os espaços dos sons dos sistemas linguísticos, que ao mobilizar-se recomendam a concepção de sons substitutos dos ambientes sonoros antes tomados por outros reminiscetes. Percebemos que os erros dos estudantes brasileiros vêm exatamente assinalar esses lugares que estão ocasionando dificuldades e que realmente foram ou são problemas da história linguística. Essa percepção através do olhar histórico evidencia os alunos como parte do processo de mutação linguística, vivenciando as transformações ocorridas desde o latim vulgar na aprendizagem de uma língua estrangeira.

Por isso, podemos assegurar que os erros fonético-fonológicos de estudantes brasileiros de espanhol são resquícios dos processos linguísticos sucedidos

durante a evolução histórica, tais como: abertura das vogais médias, harmonia vocálica, neutralização das vogais átonas em posição postónica final, nasalização das vogais, alongamento vocálico etc. Todos esses processos são ocorrências da língua portuguesa e ao mesmo tempo são procedimentos incorretos na fala de aprendizes da língua espanhola. Isso quer dizer que, os desencontros oriundos do latim vulgar na mutação das línguas românicas indicaram dificuldades de aprendizagem de espanhol por falantes brasileiros de português, pois os erros encontram-se, justamente na separação das línguas estudadas.

Esses fenômenos de reestruturação das línguas não são apenas incidências de circunstâncias externas, encontros com línguas desiguais e reformulações sociais, pois esses acontecimentos são confirmados nos erros dos aprendizes brasileiros ao aprender a língua espanhola, por se apresentarem precisamente nas dificuldades que os alunos têm em comum, inaptidões precisas na efetivação de mecanismos articulatórios que beneficiam as inadequações na oralidade.

Todos os fenômenos analisados nesse estudo estão correlacionados aos fatos históricos fonético-fonológicos que armazenam as implicações decorrentes das transformações linguísticas sucedidas pelos séculos. Há uma ocorrência de modificações entre o latim e as línguas românicas que estruturam as línguas e proveem sons e ajustes fonéticos para aparelhar os inventários fonológicos. Esses processos permanecem visíveis nas inadequações orais produzidas pelos estudantes de espanhol, falantes do português brasileiro. Por isso, podemos afirmar que há um entrelaçamento entre os erros fonético-fonológicos dos alunos brasileiros estudantes de espanhol e os fenômenos linguísticos passados.

## REFERÊNCIAS

ARIZA, Manuel. **Fonología y fonética históricas del español**. Madri: Arco/libros, 2016.

BARBOSA, Maria Silvia Pereira R-A. **A nasalidade vocálica em português e em espanhol**. 2014. Tese (Doutorado) – Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências e Letras (Campus de Araraquara), 2014. Disponível em: <http://hdl.handle.net/11449/123254>. Acesso em: 22 dez. 2018.

BATTISTI, Elisa; VIEIRA, Maria José B. O sistema vocálico do português. In: BISOL, Leda (org.). **Introdução a estudos de fonologia do português brasileiro**. Porto Alegre: EDIPUCRS, 2005.

BAZZAN, M. A. T. **As vogais médias na interfonologia português-espanhol**. 2005. Dissertação (Mestrado em Linguística Aplicada), Universidade Católica de Pelotas, Pelotas.

BEUX, E. B. S. **Descrição da interlíngua português-espanhol no desempenho de formandos dos cursos de turismo e hotelaria do IFAL – Campus Maceió: o aporte das teorias Linguísticas e pressupostos de ensino-aprendizagem de LE**. 2014. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade Federal de Alagoas. Maceió.

BISOL, L. **Harmonização vocálica: uma regra variável**. 1981. Tese (Doutorado em Linguística). Universidade Federal do Rio de Janeiro, Rio de Janeiro.

BISOL, Leda. **A harmonização vocálica como indício de uma mudança histórica**. 2015. Disponível em: <http://www.scielo.br/pdf/delta/v31n1/0102-4450-delta-31-01-00185.pdf>. Acesso em: 17 fev. 2019.

BRANDÃO, L. R. **Yo hablo, pero... ¿quién corrige? a correção de erros fonéticos persistentes nas produções em espanhol de aprendizes brasileiros**. 2003. Dissertação (Mestrado) Universidade Estadual de Campinas, São Paulo.

BRISOLARA, L.; SEMINO, M. J. **¿Cómo pronunciar el español? La enseñanza de la fonética y la fonología para brasileños: ejercicios prácticos**. Campinas, SP: Pontes editores, 2016.

CAGLIARI, Luiz Carlos. **Análise fonológica**. Campinas, Edição do autor, 2002.

CÂMARA JR., J. M. **Estrutura da língua portuguesa**. 46º ed. – Petrópolis, Rio de Janeiro: Vozes, 2014.

CÂMARA JR., J. M. **História e estrutura da língua portuguesa**. 2º ed. – Rio de Janeiro: Padrão, 1979.

CÂMARA JR., J. M. **Para o estudo da fonêmica portuguesa**. 4º ed. – Rio de Janeiro: Padrão, 1977.

CARVALHO, Kelly Cristiane H. P. de. **Descrição fonético-acústica das vibrantes no português e no espanhol**. 2004. Tese (doutorado) – Faculdade de Ciências e Letras de Assis, Universidade Estadual Paulista, 2004.

COSTA, Rossana S. **A produção da lateral /l/ por alunos de espanhol/le**. 2013. Dissertação (Mestrado em Linguística Aplicada). Centro de Humanidades, Universidade Estadual do Ceará, Fortaleza.



DIAS, Eva, C. O. **Produção do yeísmo em falantes nativos de espanhol e aprendizes brasileiros: um estudo a partir de análises acústicas**. 2011. Dissertação (Mestrado em Linguística), Universidade Federal de Santa Catarina, Florianópolis.

FERNÁNDEZ, J. G. **Panorama de la fonología española actual**. Madrid: Arco/libros, S.L, 2007.

GUIMARÃES, F. I. **Aquisição do português como L2 por falantes de espanhol: uma experiência com o modelo de ontogenia**. 2011. Dissertação (Mestrado em Filologia e Língua Portuguesa), Universidade de São Paulo, São Paulo.

HERRERO, Maria Antonieta A. **Variedades del español de América: una lengua y diecinueve países**. Brasília: Embajada de España. Consejería de Educación, 2004.

HOYOS-Andrade, R. E. **El vocalismo del español y del portugués: estudio contrastivo fonético-fonológico y sus implicaciones pedagógicas**. 1978. Disponível em: <http://www.revistas.usp.br/linguaeliteratura/article/view/138142>. Acesso em: 15 jan. 2018.

MACHRY DA SILVA, S. **Aprendizagem fonológica e alofônica em L2: percepção e produção das vogais médias do português por falantes nativos do espanhol**. 2014. Tese (Doutorado em Letras), Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

MARTÍNEZ CELDRÁN, E. & FERNÁNDEZ PLANAS, A. M. **Manual de fonética española**. Barcelona; Ariel, 2007.

MASIP, V. **Fonologia, fonética e ortografia portuguesas**. Rio de Janeiro: E.P.U., 2014.

MASIP, Vicente. **Gramática española para brasileños: fonología, ortografía e morfosintaxis**. São Paulo: Parábola, 2010.

MASIP, Vicente. **Gramática histórica portuguesa e espanhola: um estudo sintético e contrastivo**. São Paulo, EPU, 2003.

MENDONÇA, Ana Maria Santos de. **A nasalização fonética de vogais átonas em Alagoas: uma análise sociolinguística**. 2019. Tese (Doutorado em Linguística) – Faculdade de Letras, Programa de Pós-Graduação em Letras e Linguística, Universidade Federal de Alagoas, Maceió, 2019.

MIGNONI, R. P. L. **A transferência e a aquisição das vogais espanholas /e/ e /o/ em substantivos e adjetivos por falantes universitários brasileiros**. Dissertação (Mestrado em Letras), Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999.

MILANI, E. **Gramática de espanhol para brasileiros**. São Paulo: Saraiva, 2006.

OLIVEIRA, Aline V. B. H. **Estudo Fonético-Fonológico Contrastivo entre a Língua Portuguesa Falada no Brasil e a Língua Espanhola**. 2009. 101 f. Dissertação (Mestrado em Linguística). Faculdade de Letras, UFAL, Maceió.

OLIVEIRA, Aline V. B. H. **Análise de erros fonético-fonológicos do processo de ensino-aprendizagem de espanhol como língua estrangeira por brasileiros baseada na relação com as mudanças históricas das línguas**. 2020. 155 f. Tese (Doutorado em Linguística). Faculdade de Letras, UFAL, Maceió.

QUILIS, A. Comparación de los sistemas fonológicos del español y del portugués. *In: Revista de filología española*. Madrid, 1979, nº 68: 33-65.

RAMÍREZ, María V. **El español de América I: Pronunciación**, Madri, ARCOS/LIBROS S.L, 1998.

SANDES, E. I de A. **Estágios de interlíngua: estudo longitudinal centrado na oralidade de sujeitos brasileiros aprendizes de espanhol**. 2010. Tese (Doutorado), Universidade de São Paulo, São Paulo.

SANTOS. G. R. dos. **Percepção e produção das vogais médias do espanhol por falantes do português brasileiro**. 2014. Tese (Doutorado em Letras), Universidade Católica de Pelotas, Pelotas.

SILVA, Kátia C. D. da. **Ensino-aprendizagem do espanhol: o uso do interlingüístico das vibrantes**. 2007. Dissertação (Mestrado em Linguística). Universidade Federal do Ceará, Fortaleza.

SILVA, T. **Fonética e fonologia do Português: roteiro de estudos e guia de exercícios**. São Paulo, Contexto, 2005.

TEYSSIER, P. **História da Língua Portuguesa**. São Paulo: Martins fontes, 2007. Trad. Celso Cunha.

TOMAS, T. N. **Manual de pronunciación española**. Madri, grafipen, 1989.



## EXPLICIT INSTRUCTION OF POST-LEXICAL PROCESSES TO AID LISTENING COMPREHENSION IN THE TEACHING OF ENGLISH AS AN ADDITIONAL<sup>1</sup> LANGUAGE

*Adriana Lisboa Tibana<sup>2</sup>*

“Syllables, words, sentences run together like a watercolor left in the rain.”

Bill Bryson

### 1. INTRODUCTION

Assuming that post-lexical processes significantly affect the English spoken language when in connected speech, and that this is one of the main reasons why students of English as an additional language find listening comprehension so difficult to cope with, this study investigated the connection between listening comprehension and the teaching of pronunciation using a top-down approach. That is to say, starting from the whole picture and based on the organization of the language in chunks in the fluidity of speech. This is an action-research, conducted with two groups of upper-intermediate students during one semester. It was based on the assumption that by raising students' awareness of post-lexical processes, with the support of the Lexical Approach

---

<sup>1</sup> I have chosen to use the term *additional language* rather than *foreign* or *second language* for the same reasons given by the International Academy of Education (IAE) “The term ‘foreign’ can, moreover, suggest strange, exotic or, perhaps, alien—all undesirable connotations. Our choice of the term ‘additional’ underscores our belief that additional languages are not necessarily inferior nor superior nor a replacement for a student’s first language.” (JUDD, Elliot L.; TAN, Lihua; WALBERG, Herbert J. 2001, p.6)

<sup>2</sup> Universidade Federal de Alagoas.

(LEWIS, 1994, 1997) which encourages teaching language in chunks, teachers can help students to be better listeners.

By acknowledging the differences in form and character between spoken and written English, teachers could be helping learners to become aware of the fact that in the transition from written to spoken, English somehow turns into another language. Not only do pronunciation and spelling fight a continuous battle to dictate the norms upon which the words should adhere to, but also, and most importantly, connected speech processes significantly affect spoken language. This operation of merging word boundaries, to produce the stream of speech, gives rise to the rupture of the English language into a spoken and written form<sup>3</sup>. The written language is the one where word boundaries dictate the norms, whereas the spoken language does not conform to those rules as its pace needs to be much faster: words running over subsequent ones rendering a sort of domino effect.

This ‘fluidity’ in spoken language is produced by post-lexical processes<sup>4</sup> - such as elision, assimilation, devoicing, lenition, insertion, among others - as well as change in stress, rhythm, intonation, and a variety of factors which make words pronounced in isolation rather different from when they are in the ‘stream of speech’. However, these factors are hardly ever shown to students learning an additional language. Our failure to do so is perhaps rooted in our strong tendency to think of language as written.

It is also noticeable that among the four skills – listening, writing, speaking and reading - listening tends to be rated by a great number of students as the most difficult to acquire. Most students struggle to deal with the idiosyncrasies of the spoken language as opposed to the written one. These peculiarities can be produced by the linking devices that engender the interwoven units of connected speech.

There is a gap, which must be bridged, between the teaching of pronunciation for spoken purposes only, and the teaching of listening with more emphasis on pronunciation features. These features affect not only students’ ability to put

---

<sup>3</sup> Written and spoken language also differ in other features such as grammar and lexis.

<sup>4</sup> According to Spencer (1996, p.201) “Post-lexical processes are phonological processes which are triggered solely by phonological structure, and which thus do not have lexical exceptions or morphological conditions. For this reason they are sometimes called automatic processes. Many of these processes operate across word boundaries or are affected by the phonological structure of a whole phrase, so they are often referred to as connected speech processes or phrasal phonology.”

their intended message across but also their ability to understand spoken English. This means that teaching pronunciation is much more about raising students' awareness of the idiosyncrasies of the spoken language, and enabling them to understand and produce intelligible language, rather than trying to make them sound like native speakers.

Moreover, most pronunciation teaching traditionally tends to be mainly at the segmental level, that is to say, teachers focus on phonemes and their articulation. For a long time pronunciation activities were limited to the teaching of contrast using minimal pairs (e.g. the contrast between *ship* and *sheep*) and word stress; as if being able to pronounce words in isolation accurately was all a learner needed to know as far as pronunciation is concerned. More recently, however, this focus has shifted to some more work at a supra-segmental level, including sentence stress and intonation. Such activities, which are frequently found in modern coursebooks together with work on phonemes and word stress, comprise the stock-in-trade of English pronunciation teaching. It is the 'whole picture', however, that has been left aside, as little work on the features of connected speech seems to have been done.

## 2. CONNECTED SPEECH

Connected speech is defined here as naturally occurring speech, language that is constructed as it is gradually delivered, "the inescapable fact of the real-time, step-by-step assembly of a spoken utterance" (BRAZIL, 1995, p.17). It is real language in action, it happens when the words run together in a string of spoken language, breaking their boundaries, turning into the amorphous mass of speaking. This 'blurring' of word frontiers establishes the major differences between spoken and written language. It is this considerable contrast that most teachers of additional languages fail to point out. As a result of such neglect, students' view of language tends to be focused on its written form only. (CRYSTAL, 2002)

The phonological processes that are the result of such rapid changes of the vocal organs when the words run together are: assimilation, elision, vowel shortening, linking and intrusive sounds, juncture, among others. These post-lexical processes are certain to occur in both formal and informal unscripted speech. According to Brown (1990, *apud* Shockey, 2003) there is little difference in phonological structure in different styles. The main difference between those two spoken registers are the lexical choice and the 'intensity', because informal speech tends to be less careful thus rendering more simplifications. This is so because such processes are automatic, working on a subconscious level.

The rate of delivery of the utterance, as one may expect, surely affects this fluidity of speech. However, according to Shockey (2003, p.14 to 19) other factors influence casual speech reduction, namely:

- Frequency - the more frequent a word is, the less it needs to be clarified phonetically to achieve communication status;
- Discourse – first mentions or focal mentions of a lexical item tend to be given a more complete articulation than its following repetitions;
- Syntactic functions – pronouns often show more reductions than nouns; short, frequent function words (of, and) are more likely to be shortened than longer ones (moreover, nevertheless);
- Morphological classes – in some dialects the morphological class of a word affects its realization;

It should be added to the above categories, the pronunciation of chunks, which are lexical items, “socially sanctioned independent units” (LEWIS, 1994, p.90) that could be made of one or lots of words together – thus belonging to discourse. Such expressions are used so often that sometimes their individual parts have got completely lost in its pronunciation form, having turned into an ‘unintelligible’ amorphous number of phonemes. This may leave non-native speakers completely at a loss, unless they have these expressions in their repertoire, and know their pronunciation as a chunk.

## 2.1 Rhythm and intonation

The English language overall rhythm is dictated by the placement of the nuclear stress, which holds the principal change of pitch<sup>5</sup>, and an alternation of stressed and unstressed syllables. Stressed syllables are louder, longer, and with a higher pitch (thus more prominent) than the unstressed. The result is that stressed syllables tend to be clearer and are realized in their full phonetic value, that is to say, they are pronounced more comprehensibly whereas unstressed ones are shortened and tend to have a more obscure overall quality. This characteristic is common to each and every variety of L1 English. (UNDERHILL, 1994)

---

<sup>5</sup> Underhill explains pitch (1994, p.76) as such: “The vocal chords vibrate during speech. This vibration is heard, and the pitch of this sound varies according to the frequency of the vibration of the cords: the higher the frequency of vibration the higher the pitch that you hear. When you sing a pitch or note you usually hold it for a time before jumping or sliding to the next note. But in speech the pitch of your voice varies continuously so that your speech is not heard as a tune. This pitch variation extends over single phonemes, sequences of phonemes, and whole utterances.”

Because of its overall rhythm, English tends to be called a stress-timed language as opposed to syllable-timed Portuguese. This division is a rather controversial issue that creates different opinions among experts. The idea of a stress-timed language comes from the belief that those languages follow some patterns that reoccur in established intervals of time. According to McCarthy (1996, p.91): “While this distinction may correspond to some strongly felt perception of the different characteristic rhythms of languages, there is little hard instrumental evidence for it. In fact, in recent years, quite a lot of convincing counter-evidence has been presented”. As a matter of fact, research in the area shows that the so-called stress-timed languages are not more rhythmical than the syllable-stressed ones. This feeling of ‘rhythmicality’ that the English language evokes may be due to the fact that one of its distinctive features is the contrast between stressed and unstressed syllables.

## 2.2 Word and Sentence Stress

When we talk about intonation two things arise: word stress and sentence stress. Together they dictate the rhythm of the language. Word stress or accent is defined by the language and therefore is totally out of the control of the speaker. Changing word stress will only make the words incomprehensible and break the communicative flow. In fact that is one of the main sources of comprehensibility break down.

Moreover, as Shockey (2003, p.16) points out, English is a ‘topic-comment’ language, that is, “the old information comes first, followed by the new.” Thus, the nuclear stress tends to fall towards the end of the utterance. This triggers another phenomenon which affects connected speech: the beginning of the utterance, not carrying the nuclear stress, tends to be spoken more quickly and less clearly than the end, unless the speaker wants to emphasize a word to convey specific information (SHOCKEY, 2003). This intelligibility problem contributes to increasing students’ anxiety as they panic, because they cannot cope with the beginning of the utterance, therefore, losing concentration and missing the whole thing. Hence, it is important to make students aware of this fact, so that they know that the message tends to be clearer towards the end, and that the ‘somehow obscure beginning’ is related to some information he/she already holds.

One cannot talk about stress without mentioning weak forms. They are the most unstressed part of the utterance: function words such as articles, auxiliary verbs, prepositions, etc. They are words which do not carry content information.



They might as well be left out like the words excluded in old-fashioned telegrams. In spite of this, the majority of students tend to give full value to those words, producing some spoken language that sounds formal and sometimes stilted. Moreover, when listening, they expect spoken language to do the same and end up lost when familiar words sound alien as they lose their strength.

Students need to be made aware of them and thus train their brains not to expect to hear every single word with the same strength in the fluidity of connected speech. They also need to know that they can get back their strength in certain situations. Much more than including this feature in their own speech, students need to recognize weak and strong forms, as failure to acknowledge them in connected speech is likely to cause communication breakdown. In fact, it is one of the main sources that lead to incomprehensibility, contributing to greater strain when trying to understand spoken language. As Brown (1990, p.56) points out: “From the point of view of the comprehension of spoken English, the ability to identify stressed syllables and make intelligent guesses about the content of the message from this information, is absolutely essential”.

### 2.3 Post-Lexical Processes

There is a close relationship between words in a string of speech: they influence each other, conjoin to become almost one single word, let intrusive sounds come between them, reduce length, get stressed, elide the last sound of others, thereby interacting to form the fluid mass of spoken language. This relationship is somehow dictated by post-lexical processes which are, as Spencer (1996, p.200) states, automatic, entirely triggered by phonological structure.

Such processes are: assimilation - when sounds modify the quality of a neighboring phoneme in transforming some of its distinctive features<sup>6</sup>, elision - when sounds are deleted or elided, vowel reduction, liaison - when, as the name suggests, we link words or when an alien sound is introduced to smooth the connection between words in connected speech, intrusive sounds, juncture - when two words come together in such a way that it is nearly impossible to determine where one ends and the other starts, among others<sup>7</sup>. They are mostly responsible for the problems students face when trying to understand spoken language, especially when two or three of them happen at the same time, something that is bound to occur very frequently.

---

<sup>6</sup> Distinctive features are a set of characteristics inherent to a phoneme, such as: nasal, voiced, bilabial (produced bringing the two lips together), etc.

<sup>7</sup> I did not mention some of these processes such as aspiration, clear/dark /l/, retroflexion,

### 3. LISTENING COMPREHENSION

In real life, according to Doff (1993, p.198), there are basically two kinds of listening: ‘casual’ or ‘focussed’ listening. This author explains that in the former, we do not listen very closely and do not have any specific reason to do so. Thus, our attention span varies depending on what is being said. If it interests us, we concentrate more. Otherwise, we just skim through what is being said for the gist. This is the kind of comprehension we get when we are chatting to a friend or listening to the radio, or the television while doing something else. As for the latter, we listen with a purpose and therefore concentrate on the important information we want to get from it. We scan through what is being said to suit our purposes, be it for communication or to retain some important information.

In both cases, according to Rost (1990, p.33), we go through two stages in our way to understanding the message: the physical aspect/auditory perception, and the cognitive aspect/linguistic processing. Those stages happen almost simultaneously, because in the process of listening, there is not much time to process information as the flow of new speech acts keeps coming. Such stages require as much of the listener’s active performance as the speaker’s. Thus, we cannot call the listening skill a receptive skill only. It is the listener’s active processing of information that makes him/her understand the spoken language.

As Underwood (1993, p.2) states, although the act of listening takes place fairly quickly, it goes through stages: first the utterance is taken in by the listener and organized into chunks which remain in the ‘echoic’ memory for no more than a second. Following that, this raw material which has been roughly ‘processed’, is stored in the short-term memory. At this point, if the listener had not been able to segment the speech into meaningful chunks, he/she may not manage to take in the next piece of information that comes right after that. This segmentation of chunks is aided by the listener’s knowledge of the structure of the language, lexical range, and familiarity with the ways these chunks are pronounced in spoken language. Rost (1999, p.38) explains:

Listeners who anticipate hearing ideal pronunciations of words will have considerable difficulty in decoding connected speech since **all phonemes change their perceptual features in different phonetic environments**. As such, any ideal phoneme is an unrealistic standard against which to match heard forms. (ROST, 1999, p.38)

---

among others, because they do not influence much intelligibility as far as listening comprehension is concerned.

This is one of the major sources of oral communication breakdown when students are expecting to hear the words as if read aloud from a written text. As students' expectations do not match what is actually uttered, there is a battle between their knowledge of the language, which is based on the written form, and the spoken language that is being produced. This continual battle prevents students from understanding, rendering frustration and demotivation. The listening comprehension process is thus blocked before it can be analyzed for the gist or specific information and transferred to the long-term memory.

This process of analyzing information, when it is only raw material in the short term memory, is aided or hindered by the listener's ability to recognize chunks of language. "As the learner gets more used to listening, and has at the same time learned more of the language, he/she can process some often-heard chunks more or less automatically, thus leaving 'space' and energy to deal with the more difficult or less familiar input." (UNDERWOOD, 1993, p.2) That is when the main principle of the Lexical Approach (LEWIS, 1994) comes into play: language is not produced from scratch. We are not as creative as we tend to think, we use a wide range of formulaic set expressions to communicate. By raising students' awareness of those expressions, as well as the way they are pronounced in connected speech, we are equipping students with powerful tools to be better listeners.

Another essential factor to take into account is that throughout the process of listening comprehension, we 'hear' much more than it is actually there, especially those formulaic expressions. Some of these expressions, being so often uttered, have become an amorphous mass. We successfully make them out only because we understand the context and know in advance that they would be appropriate there. As Lass (1984, p.296/297) explains:

[...] the listener applies, in a CS [casual speech] situation, all his knowledge of linguistic structure: syntax and morphology, semantics and lexis – as well as pragmatic cues – and finally, of course phonology. And he approaches the task of interpretation, under normal conditions, with the expectation that messages make sense, and he does his best to ensure it. [...] So the speaker listens not to what it is - strictly – being said, but to what his knowledge of the basic structure of the language tells him ought to be being said. (LASS, 1984, p.296/297)

Taking this into consideration, we can somehow understand an advanced learner's problem whose command of the language is very good, but has trouble with listening comprehension. It may be that, in spite of the fact that their syntactic, morphological, semantic, lexical, and pragmatic knowledge of the target language is considerable, their phonological understanding of the spoken language

needs improvement. As Rost (1999, p. 33) says: “the intelligibility of what is heard is reciprocally linked with the interpretability of what is heard through cognitive effects as the listener attends to speech.” Thus, we cannot separate the two stages - the auditory perception and the linguistic processing - as one merges into the other throughout the whole process of understanding. As in connected speech, there are no boundaries: one cannot identify where the physical process ends or the cognitive starts.

Another significant factor to include is the co-text, as Underwood (1993, p.3) says:

One important part of this overall situational context in which the listener places what he/she hears is, in Brown and Yule’s terms, the ‘co-text’, which they define as ‘whatever has already been said in a particular event’. It is by placing what follows in relation to what has already been said (either by the current speaker or by other speakers) that the listener establishes the speaker’s meaning. (UNDERWOOD,1993)

Having said that, if asked what makes listening difficult, most students are bound to say that people speak too fast so they cannot understand each and every word as they would like to. In fact, this was what a great number of students from this research said when asked this question. They also mentioned their inability to cope with familiar words in connected speech, problems with lexis (especially idioms and slang), and difficulty in concentrating and remembering what had been said. Only 5.2% of the students mentioned pronunciation and 2.6% accent.

What students do not realize is that most of the problems they mention have an intrinsic relationship with pronunciation. The impression that native speakers speak too fast, is due to pronunciation features, such as assimilation, liaison, weak forms, elision, etc, those simplifications of the spoken language which give the idea that connected speech goes at an incredibly fast pace. As for their inability to cope with familiar words in connected speech, this is also explained with the help of phonology, especially as there is a tendency for words to appear in spoken language in chunks. Formulaic expressions that have been turned into a block of spoken language sound very different from when their components are spoken as isolated words.

When not aware of the fact that language is produced in chunks, students tend to try to hear each and every word. This can lead to loss of concentration and tiredness. Talking about this problem, Underwood (p.19, 1993) says that “sometimes, even when the topic is interesting, students simply find listening work very tiring, because they make an enormous effort (often greater than is useful) to follow what they hear word by word.” It is the role of the teacher to

raise students' awareness of such chunks and formulaic expressions, and develop students' ability to recognize them, even when part of it somehow disappears in connected speech.

Apart from these factors mentioned above, there are certainly others that can hinder listening comprehension, such as accent, register, regional idioms, jargons, lack of cultural knowledge, concentration, among others. There is also the amount of information as opposed to the amount of time one needs to process such data, and the fact that spoken language is not always well organized. However, as this research focusses mainly on pronunciation, these factors are not discussed here.

## 4. METHODOLOGY

This is a qualitative action-research of pedagogical intervention and of an experimental type. The data for this study was collected from the tests, questionnaires, interviews, some exercises from the English File Upper-intermediate coursebook<sup>8</sup>, and activities especially devised to raise students' awareness of the post-lexical processes in the English language. These activities were applied in lessons taught in one semester to one of the two chosen groups.

To establish the groups' level of proficiency in the listening skill, the Cambridge First Certificate in English (FCE)<sup>9</sup> listening test was used as a diagnostic test. This is one of the tests of the main suite of international exams devised by the University of Cambridge – Cambridge Assessment. One such test was carried out at the beginning of the semester, and two others at the end in order to attempt to assess students' performance development. Cambridge produces past examination papers to be used for practice when preparing students for their exams, the listening tests to which students were submitted were taken from such papers. The choice of an FCE listening test was due to their internationally respected standards in preparation and production, ensuring consistency in level between the first and the last test in this research.

They did the whole listening test, however, for the subject of this research, the grades were given considering their performance in Part 2 only. This is because this part tests candidates' abilities to listen for specific information whereas the others focus on gist and also assess students' interpretative skills. It

---

<sup>8</sup> English File Upper-intermediate, (OXENDEN and LATHAN-KOENIG, 2001). From now onwards when students' course book is mentioned, that is the one.

<sup>9</sup> This exam is now called B2 First.

is in part two that candidates need to have the ability to break up the fluidity of connected speech into chunks to find specific information. The straightforward simplicity of filling gaps with information that comes in the same order of the text makes part two a very reliable assessment tool to measure students' abilities to understand connected speech. Such information is precise, if they do not manage to understand the very moment it is uttered, they will not have another chance. It is pure understanding of words embedded in connected speech, and it is considered by the majority of the students who are preparing for the test, the most difficult part. Statistically, at least from my experience with this test, this part is the one which students score the lowest.

The students in this research were in their first year of the upper-intermediate course, which means at least two years behind the required level to do the FCE exam. Undoubtedly, this would be a challenge for them. Therefore, their performance was expected to be weak, as scoring low at the beginning would make room for improvement at the end.

The description of the groups and the students came from a questionnaire which they answered after taking the listening test. The information about the questions some of the students left unanswered was collected on an individual basis in subsequent lessons. At the end of the semester, students from the experimental group gave a short, filmed interview, and then talked about what they had gained from the experiment.

Two groups took part in this research: the Experimental Group, with whom I tried the experiment in order to test my hypothesis, and the Control Group with whom I worked in following the coursebook without applying any activity related to pronunciation awareness connected to listening skills. I tested both groups listening abilities at the beginning and at the end of the semester. The results of the first test determined the choice of the group I was going to apply the experiment to: I chose the one which was weakest at listening. It also helped measure the groups' listening abilities (diagnostic test), and provide data to be used at the end of research when comparing students' improvement in listening comprehension.

The Experimental Group (19 students) attended lessons every Friday morning for two hours and ten minutes. In this group there were mostly women (79%) and the average age was higher than the Control Group. The majority of students (63.15%) were adults ranging from 22 to 37 years old. The Control Group (20 students) had lessons on Monday and Wednesday afternoon for one hour and fifteen minutes each. This group was younger, mostly teenagers, ages

ranging from 15 to 20 years old (80%), with a slightly higher number of men (55%) than women (45%).

For most of the students, English was their first and only additional language. There were two students (10.6%) in the Experimental Group who said they spoke Spanish. In the Control Group nobody spoke any other language apart from English and Portuguese (their native language). There was only one student studying another additional language (French) and at a pre-intermediate level. There were only three English teachers<sup>10</sup>, two in the Experimental Group and one in the Control one. The teachers were not experienced, and none of them had worked with phonology in their classes or had ever done a course which included pronunciation. Hence, we can say that in none of the groups students had recourse to previous knowledge of phonology or of another language to perform better when doing listening tasks.

This illustrates the lack of pronunciation teaching in general: in private language schools, teachers are either intimidated by the phonemic chart or too busy covering the syllabus, with the effect that pronunciation is left aside as an incidental issue, coming up only when a student mispronounces a word. In general schooling, the subject is not even tackled as the major focus is on reading skills. Moreover, being a topic considered difficult by a vast majority of teachers, it is, most of the time, deliberately avoided.

The lessons I taught with the Experimental Group took place within a semester, which made 13 meetings, not counting the regular tests and revisions, nor the first and last listening tests. These 13 meetings of two hours and ten minutes were all I had to cover a very tight timetable and incorporate a variety of pronunciation exercises which had not been officially included in the syllabus. The Control Group had the course standard lessons with the occasional pronunciation activities which are part of their coursebook. However, they were assessed with the FCE listening test at the beginning and the end of the semester for comparison with the other group.

The main difference between the lessons conducted with the groups is the fact that the Experimental Group was given explicit instruction on pronunciation matters to aid listening comprehension, whereas the Control Group was just exposed to the language without prior work on pronunciation. The term 'explicit' may give a false impression that the rules should be given to the students mechanically without much thinking. However, this explicit instruction can be

---

<sup>10</sup> I am considering teachers who taught English either at primary or secondary state or private schools, as well as teachers who gave private lessons.

done either deductively or inductively. Deduction is the most traditional<sup>11</sup> way where teachers present the rules before looking at the language itself. Induction is where students are encouraged to formulate those rules by themselves. This can be done through discovery activities with the aim of developing an awareness that will help students improve their understanding of how the language works by using their reasoning processes. This makes students understand that language is analyzable, as they can look for regularities themselves, and that they can benefit from their attempt to make sense of how it works.

The lessons taught to the Experimental Group took both a deductive and an inductive approach depending on the complexity of the phonological process involved. It was one of the aims of the study to find out to what extent this explicit instruction of some of the phonological features of connected speech, be it deductive or inductive, enabled learners to be better listeners.

Students in both groups did not have much contact with the English language outside the classroom, only 5.2% stated that they had someone in their immediate family with whom they spoke in English regularly. A total of four students (10.4%) did not answer this question. As for contact with native speakers, an overwhelming majority (94.8%) said that they never or hardly ever had face-to-face conversations or talked on the phone with them (100%). Apart from films, which they all watched with subtitles in Portuguese, songs and the occasional chat on the internet, mostly through writing, these students were in touch with the spoken language only when at the language school. Hence, they had to rely a lot on classroom time for exposure to the target spoken language, which made the development of their listening skills a challenging task.

Besides the minimum contact the students had with native speakers outside the classroom, their experience within an English environment was almost non-existent. Only 7.2% had ever been to an English-speaking country and only for a short time. The majority (87.6%) had never experienced being immersed in the target language culture or interacted with a native speaker in their environment. Their input came almost exclusively from films and songs, where they played the passive role of outside observers. Therefore, they could not have relied on such experience to improve their listening abilities.

Confirming the assumption that, in general, students find the listening skill the most difficult, 69.2% of all students stated that of the four skills listening was

---

<sup>11</sup> Although I used the term ‘traditional’ here, it is important to point out the fact I do not mean the memorization of rules disconnected from context, but the fact that those rules are made explicit to the students.



the hardest to master. Among these students, there were 50% in the control group, and a huge majority of 89.5% in the experimental group. As far as their listening competence was concerned, 74.4% of all the students did not consider their listening skills in the target language satisfactory, the figures being 94.7% for the experimental group as opposed to only 30% for the control group. This fact, together with the first FCE listening test result, determined the choice of groups when deciding with which of them the experiment would be carried out. It was clear that the students in the experimental group needed to work to improve their listening skills.

The fast delivery of connected speech was rated as the feature of spoken language that rendered it hard to understand by 36.8% of the experimental group as opposed to only 10% of the control group - 23.1% of all the students. Following this came the ability to understand every single word, together with lexis (idioms and slang), rated by 17.9% of the students in general.

None of the students had ever done any pronunciation course. In general, students' contact with pronunciation was restricted to the incidental explanation in class and some rudimentary overview of the phonemic chart from exercises on individual phonemes done in class throughout the course.

## 5. RESEARCH RESULTS

As previously mentioned, the results of the diagnostic test helped me decide on the group I should apply the experiment to. The group whose students had the lowest mark in the test was chosen. This group was henceforth called the *Experimental Group*. In fact, this group's overall performance in part 2 of the test was 5.6% as opposed to 27% in the other group - namely the *Control Group*. As for the Experimental Group overall performance in the whole diagnostic test, it was very poor, only 19.4% as opposed to 42.5% in the Control Group.

Looking more closely at the performance of both groups in part two, where students had to manage to single out 10 words in the string of speech, we can clearly see the difference in performance between them (see Table 5.1). The Control Group far outperformed the Experimental one. Considering that this test is way above their level – and that the average student should be ready to do it only after another two years of language lessons<sup>12</sup>, the performance of some students (35.29%) in the Control Group was outstanding<sup>13</sup>.

---

<sup>12</sup> This is due to the fact that they are not in an English-speaking environment.

<sup>13</sup> The average grade to pass this exam is 60%. The students who obtained 4 to 6 correct answers, four semesters short of taking the exam, and manage to keep those standards, would be expected to perform brilliantly in the actual test.

Table 5.1 – Comparison of groups’ performance in part 2 of the Cambridge FCE Listening Test.

GROUPS	Number of correct words out of 10						
	0	1	2	3	4	5	6 (pass mark)
Experimental (Number of students out of 16)	9 (56.25%)	5 (31.25%)	2 (12.5%)	-	-	-	-
Control (Number of students out of 17)	2 (11.76%)	3 (17.64%)	4 (23.53%)	2 (11.76%)	3 (17.64%)	1 (5.88%)	2 (11.76%)

As for students’ feedback after the test, there was not much difference between the two groups. Both groups considered the test either difficult or extremely difficult. Only two students in the Control Group found it not so difficult. However, the difference between the two groups in their attitudes towards listening was remarkable. The Experimental Group, in general, felt intimidated by it, whereas the Control Group, in general, did not feel much anxiety when having to face a listening task. This was the first change that I noticed in the Experimental Group. As the semester went by, this group was feeling more and more at ease with listening tasks. Maybe this was due to two different factors. First, I made them do more of these tasks than they had ever done before, and that forced familiarity which helped diminish their anxiety towards them. Second, by attempting to understand spoken language through pronunciation awareness exercises, students started to get over the feeling of helplessness towards a listening task.

From my experience as a language teacher, I have noticed that when pressed by a time constraint, teachers tend to skip listening tasks in favour of grammatical points that will certainly be covered in the test. Worse still is formal<sup>14</sup> pronunciation activities, which are completely forgotten when one does not have much time to cover the entire syllabus. Moreover, the time constraint is a big issue for those groups that meet only once a week. That was my first feeling of accomplishment, when I realized that I had managed to slot in those extra listening and pronunciation exercises in spite of being pressed for time. I then perceived that not only is it possible to include more listening activities in the lessons, but also some formal pronunciation teaching, without impairing the teaching of grammar and lexis. This was an issue that had worried me when I first thought about doing this experiment with this particular group.

<sup>14</sup> By formal I mean not incidental, as most pronunciation practice tends to happen because a student mispronounced a word or the teacher wants to focus on difficult words, etc.

In fact we managed time so well, due to students' cooperation as they had become so highly engaged in the project, that I had an entire lesson to administer the final FCE listening test with the Experimental Group. This was when I made my biggest mistake. I decided to award this group an episode of the sitcom *Friends* before doing the actual test. To make matters worse, I did so without letting them have subtitles in English. That was because I wanted them to feel that they were able to understand spoken English without the help of the written captions.

Students were quite happy to have understood a lot of the episode. However, after thirty minutes of undivided attention, as they made a great effort to try to understand everything, their attention span went down considerably, thus affecting their concentration when doing the actual listening test. I realized this when they started complaining about tiredness even at the beginning of the test. I should not have persuaded them to do this listening test. As for the Control Group, I had done the test fifteen minutes after the lesson had begun, just to give some time for the late comers to arrive.

The result of this first final listening test, as far as the Experimental Group was concerned, was rather discouraging, as students' performance, albeit improved, did not live up to my expectations. This group went from an overall performance of 19.4% to 28.3% in the whole test. In part 2 they obtained 17.8%, as opposed to 5.6% in their diagnostic test. I was disappointed because such improvement would be expected after one regular semester working with the groups in a traditional way, without much emphasis on pronunciation exercises. The only other noteworthy factor is the students' feedback on the test, as not many of them found the test extremely difficult: 43.75% of the students in the diagnostic test compared to 28.57% in the first final test.

As for the Control Group, their results were as expected, a slight increase in listening abilities after one semester's work. What is noticeable is that the students found the test less difficult than the first one: 64.7% had rated the diagnostic test as difficult, contrasted with 21.47% in this first final test. Comparing the 12.2% improvement in the Experimental Group (for part 2) with 11.6% in the Control Group, one may catch a glimpse of the importance of teaching pronunciation awareness exercises in order to improve listening skills. However, this is not enough to validate the claim that such exercises did help those students improve their listening comprehension. This was my first feeling when I calculated the results of this first final test.

Nevertheless, from my experience as a teacher, I knew that one of the facts that hinder students' performance in listening comprehension is lack of

concentration, which is something that is vital to the whole process. Motivated by this thought, and the incident that had happened to the Experimental Group, namely the loss of concentration because of tiredness, I decided to do a second final test<sup>15</sup>. This time I gave the test to both groups fifteen minutes after the beginning of the class. Not surprisingly, the results were significantly different.

Now the Experimental Group’s performance was outstanding, 45% in part two as opposed to 5.6% and 17.8% in the diagnostic and the first final test respectively. Looking closer at their performance, I could see that those pronunciation awareness exercises did play an important role in developing those students’ ability to understand spoken language. In part 2 of the second final test, 12 out of 16 students (an overwhelming 75%) performed satisfactorily. They got between 4 to 8 correct words out of ten (among those, three students got the pass mark or above), which, considering their level and the standard of the FCE listening test, is quite an accomplishment. As for the Control Group, although they performed slightly worse than in the first final test, the overall differences in part two were not substantial: 33.8% as opposed to 27% and 38.6% in the diagnostic and the first final tests respectively. Looking closer at their performance, what stands out is the fact that there was not a considerable change in their ability to understand spoken language. They maintained their level throughout the semester, improving only what would be expected.

This conclusion is even more evident when we compare both groups’ performance<sup>16</sup> (see Table 5.2), and their diagnostic test results to their best performance (see graphic 1). The difference between the two groups’ improvement is noticeable.

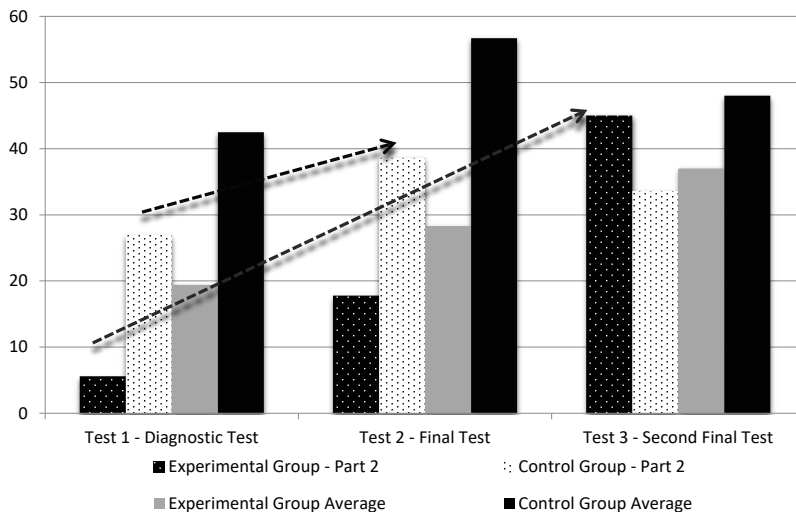
Table 5.2 – Comparing students’ performance between the diagnostic test and the average of the first and second final test.

GROUPS	DIAGNOSTIC TEST		FIRST and SECOND FINAL TEST AVERAGE		IMPROVEMENT	
	Part 2	overall	Part 2	overall	Part 2	overall
Experimental	5.6%	19.4%	31.4%	32.65%	25.8%	13.25%
Control	27%	42.5%	36.1%	52.35%	9.1%	9.85%

<sup>15</sup> It is important to point out that the tests which were administered with the groups were exactly the same, that is to say, the same diagnostic test to both groups, the same first final and the same second final test. By doing so, I was making sure students from different groups had exactly the same level of difficulty in each test.

<sup>16</sup> In order to be more accurate, I decided, to use the average grade of the first and the second test as a basis for comparison with the diagnostic test in spite of the fact that the results of the students in the Experimental Group may have been impaired by students’ lack of concentration.

Graphic 1 – Comparing students’ performance between the diagnostic test and their best performance in the final tests.



From the students’ recorded statements, it was evident that nearly all of them considered that this experience helped them develop their listening as well as speaking skills. It was also mentioned that those exercises helped reduce anxiety when facing a listening task. Another important point, touched on by a lot of students, is the fact that they were listening to ‘real language’ for the very first time. They said that this experience finally linked them to the way real language is spoken. Interestingly enough, their coursebooks are full of conversations, and its listening tasks are supposed to be based on real language.

Two different inferences may be drawn from this information: firstly, students somehow subconsciously understand that written and spoken language are different, so by becoming more aware of how spoken language works, they had the feeling of getting in touch with it for the very first time. Secondly, as Michael Lewis’ Lexical Approach claims, we are not teaching real language to our students, but some prescriptive language disguised in so-called communicative tasks. It is about time we started facing up to language the way it happens in real life, and stopped simplifying it unnecessarily.

### 5.1 Research Conclusions

The main aim of this study was to investigate the connection between listening comprehension and pronunciation. Based on the assumption that spoken

and written English are very different, and that most of the problems students have with listening comprehension is due to this difference, this study looked into phrasal phonology, searching for the main features that make word boundaries unrecognizable, thereby resulting in the blur of connected speech. It was assumed that this gap between spoken and written language, as far as listening comprehension is concerned, could be bridged by the explicit instruction of some pronunciation features of connected speech. Therefore, activities specially made to develop students' awareness of the underlying rules that govern streamed speech, were applied to a group of upper-intermediate students in order to find out if they could benefit from such activities to become better listeners.

When doing such pronunciation tasks, some comparisons were made between English and Portuguese, which is the students' mother tongue, concerning voice-setting, rhythm, intonation, syllables, and vowel and consonant sounds. By doing so, I wanted to find out if, with the help of students' L1 providing some familiar ground to fall back on for support, students would understand the phonological features of the target language better, thus turning mother tongue from hindrance to support.

Although not using specific exercises but following the main premise of the Lexical Approach, the teaching of language in chunks in tandem with their pronunciation was also investigated to see if it would help students deal more efficiently with the interwoven units of connected speech.

The results, in general terms, point to the teaching of pronunciation awareness raising exercises as a powerful tool to aid listening comprehension. The students from the Experimental Group, who were explicitly instructed on pronunciation matters, outperformed the ones in the Control Group, even though they had been the weakest at the listening skill at the beginning of the semester. Their results in the final listening test were outstanding.

Having said that, when looking back at the exercises I did with the students throughout the semester, I realize that although such activities helped students develop a greater awareness of the phonological system of the language, it was my attitude as a teacher that made all the difference. It helped students view the language from an entirely different perspective. By attitude, I mean the way I always incorporated a 'pinch' of pronunciation into every activity, be it lexical or grammatical. This may have helped students understand that language is holistic, and that pronunciation plays an important role in it. It was those short moments of explicit instruction together with the pronunciation exercises done in the lessons that promoted better understanding of the language.

I consider this conclusion, the fact that it is the teachers' attitude towards pronunciation that needs to be changed, as one of the most important findings from this study. This may sound rather simplistic or obvious, but what I mean here, is that no matter how many pronunciation exercises the teacher incorporates in his/her lessons, if he/she does not understand that the scope of pronunciation should encompass the whole language lesson - drawing students attention to stress, prominence, writing the spoken version of a new lexical item on the board in phonemic transcription, making pronunciation a living presence throughout the lesson - he/she will not be fostering students' understanding of the spoken language and its idiosyncrasies.

It is certainly the respect that the learners have for the written form that somehow prevents them from understanding that the spoken form takes a rather different shape. The use of the IPA symbols helped students visualize language and thus look at it from a new perspective. I noticed that when my students were able to 'see' spoken language and understand the processes such spoken language had gone through because of connected speech, they stepped forward to overcoming the fear of helplessness towards a listening comprehension task.

It was as if they had finally realized that it was not entirely their fault that they sometimes could not understand what was being said, but the problem was the intricacies of connected speech which made words so diverse from their citation form. This understanding helped boost their confidence and may help students feel more at ease in asking for clarification when interacting with another speaker. In fact, I realized that students tend to put their failure down to understanding spoken language through their own inability as learners, when it could be the failure of teacher training courses which do not instruct teachers on helping students 'see' and understand the regularities of some pronunciation features of spoken language. By concentrating on what really causes communication breakdown, teachers can reduce learners' workload as well as the level of anxiety which is so often connected with speaking and listening skills.

Teachers' negative attitude towards pronunciation reflects the lack of formal training on the subject. It is about time we considered the teaching of pronunciation as important as grammar and lexis in teacher development courses, for it is a fundamental part of the language, influencing deeply its communicative scope.

As for listening itself, the importance of the post-listening phase should be emphasized. It is when all the problems concerning communication breakdown can be dealt with, and to prevent it from happening again for the same reason. Those post-listening phases were a constant in the group I applied the

pronunciation exercises to, and the students' feedback on what made them not understand or misunderstand specific utterances, made them process the language more carefully, thus getting more familiar with the features of connected speech.

Another important fact to point out about listening from this research is that this skill is not what we tend to think it is. That is to say, we do not need to understand every single word of what people tell us. In reality, what actually is said is sometimes far different from what is being understood. We reconstruct the incomplete oral information using our knowledge of phonetics, syntax, grammar and lexis. What we think we listen to is not always exactly what was uttered. This is even more evident when we have a closer look at formulaic expressions which somehow get lost in the fluidity of speech. Teachers should draw students' attention to the fact that language works in chunks and that the pronunciation of those chunks in spoken language can be simplified so much that it turns into an amorphous mass, extremely different from its written form. Students also need to start to get a grip on tone units, which is the starting of understanding the rhythm of the language.

Moreover, an awareness of those formulaic expressions, together with knowledge of collocations facilitates the learner's task of processing spoken language. The process of listening comprehension for native speakers is boosted by their ability to predict what piece of language comes next. Such an ability is also based on their knowledge of those chunks and collocations. I experienced this when I had to transcribe my students' recorded feedback on the experiment. As they spoke in Portuguese, which is my mother tongue, I should not have had any problem understanding what was being said. However, sometimes I had to listen to it over and over again, and it was my knowledge of the way Portuguese works that helped me make out those unintelligible instances.

As for explicit instruction, the results in this study suggest that in attempting to unveil the complexities of speech production, the explicit teaching of pronunciation features of connected speech do help students cope more efficiently with spoken language. It also helps students notice some differences between their mother tongue and the target language, diminishing the former's interference over the latter. Moreover, explicit instruction draws students' attention to some features of the language that would otherwise be overlooked. By preventing students from understanding those underlying regularities within the language, teachers are undervaluing a powerful tool of language teaching.

Finally, this study has shown that the principle that spoken language is different from written form, should be acknowledged in the classroom. It is a fact



that more recently, there has been an increasing interest in the spoken language in general. The publishing of some books on the subject, has drawn ELT attention to this difference as well as left some doors ajar to further research in this area. Together with this intensified inquisitiveness about spoken language, pronunciation features, if looked into with a more holistic view, may start to get their own deserved place in language teaching.

This study has just tried to shed some light on this issue, and to make teachers more aware of its importance. Further studies on phrasal phonology and prosody will certainly open an array of possibilities for the teaching of additional languages.

## REFERENCES

BRAZIL, David. **A Grammar of Speech**. Oxford: Oxford University Press, 1995.

BROWN, Gilian. **Listening to Spoken English**. Second edition. London: Longmann, 1990.

BRYSON, Bill. **Mother Tongue**. London: Penguin Books, 1991.

CAMBRIDGE ESOL. **First Certificate in English 7 – Examination Papers from University of Cambridge ESOL Examinations**, Cambridge, 2005.

CRYSTAL, David. **The English Language – A Guided Tour of the Language**. 2<sup>nd</sup> edition. London: Penguin Books, 2002.

DOFF, Adrian. **Teaching English – A Training Course for Teachers**. Trainer's Handbook. Cambridge: Cambridge University Press, seventh printing, 1993.

JUDD, Elliot L.; TAN, Lihua; WALBERG, Herbert J. **Teaching Additional Languages**. International Academy of Education. Educational Practices Series – 6. Switzerland: Unesco, 2001.

Available on the internet: <http://www.lbe.unesco.org>

LASS, Roger. **Phonology – An Introduction to Basic Concepts**. Cambridge: Cambridge University Press, 1984.

LEWIS, Michael. **The Lexical Approach**. London: Language Teaching Publications, 1994.

LEWIS, Michael. **Implementing the Lexical Approach**. London: Language Teaching publications, 1997.

McCARTHY, Michael. **Discourse Analysis for Language Teachers**. 6th edition. Cambridge: Cambridge University Press, 1996.

ROST, Michael. **Listening in Language Learning**. London: Longman, sixth printing, 1999.

SHOCKEY, Linda. **Sound Patterns of Spoken English**. Cornwall: Blackwell Publishing, 2003.

SPENCER, Andrew. **Phonology**. Cornwall: Blackwell Publishing, 1996.

UNDERHILL, Adrian. **Sounds Foundation – Living Phonology**. Oxford: Heinemann, 1994.

UNDERWOOD, Mary. **Teaching Listening**. London: Longman, 1993.



## ANÁLISE VARIACIONISTA DE /R/ NA CODA SILÁBICA MEDIAL NO PORTUGUÊS FALADO EM ALAGOAS

*Jeylla Salomé Barbosa Santos Lima<sup>1</sup>*  
*Januacele da Costa<sup>2</sup>*  
*Alan Jardel de Oliveira<sup>3</sup>*

### 1. INTRODUÇÃO

Segmentos do tipo /R/ costumam apresentar múltiplas variações alofônicas não só no português brasileiro, como também em muitas outras línguas. No português brasileiro, são registradas realizações como velar, uvular ou faríngea, utilizadas pela maioria dos falantes, e dental múltipla, pela minoria, conforme Câmara Júnior (1985 [1970]), p. 35). Tais alofonias constituem uma classe de sons, comumente chamada de róticos, com diferentes tipos de articulação (LA-DEFOGED e MADDIESON, 1996). Segmentos dessa classe, embora difíceis de descrever, apresentam características particulares, que buscamos discutir aqui a partir da literatura existente.

Barbosa (2015) faz uma série de observações bastante esclarecedoras sobre a classe dos róticos. Nas línguas que usam o alfabeto latino, esses sons, geralmente grafados com “r” e “R”, são produzidos de formas bastante variadas, o mesmo podendo acontecer nas demais línguas do mundo. Dois desses sons

---

<sup>1</sup> Doutora em Linguística, professora da UNEAL, campus São Miguel dos Campos.

<sup>2</sup> Doutora em Linguística, professora aposentada da UFAL, campus Maceió.

<sup>3</sup> Doutor em Linguística, professor da UFAL, campus Maceió.

são contrastados não apenas nas diferentes nações lusófonas, mas também entre dialetos e falares no interior de uma mesma nação, como ocorre no português brasileiro, e podem ser produzidos como aproximantes (alveolar e retroflexo); como tepes (alveolar e uvular); como aproximantes (alveolar, pós-alveolar e retroflexo); como fricativas de ponto posterior (velar, uvular e glotal); e mesmo como *off* (aproximantes à direita da vogal) rotacizantes, como no caso do “r” caipira. Essas produções podem ser vozeadas e não vozeadas, dependendo de contexto e língua.

Estudos sobre o *status* fonológico da sílaba (TRUBETZKOY, 1939; GOLDSMITH, 1976; HORA E MONARETTO, 2003; COLLISCHONN, 2005; CARVALHO, 2009) têm procurado avaliar segmentos que ocupam as posições que a constituem. Uma das posições que tem despertado a atenção de estudiosos, sob diferentes perspectivas teóricas, é a posição de coda, preenchida, na maioria das vezes, por elementos consonânticos. Segundo Hora e Monaretto (2003), a posição de coda medial apresenta um desenvolvimento de consoantes bem complexo quando comparado à mesma posição em final de palavra.

Considerada essa possibilidade de variação múltipla de segmentos da classe dos róticos, a investigação aqui desenvolvida encontra suporte na sociolinguística variacionista proposta por Weinreich et al. (2006 [1968]) e Labov (2008 [1972]), cujo arcabouço teórico-metodológico assume o princípio da heterogeneidade linguística, ou variação, propondo que essa heterogeneidade é inerente a todo e qualquer sistema linguístico, variação linguística não sendo um fato aleatório, mas governado por regras e restrições tanto linguísticas quanto extralinguísticas.

Em nossos dados, foram encontradas quatro realizações de /R/<sup>4</sup> em coda silábica: fricativa glotal ['kahtɐ] ‘carta’; apagamento [mine'vĩnʊ] ‘Minervino’; aproximante ['kɔ:ɾɐ]<sup>5</sup> ‘corda’; tepe ['turmɐ] ‘turma’. Dessas, a variante com maior frequência de ocorrência, de modo geral, é a fricativa glotal [h], enquanto que o tepe [r] apresenta frequência muito baixa. A variante apagamento [ø], bem como

---

<sup>4</sup> Seguiremos a proposta de Silva (1999, p. 161), que transcreve de “R forte”, ou seja, o fonema ao qual as variantes tratadas neste trabalho relacionam-se, como /R/, sempre que estivermos nos referindo ao Português Brasileiro.

<sup>5</sup> Entre as variantes, a transcrição fonética da aproximante apresentou problemas, uma vez que no IPA (International Phonetic Alphabet), há símbolos para duas aproximantes róticas: [ɹ] e [ɻ], sendo o primeiro a transcrição de uma aproximante alveolar/pós-alveolar e o segundo, a transcrição de uma aproximante retroflexa. Considerando que a variante rótica aproximante encontrada nos dados do Português Brasileiro falado em Alagoas é uma aproximante alveolar, não retroflexa, transcreveremos esse som como [ɹ].

a variante aproximante [ɹ], são significativas. Essa configuração da variação do /R/ na variedade de Português Brasileiro falada em Alagoas direcionou a nossa investigação no sentido de realizarmos duas análises distintas/separadas.

A variante *tepe* tendo apresentado, conforme informamos, um número bastante reduzido de ocorrências – apenas 26 casos – optamos por não a analisar estatisticamente.

No decurso da investigação, detectamos a existência de um processo variável secundário disparado pela realização do /R/ como aproximante [ɹ]: as oclusivas coronais /t, d/ seguintes realizam-se como palatalizadas [tʲ, dʲ]. Incluímos, então, a análise desse processo no nosso trabalho, visto que tal processo pode contribuir para argumentarmos em favor da nossa hipótese de que a variante de /R/ é uma aproximante e não um rótico consonantal.

Este trabalho, portanto, tem como objetivo geral investigar a variação de /R/ em posição de coda silábica medial, no Estado de Alagoas, apresentando uma descrição do conjunto de variáveis linguísticas e sociais que podem estar influenciando a variação, de modo a verificar se e quais das realizações atestadas apresentam indícios de mudança em curso ou se se trata de variação estável, de acordo com os pressupostos e finalidades da Sociolinguística Variacionista, nosso principal suporte teórico-metodológico.

## 2. A COMPLEXIDADE DOS RÓTICOS

A gama de realizações de fonemas róticos, dentro de seus contextos de ocorrência, apresenta uma diversidade de sons que muitos estudiosos da fonética e da fonologia têm tentado esclarecer e até mesmo compreender em áreas dialetais distintas. Segundo Ladefoged e Maddieson (1996), os sons que podem representar o fonema estudado são classificados como róticos e essas nomenclaturas são baseadas no fato de que todos esses sons tendem a ser escritos com a letra ‘r’<sup>6</sup>, mesmo porque não existe uma propriedade física que constitua a essência de todos os róticos. Ainda para esses autores, não são nem o modo de articulação, nem o ponto de articulação que definem a classe dos róticos, o que torna difícil organizar a noção de classe para esses sons.

Róticos têm despertado o interesse de muitos foneticistas e fonólogos (DALSTON, 1975; LINDAU, 1985; LADEFOGED; MADDIESON, 1996; HERNANDORENA; LAMPRECHT, 1997; RECASENS; PALLARÈS, 1999;

---

<sup>6</sup> O alemão e outras línguas saxônicas, como o holandês, usam ‘g’ para um tipo de rótico. Outras línguas usam ‘h’ para a fricativa glotal.

VIDOR, 2001; SILVA, 1996, 2002; SOLÉ, 2002; WHITLEY, 2003; MEZZOMO; RIBAS, 2004). Esse interesse se deve ao fato de essa ser uma classe de sons em que ocorre muita variação, como já vimos, e essa variação ampla pode ser encontrada tanto sincrônica quanto diacronicamente.

Ladefoged e Maddieson (1996) buscam argumentos de diferentes naturezas para definir propriedades comuns aos membros dessa classe. Os autores questionam, por exemplo, se a classificação dos róticos é baseada nas relações sincrônica e diacrônica existentes entre os membros da classe ou se há, realmente, uma similaridade fonética entre eles que até o momento não tenha sido percebida. Acreditam que tal similaridade deva ser mais acústica que articulatória, inclusive, como também aponta Lindau (1985).

Sobre a inexistência de similaridades fonéticas entre os sons que ocorrem como variantes de um fonema da classe dos róticos, Ladefoged e Maddieson (1996) dizem que, apesar de foneticamente os róticos não apresentarem características comuns, fonologicamente eles apresentam comportamentos similares. Os autores destacam dois argumentos de natureza fonológica: (a) são, praticamente, a única consoante que surge como segundo elemento de grupo consonântico; (b) tendem a ocorrer próximos ao núcleo silábico, observando-se uma afinidade com as vogais, sendo que os róticos podem chegar a fundir-se de várias formas com vogais contíguas ou a terem variantes silábicas. Esse fenômeno pode ser observado entre dialetos de uma mesma língua, observando-se a posição dos róticos em relação às vogais – pós-vocálica ou intervocálica –, o que pode resultar em diferentes realizações. Algumas línguas como Alemão, Sueco, Francês, Farsi e Palauan também apresentam realizações diferentes do fonema, tais como aproximantes, *tepes* e fricativas.

Para esses autores, esses sons não podem ser considerados, do ponto de vista fonético, uma classe homogênea, embora apresentem algumas características bastante frequentes entre si na maioria das línguas: o ponto de articulação mais comum, por exemplo, é o alvéolo-dental, apesar de existirem ocorrências no ponto pós-alveolar (que podem ser retroflexos) e, em algumas línguas, no ponto de articulação uvular.

Trubetzkoy (1969) já observava a variabilidade apresentada entre os róticos. Em Alemão, por exemplo, quando o “r”<sup>7</sup> era antecedido de vogais podia realizar-se como um *tepe* dental, um *tepe* uvular ou, ainda, uma fricativa velar, sem que se estabelecesse oposição fonêmica entre essas variantes. Nas demais posições,

---

<sup>7</sup> O “r” do alemão a que se refere Trubetzkoy é representado em estudos de fonologia alemã como /R/ e ocorre em palavras como [ˈbo:Rə] “broca”.

podia ser pronunciado como uma velar “incompletamente articulada” ou como uma vogal não silábica.

No Português do Brasil, do ponto de vista fonológico, há dois fonemas da classe dos róticos. Silva (1999, p. 159-160) observa que existem o “r fraco” e o “R forte”, uma vez que existe um contraste fonêmico entre esses dois tipos do “r” em posição intervocálica, como mostra a existência de pares mínimos em “caro/carro”; careta/carreta”. Este contraste fonêmico pode manifestar-se pelo número de vibrações da língua na articulação do segmento consonantal: tepe simples em ‘caro’ [ˈkaɾo] e vibrante múltipla em ‘carro’ [ˈkaʀo]. Monaretto et al. (2000) afirmam que os sons do r-forte (vibrante múltipla) no PB podem corresponder tanto a um tepe propriamente dito, quanto a uma fricativa ou a uma aspirada.<sup>8</sup>

O “R forte” varia consideravelmente no português brasileiro. Silva (1999), o “r fraco” sempre ocorre no início da sílaba. O tepe é representado por /r/ e se relaciona ao “r fraco”. A autora afirma que a perda do contraste fonêmico entre “R forte” e “r fraco” é neutralizada no português em posição final de sílaba. Isso quer dizer que o contraste que existe em posição intervocálica não ocorre em posição de coda silábica. Esse contraste também não existe em posição de início de palavra, pois nessa posição apenas é encontrado o “R forte”, ou vibrante múltipla, nos termos de Silva (1999). Entretanto, na variedade alagoana, em início de palavra, ocorre uma fricativa glotal surda [h], como mostram os exemplos [ˈhɔda] ‘roda’ e [ˈhadʒo] ‘rádio’.

Para representar esse “R pós-vocálico”, utilizamos o símbolo /R/, conforme notado anteriormente, que representa o arquifonema, isto é, o resultado de uma neutralização do contraste entre os dois “r” em posição final de sílaba – seja em meio de palavra, como em ‘car.ta’, ou em final de palavra, como em ‘mar’.

Como vimos informando, a nossa investigação é sobre a variação desse arquifonema /R/ em coda silábica medial. Foram feitas duas análises, uma entre [h] – que é a realização predominante na variedade sob estudo – e [ø] e outra entre [h] e a variante que estamos considerando uma aproximante alveolar [ɹ]<sup>9</sup>. Aqui, apresentaremos a segunda análise. As duas variantes de /R/ apresentam peculiaridades notáveis na variedade em causa: a aproximante ocorre apenas em

---

<sup>8</sup> Na variedade de Português que estamos investigando, o fonema /R/, conforme representado por Silva (1999), realiza-se, predominantemente, como [h], fricativa glotal surda. Assim, o contraste nessa variedade se faz entre uma tepe simples /r/ e uma fricativa glotal surda /h/, contraste que podemos exemplificar com /ˈkaɾo/ «caro» versus /ˈkaʀo/ «carro».

<sup>9</sup> O som encontrado nos dados desta pesquisa assemelha-se mais ao [ɹ] do inglês do que à retroflexa encontrada em alguns dialetos da região sudeste do Brasil.



coda medial, enquanto que o apagamento em final de palavra é quase categórico, corroborando com estudos de Callou, Leite e Moraes (1996, 2002).

Segundo Silva (1999), os segmentos [x, ɣ, h, fi] relacionam-se a ‘r’ em posição final de sílaba e esses segmentos concordam em vozeamento com a consoante seguinte. De acordo com a autora, o que ocorre é uma variação posicional pelo fato de as fricativas assimilarem o traço [±voz] da consoante seguinte, pelo menos no dialeto de Belo Horizonte, no qual ocorre uma fricativa vozeada, antes de uma consoante vozeada, como em [‘kaŋga] ‘carga’ uma fricativa desvozeada antes de consoante desvozeada, como em [‘tohto] ‘torto’ e em final de palavra, como em [‘mah] ‘mar’.

Parece ser correto afirmar que a realização mais frequente do “R forte” é, em um grande número de dialetos do Português brasileiro, ou a fricativa velar ou a fricativa glotal. Essas realizações são bem diferenciadas do tepe, ou tepe simples, som considerado uma das principais realizações da variante fraca.

Malmberg (1954) nota que os tepes são articulados de tal modo que o órgão articulante – a ponta da língua ou a úvula – provoca uma série de oclusões muito breves, separadas por pequenos elementos vocálicos. O autor constata a ocorrência de mudanças linguísticas na pronúncia de várias línguas da Europa, sobretudo, a ocorrência dos processos fonológicos de enfraquecimento, a substituição de “R” anterior por um “R” posterior, processo este também observado e analisado por muitos estudiosos brasileiros.

Neste trabalho, uma de nossas propostas é a análise de uma variante com características de aproximante, considerada assim por não ter a mesma articulação de um som retroflexo, que, segundo Cagliari (1981), pode ser pronunciado de várias maneiras, sendo a maneira mais comum levantando-se e encurvando-se a ponta da língua em direção à região palato-alveolar ou mesmo palatal. Ainda segundo o autor, dependendo do tipo de constricção produzida pela língua, o som retroflexo pode ser oclusivo [t,d], nasal [ŋ], lateral [l], fricativo ou construtivo [ɹ]. Pode mesmo ser o som de um tepe, pois, de acordo com o autor, esse tipo de som é produzido com a ponta da língua (raramente o véu palatino) quando se põe a bater repetidamente contra a área alveolar da abóbada palatina atrás dos dentes incisivos (ou contra a parte posterior da língua), como alveolar sonora ‘r’, alveolar surda ‘r’ e uvular sonora ‘r’. Essa última raramente acontece no português.

Observando a articulação da variante aproximante nesta pesquisa, concluímos que a projeção da ponta da língua (articulador ativo) se aproxima do palato duro (articulador passivo) na sua parte alveolar. Do ponto de vista do contexto que segue, pode-se observar que o mais favorável à realização dessa variante é

aquele formado por uma consoante oclusiva coronal – /t/ ou /d/. De um modo bastante interessante, ocorre um processo reverso: as oclusivas /t/ e /d/ palatalizam-se depois da aproximante, realizando-se como [tʲ] e [dʲ], respectivamente. Assim, esse processo assemelha-se ao processo de palatalização ou africacão das consoantes /t/ e /d/ quando seguidas da aproximante /j/, fato observado em palavras como ['ojto] ‘oito’ e ['dojdo] ‘doido’, que podem ser pronunciadas como ['ojtʲo] e ['dojdʲo] ou como ['ojtʃo] e ['dojdʒo], em algumas variedades alagoanas (SANTOS, 1996; OLIVEIRA, 2017).

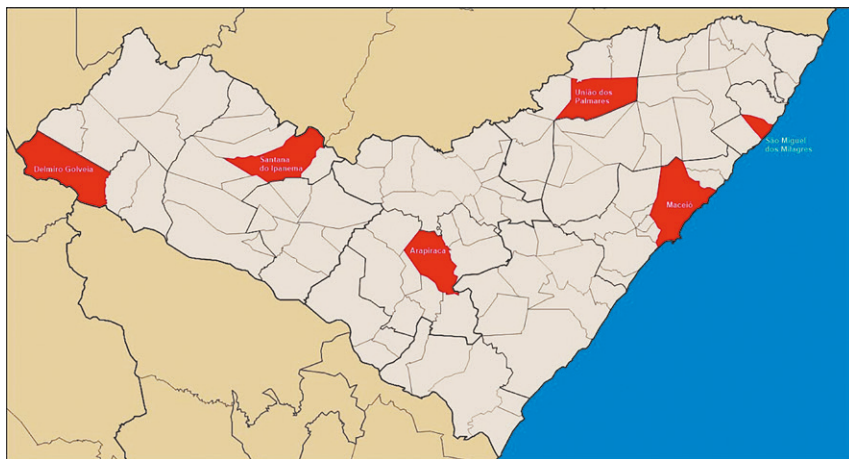
### 3. PROCEDIMENTOS METODOLÓGICOS

Para a sociolinguística, dados os seus pressupostos, é necessário que os dados básicos para qualquer forma de análise linguística geral seja a língua tal como usada por seus falantes nativos comunicando-se uns com os outros na vida diária, o que quer dizer o vernáculo, conforme definimos anteriormente. É esse vernáculo, a língua falada espontaneamente pelos falantes de um grupo definido como uma comunidade real de fala, ali observada diretamente, que é o objeto de estudo da variação linguística, dado que a língua é concebida, nessa perspectiva teórica, como um fato social (LABOV, 2008 [1972]). Sendo esse o objeto de estudo da sociolinguística, constitui-se o ponto de partida do estudo do processo de variação e mudança de uma língua.

Nesta pesquisa, trabalhamos com dados de fala de seis cidades alagoanas, cada uma delas considerada, *a priori*, como sendo uma comunidade de fala, basicamente porque elas apresentam realidades socioculturais e econômicas diferentes e estão localizadas, do ponto de vista geográfico, relativamente distantes.

A Figura 1 mostra, no mapa de Alagoas, a localização das cidades – São Miguel dos Milagres, Maceió, União dos Palmares, Arapiraca, Santana do Ipanema e Delmiro Gouveia – nas quais os dados para esta pesquisa foram coletados.

Figura 1 – Mapa 1: Cidades pesquisadas.



Fonte: Autora, 2019 – Adaptado de Variação Linguística no Português Alagoano – PORTAL, 2013.

### 3.1 Constituição da amostra

Alguns critérios básicos de seleção para que o indivíduo pudesse participar na composição da amostra deste trabalho foram os seguintes: i) ter nascido no município; ii) não ter se ausentado do município por mais de 1 ano; e, preferencialmente, iii) ter ambos os pais nascidos também no município.

Foram, assim, selecionados 144 participantes, sendo 24 por cidade pesquisada. Para compor a amostra, utilizamos o método “amigo do amigo” (MILROY, 2004), de forma que os participantes foram selecionados sem que a entrevistadora os conhecesse. A amostra de língua registrada por esse método fica sendo uma amostra que é constituída através de um mecanismo não probabilístico, pois, após um contato inicial com alguns sujeitos da comunidade a ser estudada, os indivíduos seguintes que participarão da pesquisa são indicados pelos primeiros. Estes indicam novos prováveis participantes e, assim, sucessivamente, até que a amostra esteja completa.

A amostra por cidade foi composta por cotas, considerando-se as variáveis sociais *sexo/gênero, faixa etária e escolaridade*, estratificada conforme a tabela a seguir:

Tabela 1 – Composição da amostra por cidade

		faixa etária		
sexo/gênero	escolaridade	18-30 anos	45-55 anos	>65anos
Feminino	<9 anos	2	2	2
	>11 anos	2	2	2
Masculino	<9 anos	2	2	2
	>11 anos	2	2	2
<b>Total</b>		<b>24 participantes por cidade</b>		

### 3.2 Coleta e transcrição dos dados

As pesquisas sociolinguísticas vêm se dedicando ao estudo da língua em uso no seio das comunidades de fala, com atenção devida aos aspectos linguísticos e sociais. Estudiosos têm se dedicado à variabilidade presente no uso da língua e observado os efeitos positivos e negativos sobre a emergência desses usos, prevendo seu comportamento regular e sistemático. Com esse entendimento, decidimos, neste estudo, trabalhar com dados oriundos de fala natural, a fim de captar e analisar melhor as variantes de /R/ em posição de coda silábica medial existentes nas cidades pesquisadas, sob a luz da perspectiva teórico-metodológica da sociolinguística variacionista (LABOV, 2008 [1972]), conforme já informamos.

A coleta de dados foi realizada entre julho de 2014 e outubro de 2015. O método “amigo do amigo” foi de suma importância para esta coleta. A partir de tal intermediação, muitos informantes se mostraram bastantes solícitos durante a entrevista, assim como quando solicitados a indicar novos colaboradores. Em contato com os participantes, evitamos a palavra “entrevista”. Foi preferida a palavra “conversa” porque estamos fazendo uma pesquisa que lida com dados espontâneos e era preciso que os participantes se sentissem à vontade para conversar, de modo que não houvesse o policiamento da fala, controlando-a e, assim, enviesando a amostra.

Os participantes eram informados de que a conversa seria gravada e que somente os pesquisadores teriam acesso ao conteúdo registrado e ao questionário social, preenchido logo após a entrevista, juntamente com o Termo de Consentimento Livre e Esclarecido<sup>10</sup> – TCLE.

O roteiro de gravação com cada participante foi organizado da seguinte forma: em primeiro lugar, pedíamos uma narrativa e, depois, uma descrição. A

<sup>10</sup> O TCLE desta pesquisa foi aprovado pelo comitê de ética em pesquisa com o número de CAAE: 17274214.7.0000.5013.

nossa pressuposição era que estaríamos levando o informante a construir um percurso histórico e que seria mais fácil e confortável conversar, relembando de casos e histórias, que poderiam recriar emoções fortes do passado vivido, e, partir disso, descrever elementos presentes na sua memória e reavivados pela narração de eventos. No final, apresentávamos um tema para a argumentação, normalmente um tema atual que sabíamos estar presente nas conversas do povo brasileiro naquele momento.

Com a finalidade de garantir-se um padrão de dados coletados, o roteiro básico da entrevista com cada participante foi estruturado da seguinte forma:

1. *NARRAÇÃO* – “*Conte uma lembrança importante*”.
  - a) da sua infância em casa.
  - b) da sua infância na escola.
  - c) da sua infância com amigos.
  - d) de como era a cidade na sua infância
  - e) da sua infância com os pais.
  - f) da sua infância com os avós.
  - g) da infância dos seus filhos em casa (caso tenha filhos).
  - h) da infância dos seus filhos na escola (caso tenha filhos).
  - i) da sua juventude ou adolescência.
  - j) de relacionamentos amorosos (marido/esposa, namorado(a) etc., como conheceu o marido/esposa, namorado(a) etc.)
2. *DESCRIÇÃO* – “*Conte com o máximo de detalhes possível...*”
  - a) como era a casa em que você morava quando tinha 10 anos.
  - b) como era a cidade na sua infância.
3. *ARGUMENTAÇÃO* – “*O que você pensa sobre...*”
  - a) pena de morte.
  - b) aborto.
  - c) casamento entre pessoas do mesmo sexo.

Para a gravação, foi utilizado um gravador da marca TASCAM, modelo DR 100. As gravações foram feitas em formato .wav, com taxa de amostragem de 24 bits e resolução de 48 kHz. Utilizamos também um microfone headset condensador cardioide unidirecional da marca Arcano, modelo WZ – 1.000. As gravações

tiveram duração entre 9 e 11 minutos, para cada indivíduo, totalizando 1.584 horas de registros de fala.

### 3.3 Tratamento dos dados

As transcrições dos dados foram feitas utilizando-se o software Praat (Boersma & Weenink, 2007), que é uma ferramenta utilizada para a análise de voz, conforme já mencionado. Foi feita a segmentação do áudio, de modo que a transcrição (orto)gráfica corresponde exatamente ao trecho de fala que pode ser ouvido em determinado ponto selecionado.

As ocorrências do fenômeno a ser analisado nesta pesquisa foram identificadas e organizadas em um banco de dados específico que contém, ainda, as variáveis de interesse do estudo. Quando julgamos necessário, algumas ocorrências foram analisadas acusticamente no *software Praat*.

Os dados foram codificados conforme o modelo apresentado e descrito na Tabela 2.

Tabela 2 – Codificação dos dados

Código do informante: <b>SI18M08</b>
SI - para Santana do Ipanema
18 – idade
M – sexo/gênero masculino
08 – nível de escolaridade (ensino fundamental)

Fonte: Elaboração própria, 2019.

### 3.4 A variável dependente e suas variantes

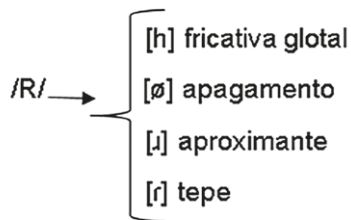
A variável linguística deste estudo é o arquifonema /R/ do Português Brasileiro quando em posição de coda silábica medial, como em /kɔR.da] ‘corda’, de acordo com Silva (1999), que afirma ser ele comumente nomeado “R pós-vocálico”.

Em nosso trabalho, tomando como base investigações que já efetuamos em Alagoas sobre variação de “r” em coda silábica (SANTOS, 2010), bem como outros trabalhos que consideraram dados da região, como, por exemplo, Callou, Serra e Cunha (2015), definimos como sendo o conjunto de variantes que compõem a variável /R/ – “R pós-vocálico” em coda medial – as seguintes realizações: fricativa glotal [h], apagamento [ø], aproximante [ɹ], tepe [r].

Consideramos, assim, que o arquifonema /R/, nossa variável dependente do ponto de vista da estrutura fonológica da língua portuguesa, apresenta o seguinte

envelope de variação na variedade falada em Alagoas, nas cidades que compõem o universo da pesquisa:

Figura 2 – Variantes encontradas em Alagoas.



Variantes em alternância representam “maneiras diferentes de dizer a mesma coisa” (LABOV, 1972, p. 271). Essa alternância, porém, pode ser vista como parcialmente aleatória, pois em determinado momento o falante pode usar uma ou outra alternativa, sem necessariamente querer indicar qualquer diferença de significado com a escolha, conforme afirmam Guy e Zilles (2007).

A pesquisa aqui desenhada, bem como a análise executada sobre os dados obtidos, tomou por base essa variável dependente. Entretanto, no programa utilizado para a análise estatística, não rodamos todas as variantes ao mesmo tempo. Em vez disso, efetuamos duas rodadas, uma envolvendo a variante fricativa glotal versus apagamento e outra procurando observar a variação entre a fricativa glotal e a aproximante. A variante tepe, vale lembrar, por apresentar frequência de ocorrência pouco significativa, não foi analisada estatisticamente.

### 3.5 Variáveis independentes

Para esta pesquisa, foram controladas cinco variáveis linguísticas – contexto seguinte, contexto antecedente, acentuação da sílaba, tamanho da palavra e item lexical – e cinco variáveis sociais – sexo/gênero, faixa etária, escolaridade, cidade e indivíduo.

O conjunto de variáveis linguísticas, com seus grupos de fatores correspondentes, ficou constituído da seguinte forma:

Variável	Grupo de Fatores
Contexto seguinte	[coronal] e [±contínuo]
Contexto antecedente	[±posterior]

Acentuação da sílaba                      sílaba pré-acentuada, acentuada, não acentuada  
 Tamanho da palavra                      duas sílabas, três sílabas, quatro ou mais sílabas

### 3.6 Análise estatística dos dados

Para a análise estatística dos dados, utilizamos o software R.<sup>11</sup> Os dados foram submetidos à análise estatística, de acordo com as propostas apresentadas em Oliveira (2009, 2011b, 2012), relacionadas à utilização de modelos de regressão logística binomial, modelos de regressão logística multinível, e de interação entre variáveis independentes e métodos de seleção de variáveis.

Toda a análise estatística foi feita com o auxílio do software R, utilizando os pacotes ‘gmodels’ (para gerar Tabelas de contingência), ‘lme4’ (para regressão logística multinível, TRMV e TW) e ‘visreg’ (para gerar os gráficos).

## 4. RESULTADOS E DISCUSSÃO

A Tabela 3, a seguir, foi elaborada no início da análise com o objetivo de observarmos a distribuição das variantes em relação ao contexto seguinte.

Tabela 3 – Distribuição das variantes de /R/ em relação ao contexto seguinte

	Contexto seguinte	Total	% [h]	% Ø	% [ɹ]	% [r]
oclusivas	[p]	16	87,5	12,5	0	0
	[b]	40	67,5	22,5	0	10,0
	[k]	242	89,7	6,6	1,7	2,1
	[g]	96	88,5	3,1	6,2	2,1
	[t]	1249	83,8	0	16,2	0
	[d]	450	80,9	0,7	18,4	0
fricativas	[f]	19	89,5	9,1	0	0
	[v]	96	26,0	69,8	0	4,2
	[s]	429	23,8	74,8	0	1,4
	[z]	28	10,7	89,3	0	0
	[ʃ]	4	0,0	100,0	0	0
	[ʒ]	129	22,5	74,4	0	3,1
soantes	[l]	8	100,0	0	0	0
	[m]	815	99,0	0,0	0	0,1
	[n]	342	100,0	0	0	0
TOTAL		3963	77,9	14	7,4	0,7

<sup>11</sup> www.r-project.org.



Observando a variante [ø], vemos que ela tem ocorrência mais alta quando o contexto seguinte é uma fricativa. Vejamos os exemplos a seguir:

(1)

“Ele vende leite em casa vende [pɛ'fumi] e coisa da tupperware” (perfume)  
– AR42M07

(2)

“Eu estou até na [hɛ'zɛvø] da polícia de dois mil e seis” (reserva) – AR40M15

(3)

“Com a graça de deus a gente vai encontrando ['fosø]” (força) – DE74F11

(4)

“E terminei o primário em Arapiraca né? No Adriano ['ʒɔʒi] na escola (Jorge) – AR79F15

Hora, Pedrosa e Cardoso (2010) afirmam que o apagamento é condicionado pelo contexto fonológico seguinte – quando preenchido por uma fricativa, como nos exemplos (1) a (4) apresentados. Entretanto, nos dados encontramos muitos casos de apagamento também diante das oclusivas [p] e [b]. Vejamos os exemplos a seguir:

(5)

“Mas eu não corria muito terminava apanhando ou os meninos me pegava na [su'prezø]” (surpresa) – UP45M15

(6)

“Muito você se [suprɛ'ɛdi] não é? Você que vive lá em Maceió” (surpreende) – DE47F15

(7)

“Crime ['babaɾɔ] aí absurdo” (bárbaro) – SI55M03

(8)

“Soldando uns bagulhos para lá negócio de [kabura'do] de carro essas coisas” (carburador) – SM19M07

Callou, Serra e Cunha (2015) também concluem que, na capital Maceió, são as fricativas as que mais frequentemente propiciam o apagamento do 'r': ([s], 76%, e [v], 46%). Callou et al. (2013), ao analisar o fenômeno no Sul e Sudeste do país, afirmam que o apagamento está relacionado às mudanças na articulação da con-

soante, de tepe para fricativa e de anterior para posterior. Os sucessivos estágios de mudança na realização do rótico, tanto do modo quanto do ponto de articulação – tepe anterior para fricativa posterior, aspiração e apagamento – estariam relacionados ao processo de enfraquecimento e conseqüente apagamento da consoante.

Conforme mostrado na Tabela 3, também concluímos que o contexto seguinte propicia o apagamento de /R/ em coda medial, principalmente se este contexto for de fricativas, como [v], 69,8% [s], 74,8% [z] 89,3% [ʃ], 100% [ʒ], 74,4%. Nossos dados aproximam nossa pesquisa das mencionadas no parágrafo anterior, no sentido de que o processo de apagamento está avançando na região Nordeste, conseqüentemente, nos falares alagoanos.

Ainda sobre a Tabela 3, o apagamento do /R/ não ocorre diante de consoantes soantes (/l/, /m/, /n/). Não houve nenhum apagamento nos 1.165 dados de /R/ seguidos de consoantes soantes. Concluímos, a partir da Tabela 3, que tais consoantes bloqueiam o apagamento; por isso, os dados em que o ambiente seguinte ao /R/ é constituído por uma dessas consoantes serão excluídos da análise que envolve a variante apagamento.

Também podemos observar, na Tabela 3, que a aproximante ocorre somente diante das oclusivas /t, d, k, g/. Não houve nenhuma realização como aproximante em 1.926 ocorrências de /R/ nos demais ambientes: 705 casos em que a consoante seguinte é uma fricativa; 1.165 casos em que a consoante seguinte é uma soante; e nos 56 casos em que essa consoante seguinte é uma oclusiva labial. Assim, a análise da variação entre a realização do ‘r’ e a aproximante será restrita aos contextos nos quais ela ocorre, quais sejam, quando seguidas das oclusivas /t, d, k, g/.

A Tabela 3 nos mostra que a aproximante, além de ocorrer apenas antes de consoantes oclusivas, é favorecida quando seguida de oclusivas alveolares /t/ (16,2%) e /d/ (18,4%). O percentual diante das oclusivas velares é significativamente menor: 1,7% para /k/ e 6,2% para /g/.

Vejamos alguns exemplos:

(9)

“E ali era o porto das [ba.r'kasɐ]” (barcaça) – [SM97F01]

(10)

“Aí depois eu desistir de ['ka.rɣɐ] de horário” (carga) – [AR76M08]

(11)

“Não aí no [sɛ.r'tãw] de Pernambuco” (sertão) – [AR76M08]

(12)

“Eu passeava muito tinha muitos amigos não [pɛ.r'djɐ] nada” (perdia) – [MC70F15]

A pequena quantidade de dados da variante [r], 26 casos, o que corresponde a 0,66% do total de dados, dificulta a utilização de métodos estatísticos de regressão e testes de hipóteses. Tendo isso em vista, nossa opção foi analisar as ocorrências de tepe e buscar generalizações a partir delas. A Tabela 4 apresenta as palavras que contêm essa variante.

Tabela 4 – Produção de tepe

Palavras	Ocorrências	Participantes	Cidade	Faixa etária
‘barbosa’	4	AR89M11	Arapiraca	Idoso
‘perco’	1	SI68M16	Santana do Ipanema	Idoso
‘mercado’	1	AR76M08	Arapiraca	Idoso
‘porco’	1	MC77F00	Maceió	Idoso
‘supermercado’	1	SI65M04	Santana do Ipanema	Idoso
‘circo’	1	MC70F15	Maceió	Idoso
‘perguntar’	1	MC65M07	Maceió	Idoso
‘vergonha’	1	SI65M04	Santana do Ipanema	Idoso
‘preservou’	1	SM18F09	S. Miguel dos Milagres	Idoso
‘servia’	1	SI65M04	Santana do Ipanema	Idoso
‘sirvo’	1	SI65M04	Santana do Ipanema	Idoso
‘árvore’	1	SI74F01	Santana do Ipanema	Idoso
‘comércio’	2	SI89F11	Santana do Ipanema	Idoso
‘conversando’	2	SI65M04	Santana do Ipanema	Idoso
‘terceiro’	1	SI89F11	Santana do Ipanema	Idoso
‘diversas’	1	SI89F11	Santana do Ipanema	Idoso
energia’	3	SI65M04	Santana do Ipanema	Idoso
‘energia’	1	MC70F15	Maceió	Idoso
‘turma’	1	SI46F04	Santana do Ipanema	Idoso

Analisando as ocorrências de tepe, verificamos que o processo ocorre apenas na fala de idosos. Quanto à cidade onde se realiza a vibrante, o quadro é o seguinte: em Santana do Ipanema foram identificadas mais ocorrências (16 casos); em Arapiraca ocorreram 5 casos; em Maceió, 4 e em São Miguel dos Milagres, 1 caso. Não parece haver uma generalização do ponto de vista diatópico no uso da variante em Alagoas. Maceió e São Miguel dos Milagres localizam-se no leste

alagoano, Arapiraca, no centro, e Santana do Ipanema no Oeste. Em relação a outras variáveis sociais, há participantes de diferentes níveis de escolaridade e de ambos os sexos/gêneros. O contexto seguinte é diversificado, tanto em termos de modo de articulação, quanto em termos de ponto de articulação, vozeamento e nasalidade do segmento seguinte (há fricativas, oclusivas e nasais; bilabiais, alveolares, alveopalatais e velares). O processo ocorre em sílabas acentuadas (sirvo, comércio, circo, porco etc.) e em sílabas não acentuadas (energia, servia, vergonha, marcado etc.).

O fato de o tepe ocorrer muito raramente e estar presente somente na fala dos idosos nos permite concluir que tal variante está em extinção nos falares alagoanos. Em relação às variáveis sociais sexo/gênero e escolaridade, não parece haver favorecimento de uma ou outra categoria no interior de tais variáveis. O mesmo pode ser dito em relação a variáveis linguísticas como o contexto seguinte e a acentuação da sílaba.

Diante da pequena quantidade de ocorrências da variante tepe, nos limitamos, nesta pesquisa, às análises já apresentadas.

Na análise estatística dos dados, a seguir, a seleção das variáveis estatisticamente significativas e a hierarquização de tais variáveis foi feita utilizando-se o teste da razão da máxima verossimilhança. Utilizamos o teste de Wald para testarmos se haveria diferença estatisticamente significativa entre o efeito dos fatores e a média dos efeitos dos fatores.

Foram, desse modo, efetuadas duas análises distintas/separadas: a análise da variação entre a fricativa glotal [h] e o apagamento [ø]; a análise da variação entre a fricativa glotal [h] e a aproximante [ɹ] e de um fenômeno relacionado a essa última realização: a palatalização da oclusiva coronal seguinte, desdobrando-se, assim, em duas subseções – 4.1.1 e 4.1.2, respectivamente. Porém, devido ao espaço aqui, apresentaremos apenas a análise da variação entre a fricativa glotal [h] e a aproximante [ɹ] e fenômeno relacionado.

#### **4.1 Análise da variação fricativa glotal [h] versus aproximante [ɹ] e fenômeno relacionado**

Como vimos anteriormente, a ocorrência da aproximante é restrita à presença das oclusivas /t, d, k, g/ no contexto seguinte. Não houve nenhuma ocorrência da variante em ambientes linguísticos diferentes desses. Diante disso, foram excluídos desta análise todos os dados nos quais as consoantes seguintes não pertencem ao grupo de consoantes determinado.

Vejam os alguns exemplos das variantes observadas:

(13)

“Ciúme por [ˈpa.ti] dela que era muito ciumenta” (parte) – [MC65M07]

(14)

“Assim eu fazia me [akɔ.ɾ.davɐ] de madrugada dava café o doente” (acordava) – [UP88F04]

(15)

“A gente veio [ẽba.ɾ.kado]” (embarcado) – [SM97F01]

(16)

“Todo dia seis [ˈka.ɾgɐ] de mandioca” (carga) – [AR89M11]

Nas análises acústicas, identificamos a existência de um processo variável secundário, que não envolve somente a variação do /R/, mas que é consequência dessa variação, em especial, da *aproximante*. Trata-se da palatalização das oclusivas /t/ e /d/ quando a variante realizada antes delas é a aproximante [ɹ].

Vejam os alguns exemplos retirados dos dados:

(17)

“[ˈpɛ.tiɔ] dos fumicultores do clube fumicultores” (perto) – [AR89M11]

(18)

“Cem quilos tenho pego por aí com o [hɔˈbɛ.tiɔ]” (Roberto) – [SM66M02]

(19)

“Daqui para loja toda [ˈta.ɾɔ] trocava de roupa” (tarde) – [UP80F15]

Aparentemente, a palatalização de /t/ e /d/ aqui é similar à palatalização progressiva de tais segmentos precedidos da aproximante [j], como em “oito”, produzido como [ˈojtʃu] e “doido”, produzido como [ˈdojdʒu], investigada por Santos (1996) e por Oliveira (2017) em dados de fala de Maceió. Nesses trabalhos, a realização das oclusivas /t/ e /d/ como [tʃ] e [dʒ], respectivamente, é tratada como um processo de assimilação, pelo qual o traço [-anterior] da aproximante [j] espalha para o nó de ponto da consoante seguinte, fazendo-a realizar-se como uma africada, uma consoante complexa, portanto.

No caso da palatalização provocada pelo espalhamento do traço [-anterior] da aproximante [ɹ], que ocorre nos nossos dados, a consoante seguinte apenas se

palataliza, ou seja, o segmento resultante não é uma consoante complexa, mas apenas uma consoante palatalizada: [tʲ] e [dʲ].

Se considerarmos que a aproximante [j] – realização fonética de uma vogal [i] em posição de coda, conforme Camara Júnior (1985 [1970]) – é um segmento palatal como o são as fricativas [ʃ] e [ʒ], o fato de essa realização de /R/ nos nossos dados apenas palatalizar a consoante seguinte e não a criar um segmento complexo, parece-nos ser uma evidência forte que nos leva a descrever o segmento em questão como sendo uma aproximante alveolar.

Na sequência, estamos, portanto, apresentando duas análises: i) a variação entre [h] e [ɹ], incluindo somente dados de oclusivas /t, d, k, g/; e ii) a variação entre /t,d/ e a palatalização de tais segmentos, incluindo somente os casos em que o /R/ ocorre como aproximante.

#### 4.1.1 Análise da variação entre a fricativa glotal [h] e a aproximante [ɹ]

Para a variação entre [h] e [ɹ], as variáveis independentes sociais analisadas foram *sexo/gênero*, *escolaridade*, *cidade* e *idade*, enquanto que as variáveis independentes linguísticas foram *acentuação da sílaba*, *tamanho da palavra* e *contexto seguinte [coronal]*. Além dessas variáveis, controlamos as variáveis agregadas *item lexical* e *indivíduo*.

Utilizando o teste da razão da máxima verossimilhança, identificamos as variáveis estatisticamente significativas para a variação entre a variante fricativa [h] e a aproximante [ɹ].

Na Tabela 5, a seguir, apresentamos as variáveis independentes estatisticamente significativas e que, conseqüentemente, foram incluídas no modelo de regressão logística multinível.

Tabela 5 – Variáveis independentes com significância estatística na realização da aproximante

	Total	% <sub>aproximante</sub>	Peso Relativo	Significância
Contexto seguinte [coronal]				2,18e-12
[não coronal]	311	3,2	0,17	
[coronal]	1.693	16,7	0,83	
Idade*Cidade <sup>12</sup>				4,59e-11
Sexo/gênero				3,74e-06
Feminino	986	9,1	0,35	
Masculino	1.018	19,9	0,65	
Escolaridade*Cidade				0,00028
Total	2.004	14,6		

<sup>12</sup> Devido a sua complexidade, os resultados dessa variável serão apresentados a seguir.

Entre as variáveis linguísticas, somente *contexto seguinte [coronal]* apresentou significância estatística para a aproximante. Observa-se que o traço [coronal] favorece a realização da aproximante (PR=0,83).

Vejam os exemplos retirados dos dados:

(20)

“Porque Deus é bom, mas eles atiraram para matar que [ko.i'to] o retrovisor do meu lado” (cortou) – AR76M08

(21)

“Nem por ['pa.ti] de pai nem por ['pa.ti] de mãe” (parte) – MC77F00

(22)

“A brincadeira era brincar de boneca jogar ['kɔ.iðɐ] roubar bandeira” (corda) – UP71F06

(23)

“Eu ia lá no ['fa.iðɔ] da carne” (fardo) – [SM84F01]

Como podemos verificar na Tabela 5, a realização da aproximante não está restrita à classe de consoantes [+obstruinte, -contínuo, coronal], pois encontramos tal variante antes de consoantes não coronais (/k/ e /g/), como em [ba.ɫ'kasɐ] ‘barcaça’ e ['ka.ɫgɐ] ‘carga’. Entretanto, o peso relativo obtido para essas ocorrências é muito baixo: 0,17.

Considerando que a variante ocorre somente seguida de oclusivas (Tabela 3), podemos concluir que o processo se restringe à presença dos traços [+obstruinte, -contínuo] no contexto seguinte e, dentro desse recorte, é fortemente favorecido pela presença do traço [coronal], ou seja, pelas consoantes /t/ e /d/.

Vejam, no que segue, os resultados encontrados pela análise das variáveis sociais em relação ao fenômeno estudado.

O gráfico a seguir apresenta o resultado da interação entre as variáveis *idade* e *cidade*. Logo após o gráfico, apresentamos a Tabela 6 com as significâncias das variáveis analisadas para cada cidade com o objetivo de analisar o efeito da idade em cada uma das cidades.

Gráfico 1 – Interação das variáveis idade e cidade na realização da aproximante.

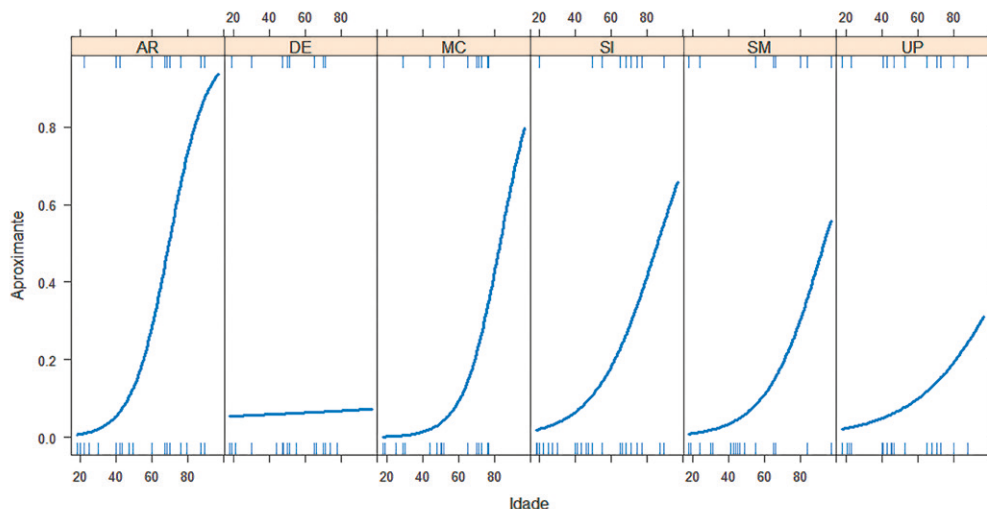


Tabela 6 – Interação entre as variáveis idade e cidade na realização da aproximante

Cidade	Sig <sup>TRMV idade</sup>
Arapiraca	0,000124
Delmiro Gouveia	0,741
Maceió	0,00498
Santana do Ipanema	0,0375
São Miguel dos Milagres	0,03668
União dos Palmares	0,000231

Pela análise do Gráfico 1, podemos concluir que o uso da aproximante aumenta de acordo com a idade do falante. Observamos, entretanto, que o efeito da idade não é o mesmo em todas as cidades, daí o efeito estatisticamente significativo da interação. Somente em Delmiro Gouveia tal relação não é estatisticamente significativa (Tabela 6). Para todas as demais cidades, podemos concluir que há um processo de mudança linguística em progresso com tendência ao desaparecimento da aproximante. Em Delmiro Gouveia, o processo já está mais avançado, visto que até entre os idosos é muito baixo o índice de realização da aproximante.

Em relação à variável *sexo/gênero*, os resultados apresentados na Tabela 5 mostram que os homens realizam mais a aproximante do que as mulheres, apresentando um peso relativo igual a 0,65, enquanto as mulheres apresentam PR=0,35. Não há interação entre as variáveis *sexo/gênero* e cidade, o que indica



que a diferença entre os *sexos/gêneros* pode ser observada em todas as cidades, em proporções similares.

Retomando o que vimos para São Miguel dos Milagres e União dos Palmares, quando analisamos a interação entre as variáveis cidade e escolaridade, observamos que nessas cidades quanto maior a escolaridade, menor o apagamento. Esse resultado, somado ao resultado da análise da variável sexo/gênero, indica que o apagamento pode ser uma variante linguística estigmatizada entre os falantes das comunidades de fala.

Esses fatos observados mostram as diferenças de comportamento sociolinguístico já apontadas por Labov (1972) e Trudgill (1972): são atribuídas à maior consciência feminina do *status* social das formas linguísticas. Muitos trabalhos em sociolinguística fazem menção à relação existente entre padrões de estratificação social e o gênero, concluindo que as mulheres, independentemente de outras categorias sociais, tendem a usar mais formas padrão do que os homens. Labov (1990) argumenta que os homens usam uma frequência maior de formas não padronizadas do que as mulheres em situações estáveis, e que as mulheres geralmente são as inovadoras na mudança linguística.

A seguir, apresentamos, no Gráfico 2, os resultados da interação entre as variáveis cidade e níveis de escolaridade e, na Tabela 7, as significâncias da escolaridade para cada cidade pesquisada.

Gráfico 2 – Interação entre as variáveis escolaridade e cidade na realização da aproximante.

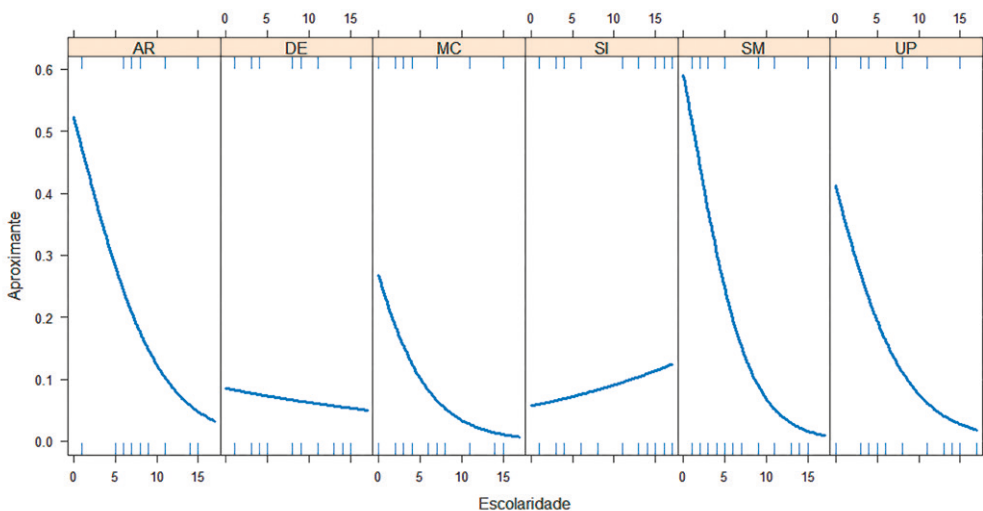


Tabela 7 – Interação entre as variáveis cidade e escolaridade na realização da aproximante

Cidade	Sig. <sup>TRM</sup> Vescolaridade
Arapiraca	0,029
Delmiro Gouveia	0,866
Maceió	0,036
Santana do Ipanema	0,573
São Miguel dos Milagres	0,044
União dos Palmares	<0,001

Quanto à escolaridade, nossa hipótese foi a de que, quanto menor a escolaridade, maior a probabilidade de produção da aproximante. Os efeitos da interação entre as variáveis cidade e escolaridade, apresentados no Gráfico 2, nos mostram que essa relação pode ser observada em Arapiraca, Maceió, São Miguel dos Milagres e União dos Palmares.

O nível de escolarização do falante e o contato deste com o ambiente escolar tendem a contribuir para o aumento do uso de variedades cultas, cujas formas são socialmente prestigiadas. É o menor uso da aproximante pelos mais escolarizados, pelas mulheres e pelos mais jovens que contribui para associarmos essa variante a algum estigma, especialmente em São Miguel dos Milagres, Arapiraca e União dos Palmares.

O processo analisado neste trabalho está diretamente relacionado ao tempo de escolarização dos falantes, uma vez que a aproximante não é largamente utilizada por todos os falantes das comunidades pesquisadas.

Seguimos ainda com a segunda análise sobre a variação entre a fricativa glotal [h] e a aproximante [ɹ], mostrando os resultados das variáveis independentes sem significância estatística na realização da aproximante.

Tabela 8 – Variáveis independentes sem significância estatística na realização da aproximante

	Total	% <sub>aproximante</sub>	Peso Relativo	Sig <sub>Wald</sub>	Sig <sub>TRMV</sub>
Tamanho					0,6596
dissílabo	1.052	14,8	*	*	
trissílabo	613	14,0	*	*	
polissílabo	339	15,0	*	*	
Acentuação da palavra					0,1691
Não acentuada	945	13,1	*	*	
acentuada	1.059	16,0	*	*	
Total	2.004	14,6			

A Tabela 8 revela que não há relação entre a variável tamanho da palavra e a realização da aproximante e que também não há diferença estatisticamente significativa entre sílaba não acentuada e sílaba acentuada no que diz respeito à variável acentuação da sílaba.

Apresentamos a seguir, na Tabela 9, os resultados para as variáveis agregadas *indivíduo* e *item lexical* na realização da aproximante.

Tabela 9 – Variáveis agregadas de efeito aleatório incluídas no modelo final na realização da aproximante

	n	Variância	CCI	Sig <sub>TRMV</sub>
Indivíduo	144	2.2600	40,7%	2.2e-16
Item lexical	425	0.6812	17,2%	0.0002803

Os resultados para o CCI dos níveis foram 40,7% para indivíduos e 17,2% para item lexical. Esse resultado nos permite afirmar que a interferência do item lexical é reduzida (17,8% da variabilidade entre a fricativa glotal e a aproximante é explicada pela variação entre os itens lexicais). Isso indica que a maior parte da variabilidade pode ser explicada pela variável *contexto seguinte [coronal]*, a única variável linguística estatisticamente significativa. Em relação ao indivíduo, a interferência é elevada: 40,7% da variabilidade pode ser explicada pela variação entre os indivíduos.

Mostramos, a seguir, também um processo secundário detectado durante a análise dos dados com foco no fenômeno central desta pesquisa, a variação de /R/ em coda silábica. Trata-se do processo de palatalização que envolve apenas as consoantes /t/ e /d/ precedidas de aproximante [ɹ], processo já apresentado

neste trabalho, que consideramos um achado valioso, pois parece fornecer uma evidência fonológica para a caracterização do segmento aproximante, uma das nossas variantes.

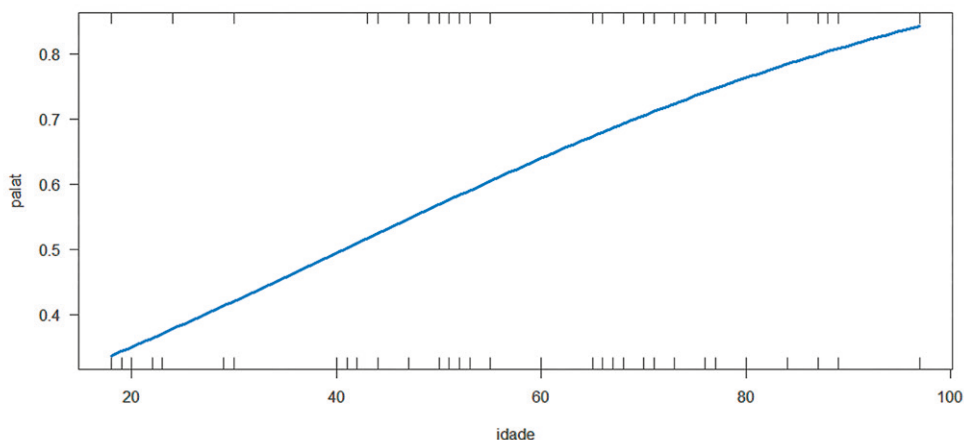
#### 4.1.2 Variações [t] ~ [tʲ] e [d] ~ [dʲ]

Para a análise desse processo de palatalização, selecionamos apenas as consoantes /t/ e /d/ precedidas de aproximante [ɹ], visto que a palatalização ocorre, basicamente, apenas nesse contexto. As variáveis sociais analisadas foram *sexo/gênero*, *escolaridade*, *cidade* e *idade*. As linguísticas foram *acentuação da sílaba*, *tamanho da palavra* e *tipo de consoante* (t ou d). As variáveis agregadas controladas foram o *item lexical* e o *indivíduo*.

Apenas as variáveis “*idade*, *cidade* e *escolaridade*” apresentaram significância estatística.

Vejam os resultados referentes a estas variáveis, no Gráfico 4 e na Tabela 10, a seguir.

Gráfico 3 – Variável idade no processo de palatalização.



Pela análise do gráfico, podemos concluir que a probabilidade de ocorrência da palatalização das oclusivas antecidas da aproximante [ɹ] aumenta nas faixas etárias mais altas. Apesar da frequência de palatalização ser alta (62,2%), temos indício de que a variante está em desuso em Alagoas.

Em seguida, os resultados para a variável *cidade*, também considerada significativa.

Tabela 10 – Variável cidade no processo de palatalização

	<b>Total</b>	<b>%<sub>palatalização</sub></b>	<b>Peso Relativo</b>	<b>Sig<sub>Wald</sub></b>	<b>Sig<sub>TRMV</sub></b>
Cidade					0,0003
Arapiraca	67	41,8	0,24	0,001	
Delmiro Gouveia	19	31,6	0,30	0,081	
São Miguel dos Milagres	70	72,9	0,56	0,425	
Maceió	39	69,2	0,58	0,401	
Santana do Ipanema	46	73,9	0,66	0,061	
União dos Palmares	42	71,4	0,67	0,055	
<b>Total</b>	<b>283</b>	<b>62,2</b>			

Pela análise da Tabela 10, concluímos que o processo é desfavorecido em Arapiraca e em Delmiro Gouveia. Já em Santana do Ipanema e União dos Palmares, há um maior favorecimento. Em nenhuma cidade, porém, houve significância em relação ao efeito médio, provavelmente devido ao tamanho reduzido desta subamostra.

Para finalizar o resultado das variáveis sociais consideradas significativas, apresentamos a variável escolaridade.

Tabela 11 – Variável escolaridade no processo de palatalização

	<b>total</b>	<b>%</b>	<b>PR</b>	<b>Sig</b>
<=9	781	25.6	0.72	0
>=11	912	9.1	0.28	0

Observando a Tabela 11, vemos que a variável escolaridade apresenta significância estatística, pois apresenta um diferencial considerável entre os participantes com maior e menor escolaridade. O processo de palatalização é favorecido significativamente pelos de menor escolarização, com um peso relativo de 0.72, enquanto os participantes que detêm maior nível de escolaridade apresentam 0.28 de peso relativo. Tal resultado corrobora com os dados da realização da aproximante, analisada na Tabela 7, que mostrou a interação entre cidade e escolaridade, revelando que quanto menor a escolaridade, maior a probabilidade de produção da aproximante. O mesmo acontece em relação ao processo de palatalização.

A seguir, discutimos os resultados das variáveis sem significância estatística.

Na Tabela 12, a seguir, estão postos os dados referentes a essas variáveis.

Tabela 12 – Variáveis independentes sem significância estatística no processo de palatalização

	Total	% <sub>palatalização</sub>	Sig <sub>TRMV</sub>
Consoante			0,896
t	201	62,2	
d	82	62,2	
Tamanho			0,785
Dissílabo	150	59,3	
Trissílabo	82	67,1	
Polissílabo	51	62,7	
Acentuação da sílaba			0,594
Não acentuada	120	65,8	
Acentuada	163	59,5	
Sexo/gênero			0,123
feminino	87	59,8	
masculino	196	63,3	

A variável *tamanho da palavra* não apresenta significância estatística para a realização da palatalização das consoantes /t/ e /d/ precedidas de aproximante [ɹ]. Quanto à variável *acentuação da sílaba*, concluímos que não há diferença estatisticamente significativa entre sílaba não acentuada e sílaba acentuada. Em relação à variável *sexo/gênero*, também concluímos que não há relação entre tal variável e a realização da palatalização de /t/ e /d/.

A seguir, apresentamos os resultados do Coeficiente de Correlação Intra-classe (CCI) para as variáveis *indivíduo* e *item lexical* envolvendo o processo de palatalização.

Tabela 13 – Variáveis agregadas de efeito aleatório incluídas no modelo final no processo de palatalização de [t e d]

	n	Variância	CCI	Sig <sub>TRMV</sub>
Indivíduo	63	0,1584	4,6%	0,136
Item lexical	110	0,4930	13,0%	0,408

Observando os resultados para as variáveis agregadas, podemos dizer que a interferência do item lexical e do indivíduo no processo de palatalização é bastante baixa (13% e 4,6%, respectivamente) e não apresenta significância estatística. Isso indica que há pouca diferença na variação entre os itens lexicais e que a variação no nível social é suficientemente explicada pelas variáveis *idade* e *cidade*.

## 5. CONCLUSÕES

A investigação mostrou que a variante tepe [r] encontra-se em extinção, pois em todo o *corpus* foram encontradas apenas 26 ocorrências, o que equivale a menos de 1% do total. Além disso, a sua presença restringe-se à fala dos idosos. Por não ser possível, dada a pequena quantidade de dados, fazer-se uma análise estatística referente a esta variante, limitamo-nos a uma análise descritiva.

Em relação à variante apagamento<sup>13</sup> [ø], mesmo não apresentando a análise estatística, faz-se necessário apresentar os resultados. Concluimos que esta não ocorre antes de consoantes soantes. Das variáveis linguísticas analisadas, o *contexto seguinte com a presença de traço [+contínuo]* é o que mais influencia o processo, com 0.97 de peso relativo, contra 0.03 para o traço [-contínuo]. Nossos resultados corroboram os resultados de Callou, Serra e Cunha (2015) no que concerne ao ambiente seguinte e também com a afirmação de Hora, Pedrosa e Cardoso (2010), sobre o apagamento ser condicionado pelo contexto fonológico seguinte, quando esse contexto é uma consoante fricativa.

Nossa hipótese inicial com relação à variante apagamento foi que o traço [+contínuo] poderia favorecer o apagamento devido ao princípio do contorno obrigatório, pois, segundo Goldsmith (1976), sequências adjacentes de unidades idênticas são proibidas nas representações fonológicas. Essa hipótese foi confirmada.

Entre as variáveis sociais, a variável *escolaridade* tem interferência somente em São Miguel dos Milagres, com <0,001 de significância e em União dos Palmares, com 0,048, permitindo afirmarmos que nessas cidades, quanto mais alto o nível de escolaridade, mais baixo o índice de ocorrência da variante apagamento. O efeito da variável escolaridade, nesse sentido, é bastante evidente para São Miguel dos Milagres. Vale notar que grande parte da variabilidade observada nos dados pode ser explicada, além de pelas variáveis sociais e linguísticas, pela variação entre indivíduos e itens lexicais, com índice de 44,8% e 56,9%, respectivamente.

Pode-se dizer, de modo geral, que a variação entre a fricativa glotal e o apagamento é estável.

A variante aproximante [ɹ], realiza-se diante de consoantes com os traços [+obstruinte, -contínuo], sendo que, dentro desse recorte, é fortemente favorecida pela presença do traço [coronal], ou seja, oclusivas /t/ e /d/, com peso relativo de 0.83. Ademais, os dados revelam que a variante está passando por uma mudança

---

<sup>13</sup> Análise estatística completa na Tese de Doutorado.

linguística em progresso com tendência ao seu desaparecimento, ratificando-se, desse modo, os resultados de Santos (2010).

Quanto ao processo de palatalização, um processo secundário analisado apenas com as consoantes /t/ e /d/ precedidas de aproximante [ɹ], concluímos que entre a variação das oclusivas há pouca diferença na variação entre os itens lexicais, com índices de apenas 13%, e indivíduo, com índice de 4.6%. A variação no nível social é suficientemente explicada pelas variáveis idade, cidade e escolaridade.

De acordo com toda a análise feita sobre a variável estudada, dentre as variáveis linguísticas somente a variável contexto seguinte tem relevância, pois explica o processo que causa o apagamento [ø], a realização como aproximante [ɹ] e o processo consequente de palatalização das consoantes /t/ e /d/. Acentuação da sílaba e outras variáveis linguísticas (como tamanho da palavra e contexto anterior) não mostraram relevância para explicar a variação.

Os resultados apresentados são importantes para caracterizar a variabilidade do /R/ no português falado em Alagoas, bem como para ampliar a compreensão desse fenômeno no PB de modo geral. A principal questão desta pesquisa foi entender como se caracterizava a variação ou o polimorfismo do /R/ nos falares alagoanos e, com isso, avançar na pesquisa iniciada por Santos (2010) quando apresentou linguística e estatisticamente a existência da variante aproximante no Litoral Norte de Alagoas.

A análise das variantes do /R/ no ambiente selecionado – fricativa glotal [h]; o apagamento [ø]; aproximante [ɹ] e tepe [r] – encontradas na fala de indivíduos de cidades alagoanas nos levou às seguintes descobertas: i) a fricativa glotal [h] é a variante mais frequente no Estado, seguida pela variante apagamento [ø], com a variante aproximante [ɹ], variante identificada apenas no estudo de Santos (2010), ocorrendo de modo significativo, sendo o tepe [r], variante bastante estudada no português brasileiro, praticamente inexistente nos falares alagoanos; ii) a variante apagamento encontra-se em processo de variação estável; iii) a aproximante configurou-se como passando por um processo de mudança linguística, melhor dizendo, está entrando em declínio.

É evidente que outros estudos sobre a variação do /R/ devem ser realizados, não somente nos falares alagoanos, como no português brasileiro. Em Alagoas, especificamente, é necessário buscar-se respostas para melhor explicar a relação do processo de apagamento, da realização da aproximante e do processo de palatalização com fatores linguísticos e sociais.



O trabalho que apresentamos aqui é apenas uma gota d'água no oceano de questões que ainda precisam ser respondidas com relação à classe dos róticos de modo geral e com relação ao comportamento do /R/ no Português do Brasil.

## REFERÊNCIAS

BARBOSA, P. A.; MADUREIRA, S. **Manual de fonética acústica experimental: aplicações e dados do português**. São Paulo: Cortez Editora, 2015.

CALLOU, D. **Variação e distribuição do tepe na fala urbana culta do Rio de Janeiro**. Rio de Janeiro: UFRJ-PROED, 1987.

CALLOU, D.; LEITE, Y.; MORAES, J. Processo(s) de enfraquecimento consonantal no português do Brasil. *In*: ABAURRE, M. B.; RODRIGUES, A. (org.). **Gramática do português falado**. v.8: Novos estudos descritivos. Campinas: UNICAMP; FAPESP, 2002. p. 537- 555.

CALLOU, D.; LEITE, Y.; MORAES, J. Variação e diferenciação dialetal: a pronúncia do [r] no português do Brasil. *In*: KOCH, I. (org.). **Gramática do português falado**. v.6: Desenvolvimentos. Campinas: UNICAMP, 1996. p. 465-493.

CALLOU, D.; MORAES, J. A.; LEITE, Y.; Consoantes em coda silábica: /s, r, l/. *In*: ABAURRE, M. B. M. (org.). **Gramática do português culto falado no Brasil**: volume VII: a construção fonológica da palavra. São Paulo: Contexto, 2013.

CALLOU, D.; OLIVEIRA, A. F. **A distribuição do processo de apagamento do rótico nas quatro últimas décadas: três capitais em confronto**. Revista Gelne. 2014.

CARVALHO, Lucirene da S. **Os róticos em posição de coda: uma análise variacionista e acústica do falar piauiense**. Tese (Doutorado). Universidade Federal da Paraíba, 2009.

CAMARA JÚNIOR, Joaquim Mattoso. **História e Estrutura da Língua Portuguesa**. Rio de Janeiro: Vozes, 1985.

CHOMSKY, N. & M, HALLE (1968). **The Sound Pattern of English**. New York: Harper & Row.

FREITAG, R. M. K. **(Re)**discutindo sexo/gênero na sociolinguística. *In*: Raquel Meister Ko. Freitag e Cristine Gorski Severo (orgs.). **Mulheres, Linguagem e Poder**: Estudos de gênero na Sociolinguística brasileira. Editora Edgard Blücher Ltda. 2015.

GOLDSMITH, John A. **Autosegmental phonology**. Bloomington: Indiana University Linguistics Club, 1976.

GUY, Gregory; ZILLES, Ana. **Sociolinguística quantitativa**: instrumental da análise. São Paulo: Parábola, 2007.

HORA, D.; PEDROSA, J. L. R.; & CARDOSO, W. **Status da consoante pós-vocálica no português brasileiro**: coda ou onset com núcleo não preenchido foneticamente? Porto Alegre, 2010.

LABOV, W. **Padrões Sociolinguísticos**. São Paulo: Parábola, 2008.

LABOV, W. **Principles of linguistic change**. Vol. I: internal factors. Oxford: Blackwell, 1994.

LADEFOGED, P. e MADDIESON, I. **The Sounds of the world's languages**. Blackwell, Cambridge, 1996.

LINDAU, M. **The story of 'r'**. UCLA Working Papers in Phonetics, 51, 1985.

MILROY, L. Social networks. *In*: CHAMBERS, J. K.; TRUDGILL, P.; SCHILLINGESTES, N. (ed.). **The handbook of language variation and change**. Oxford: Blackwell, 2004.

MONARETTO, V. N. O. Apagamento da vibrante pós-vocálica nas capitais do Sul do Brasil. **Letras de Hoje**, Porto Alegre, v. 35, n. 1, p. 275-284, mar. 2000.

OLIVEIRA, A. J. **Português alagoano**: proposta de constituição de um banco de dados de falares alagoanos. Maceió, UFAL, 2013. Projeto de Pesquisa.

OLIVEIRA, I.; SANTANA, M.; SERRA, C. R. **Apagamento do rótico em coda silábica interna e externa**: a região serrana do Rio de Janeiro. Artigo, Universidade Federal do Rio de Janeiro, 2014.

OLIVEIRA, A. A.; **Processos de palatalização das oclusivas alveolares em Maceió**. 2017. Tese (Doutorado em Letras e Linguística), Universidade Federal de Alagoas, Maceió, 2017.

SANTOS, J. S. B. **As realizações de ‘r’ em coda silábica na comunidade de Porto da Rua, litoral norte de Alagoas.** Análise linguística e sociolinguística. Dissertação (mestrado). Universidade Federal de Alagoas, Maceió, 2010.

SANTOS, L. F. **Realização das oclusivas /t/ e /d/ na fala de Maceió.** (Dissertação de Mestrado). Universidade Federal de Alagoas, Maceió: PPGL/UFAL, 1996.

SILVA, A. H. P. **Para a descrição fonético-acústica das Líquidas no Português Brasileiro:** dados de um Informante Paulistano. Dissertação (mestrado). Universidade Estadual de Campinas, São Paulo, 1996.

SILVA, T. C. **Fonética e fonologia do português:** roteiro de estudos e guia de exercícios. 8ª ed. São Paulo: Contexto, 1999.

TRUDGILL, Peter. Sex, covert prestige and linguistic change in the urban British English of Norwich. **Language in society**, v. 1, n. 02, p. 179-195, 1972.

TRUBETZKOY, N. **Principles of Phonology.** Trad. Christiane A. M. Baltaxe. Berkeley/Los Angeles: University of California Press, 1969.

WEINREICH, U.; LABOV, W. e HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística.** Tradução de Marcos Bagno. São Paulo: Parábola Editorial, 2006.

# ORGANIZAÇÃO DAS UNIDADES FONOLÓGICAS NO INÍCIO DA AQUISIÇÃO DE LINGUAGEM SOB A PERSPECTIVA DA FONOLOGIA PROSÓDICA E A AUTOSSEGMENTAL

## EFFECT OF THE ACCENT ON THE ORGANIZATION OF PHONOLOGICAL UNITS AT THE BEGINNING OF ACQUISITION UNDER THE PROSODIC PHONOLOGY AND AUTOSEGMENTAL PERSPECTIVE

*Luzia Miscow da Cruz Payão<sup>1</sup>  
Januacele Francisca da Costa<sup>2</sup>*

### 1. INTRODUÇÃO

Pensar em aquisição de linguagem e atuar na clínica fonoaudiológica nos transtornos de desenvolvimento de linguagem demanda uma fundamentação teórica consistente na Linguística. O aporte teórico linguístico possibilita ao profissional dessa área da Saúde compreender a estrutura e o funcionamento da língua, ampliando a compreensão do seu objeto de estudo e as possibilidades

---

<sup>1</sup> Universidade Estadual de Ciências da Saúde de Alagoas (UNCISAL).

<sup>2</sup> Universidade Federal de Alagoas (UFAL).

de atuar, efetivamente, diante de quadros, de fato evidenciados, dos distúrbios da comunicação.

Nas abordagens fonoaudiológicas visando à prevenção, promoção e intervenção de alterações de linguagem/fala, as tarefas de avaliação são norteadoras para o aprimoramento da investigação clínica e tomada de decisões terapêuticas. Assim, neste Capítulo se propõe a discussão de aspectos fonológicos de natureza prosódica e autosssegmental, que se manifestaram na fase inicial de aquisição e desenvolvimento típico de duas crianças entre 1;0.4 e 2;1.10 de idade, em contato com o português brasileiro falado em Maceió/ AL e Recife/ PE. As duas teorias que embasaram o trabalho – Fonologia Prosódica e Fonologia Autossegmental – constituíram-se como fundamentos para entender, atualizar e sedimentar práticas clínicas e dar continuidade às pesquisas interdisciplinares que vêm contribuindo para uma atuação fonoaudiológica mais científica ao abordar a linguagem.

Este é um estudo longitudinal, observacional e descritivo, que investigou a interação entre o acento e o preenchimento segmental das unidades prosódicas inferiores – sílaba, pé e palavra. As análises descritivas se restringiram ao preenchimento fonológico-fonético da sílaba segmentada, pé e palavra, baseando-se nas produções de fala espontânea dessas duas crianças, denominadas como S1 e S2, ambas do sexo feminino.

Para descartar a possibilidade de intercorrências orgânicas e/ou socioafetivas influenciando o desenvolvimento motor, cognitivo e linguístico dos sujeitos pesquisados, realizou-se, inicialmente, a anamnese com os pais das crianças, avaliação específica utilizada no curso de Fonoaudiologia da Universidade Estadual de Ciências da Saúde de Alagoas – UNCISAL, na triagem fonoaudiológica para detecção de distúrbios de comunicação.

Considerou-se, a partir dos históricos relatados pelos pais, que elas não apresentaram intercorrências orgânicas que pudessem suspeitar de comprometimentos e necessitassem, então, de um encaminhamento para outro profissional da área de saúde, visando às investigações pertinentes. Quanto à acuidade auditiva – fundamental para a aquisição fonológica –, ambas as crianças, quando recém-nascidas, se submeteram à Triagem Auditiva Neonatal (TAN) por meio do teste de emissões otoacústicas e da pesquisa do reflexo cócleo-palpebral (RCP), nos quais obtiveram respostas presentes bilateralmente.

A inclusão das crianças nesta pesquisa foi precedida de aprovação do projeto pelo CEP da UFAL, conforme processo de nº 006680/2008-81, aprovado em 30 de junho de 2008, e de autorização livre e esclarecida de seus responsáveis, que assinaram Termo de Consentimento Livre e Esclarecido, permitindo as visitas

domiciliares durante sete meses de registros por meio das gravações em áudio da fala espontânea das crianças. As amostras foram captadas em gravador digital de alta fidelidade, *Marantz Professional, Model PMD660 – Portable Solid State Recorder*; utilizando-se de microfone de lapela sem fio – *Sennheiser EW100-G2*, o que facilitava a movimentação da criança durante as interações lúdicas que motivaram as expressões verbais.

Os dados de fala das crianças foram motivados a partir de situações lúdicas e comunicativas na manipulação de brinquedos adequados para a sua faixa etária, explorados em conjunto com um dos pais ou ambos, caso a disponibilidade fosse possível. Esses dados de fala espontânea foram selecionados da sequência dos sete primeiros registros gravados em áudio, que ocorreram com intervalos de aproximadamente trinta dias, por um período máximo de quarenta e cinco minutos de gravação a cada visita.

As análises foram efetuadas sob uma base auditivo-perceptual, visando descrever a organização prosódica e segmental inicial da aquisição fonológica, a partir da sílaba segmentada, da estruturação acentual das palavras e dos preenchimentos segmentais nas sílabas conforme os alvos.

Foram selecionadas dos intervalos de gravação as palavras emitidas pelo menos duas vezes pela criança, para serem consideradas como fazendo parte de seu vocabulário, sendo transcritas com o Alfabeto Fonético Internacional (IPA, 1996). Desse modo, a cada gravação, as palavras produzidas pelas crianças foram discriminadas e computadas em tabelas, indicando-se a idade correspondente e classificando-as quanto ao número de sílabas e à acentuação tônica. Registrou-se, assim, tanto a inclusão de palavras novas como os ajustes silábicos e segmentais efetivados para o alvo ao longo das gravações realizadas.

Para complementar as análises dos dados de cada gravação, utilizou-se um segundo tipo de tabela, destinada ao inventário segmental das sílabas conforme a posição acentual nas palavras, relacionando os segmentos já manipulados nos preenchimentos silábicos das palavras. Tais segmentos foram distribuídos nas posições de sílabas tônica e de átonas pré-tônica e pós-tônica. Essas tabelas organizaram os dados de cada gravação e refletiram a expansão do sistema segmental, à medida que a criança conseguia produzir palavras com mais de três sílabas ou mais semelhantes ao alvo. Computaram-se também os totais absolutos de segmentos já presentes nessas posições silábicas da palavra, a fim de comparar a quantidade e a qualidade de classes de segmentos utilizados segundo a posição tônica e átona das palavras produzidas.

As palavras transcritas, seguindo esses critérios de inclusão, foram classificadas quanto ao número de sílabas, ao tipo de estrutura silábica manipulado, aos segmentos presentes e discriminados em relação à posição silábica tônica e às átonas pré-tônica e pós-tônica produzidas que visavam à efetividade da palavra-alvo.

Priorizou-se neste trabalho a fase inicial de fala das crianças, a fim de selecionar as produções que demonstrassem o efeito dos movimentos centrípeta-centrífuga gerados pelo acento na aquisição fonológica.

## **2. O ACENTO ENTONACIONAL E AS ANÁLISES CENTRÍPETA E CENTRÍFUGA**

Na aquisição, a matéria sonora resultante na fala se origina de movimentos inter-relacionados de análise centrípeta-centrífuga das unidades fonológicas constituintes (PAYÃO, 2010; PAYÃO; COSTA, 2016), obedecendo à hierarquia e organização estruturada da língua. Nessa movimentação ocorrem as transposições das informações e noções abstratas dos componentes gramaticais fonológicos e morfossintáticos.

Constata-se que as informações prosódicas iniciais contidas nos enunciados das crianças a partir do primeiro ano de vida confirmam as tarefas interpretativas do sistema prosódico, atuando nas informações originárias dos outros componentes gramaticais, como preconizado pela fonologia prosódica de Nespor e Vogel (1986).

Dessas operações mentais, que ocorrem inicialmente num nível prosódico e gradativamente em níveis segmentais, definem-se as representações básicas para o planejamento e realização motora da fala.

As manifestações de fala nas crianças refletem a confluência de operações mentais cognitivo-linguísticas, assemelhando-se ao movimento centrípeta, de caráter mais sintético. Os enunciados, reduzidos e incompletos em um primeiro momento em informações distintivas segmentais, evidenciam essa análise mais global.

Por conseguinte, constatam-se as produções de enunciados formados por uma sílaba, que representa uma palavra, a qual, geralmente, é a palavra foco do enunciado. Em seguida, as análises segmentais para preenchimento e revisão dos constituintes fonológicos ocorrem numa direção que se assemelha ao efeito de uma força centrífuga, orientada pela sílaba portadora do acento da palavra mais destacada, geralmente um nome, um verbo ou advérbio, ou seja, analisando e sintetizando os segmentos constituintes da sílaba mais proeminente,

construindo os demais níveis prosódicos acima dessa sílaba (PAYÃO, 2010; PAYÃO; COSTA, 2016).

A partir do enunciado fonológico, segmentado mediante a força centrípeta atuante no sistema prosódico do componente fonológico em direção à sílaba mais forte, representado na fala da criança por uma única sílaba portadora do acento entonacional, haverá o desencadear de uma força de direção oposta – centrífuga – suficiente para especificar a combinação de duas sílabas iguais e/ou diferenciadas esboçadas em palavras dissílabas e em combinações de palavras. Nessas combinações de palavras, já se constata duas classes gramaticais: o substantivo e o verbo, básicos para a formação dos sintagmas.

A estrutura prosódica permite a entrada de outros aspectos gramaticais que devem também ser analisados com o objetivo de intensificar a diferenciação entre as palavras e suas funções. A hierarquia prosódica atende a uma sequência de prioridades da estrutura linguística a serem percebidas e analisadas na fala.

A gramática interna da criança em aquisição fonológica inicial consegue analisar parcialmente num nível segmental, selecionando os fonemas possíveis para as posições esqueléticas específicas da sílaba e atendendo às características de sonoridade entre os segmentos. As diferenciações na implementação dos traços segmentais parecem ainda muito desafiantes, dependendo de amadurecimento percepto-motor para a análise fonológica distintiva e a efetividade na realização fonética.

Observa-se que os níveis fonológicos e morfossintáticos da estrutura gramatical estão se organizando, como, por exemplo, em S2 fazendo diferença entre nome e verbo na hora de construir os pés como em [mi'ninʊ taba'ja] (S2 – 1;6.3 'menino vai trabalhar') e em S1 – 1;5 [ta 'ki o 'tew] 'está aqui o Téo' e ['tɔmɐ] 'toma', enquanto dava peças do brinquedo para o pai. Nesses enunciados constatam-se as presenças de nomes, verbos e advérbios, manipulando as sílabas para a formação de pés métricos representativos de palavras dissílabas em S1 e S2, já delimitando todos os pés métricos das palavras núcleos de sintagma nominal e verbal.

Na sequência, então, são construídas as demais sílabas que compõem os pés de outras palavras, delimitando-se as palavras vizinhas que constituem os níveis prosódicos acima da palavra até alcançar o enunciado fonológico, componente mais alto da hierarquia prosódica.

Pode-se considerar que nesse preenchimento do material fonético, acento e sílaba estabelecem suas implicações mútuas, regidas pelas proeminências entonacionais observadas na partitura do enunciado. Nessa etapa de análise



linguística, possivelmente, já se processam, previamente, as informações lexicais e morfossintáticas, num nível de abstração mais profundo da gramática, embora esses sistemas lexical e morfossintático ainda sejam reanalisados e reestruturados ao longo da aquisição. Essa reestruturação reflete um modelo de aquisição interativo entre esses componentes gramaticais e o sistema fonológico da língua.

As proeminências entonacionais desempenham um papel de rastreamento da sílaba representativa e com *status* de palavra, e relevância no enunciado, que aglutinam o máximo de informações originárias dos componentes semântico, morfossintático e fonológico. O resultado desse rastreamento de informações provenientes dos componentes linguísticos será o esboço do algoritmo acentual por meio de derivações, que se manifestam nas primeiras palavras emitidas pela criança.

Conforme as experiências linguísticas se diversificam, ocorrem as tarefas fonológicas de segmentação do enunciado; paralelamente ocorre a estruturação gramatical em conjunto dos componentes formais da linguagem. O acento mostra-se como uma força impulsionadora de percepções diferenciadas do enunciado, obedecendo a uma gradação hierárquica nas tarefas cognitivo-linguísticas que interagem com as habilidades perceptuais e motoras.

Essa capacidade é considerada como uma sensibilidade inata de focar sua atenção na proeminência da fala adulta e interpretar a relevância em alguma unidade sinalizada por essa proeminência. Nessa capacidade expressa pela criança, há indício de uma estrutura primitiva da sintaxe na gramática interna, que, associada ao recurso do acento, auxilie na identificação da unidade prosódica relevante para a segmentação (SCARPA, 1997; 1999a; 1999b; NESPOR et al., 2008).

Desse modo, o acento é que conduz uma criança em aquisição a fazer o recorte na unidade mais destacada na fala do adulto, desencadeando, conseqüentemente, as segmentações básicas que propiciem as análises da estrutura da língua materna nos níveis prosódicos inferiores – sílaba, pé métrico e palavra.

Observa-se, no trecho apresentado a seguir, a tentativa de S1 (1;5) para repetir o enunciado da mãe e, afinal, o seu êxito, utilizando-se também do verbo *tome* numa situação similar de oferecer o brinquedo a alguém, como ilustra o Episódio 1 (PAYÃO, 2010, p. 27):

### **Episódio 1**

S1 (1;5), a mãe e o pai no chão da sala de sua casa explorando as peças de Lego, os objetos e os animais sendo retirados da caixa de brinquedos.

Mãe: tōmi essi ((mãe oferecendo um brinquedo à S1))

S1: [tʃə... tʃə... tʃ... ] ((tentativa de produzir a primeira sílaba do verbo *tome*, pronunciado pela mãe enquanto dava-lhe uma peça do brinquedo))

Mãe: tōmi essi ((oferecendo outra peça do Lego à criança, mãe e filha montando juntas os encaixes dos blocos de Lego para fazer uma casa))

S1 (emite balbucios)

Mãe: tōmi essi... ((novamente a mesma expressão ao oferecer um brinquedo à criança))

Pai: S1... cadê papai? S1... ((pai querendo participar também da brincadeira))

S1: [aw aw... ta'ki ɔ tɛw] ((*está aqui o Téo* – como se estivesse mostrando o cachorro ao pai))

Mãe: Téo né?! ((confirmando o reconhecimento do cachorro familiar feito pela criança, denominado de Téo))

Pai: essi é uma casa é S1?

Mãe: dê pa painhu... ((mãe mediando a situação da brincadeira))

S1: ['tɔmɪ... 'tɔmɪ...] ((*tome... tome...* – enquanto dava peças à mãe, utilizando a mesma expressão usada por ela pouco antes ao oferecer o brinquedo à S1))

Na sequência, no Episódio 2, apresentam-se dados de S2, com 1;6.3 de idade em interação com a mãe, ilustrando os enunciados da mãe e as segmentações efetivadas pela criança a partir dessa unidade prosódica (PAYÃO, 2010, p. 28).

## Episódio 2

S2 (1;6.3) e sua mãe no chão da sala onde residem, os brinquedos (blocos de Lego, trilhos do trem, carrinhos, animais figurativos e bonequinhos) estavam espalhados, as duas manuseavam os bonecos e os distribuíam nos carros ou na casinha enquanto interagiam.

Mãe: ah... eu achu qui essi mininu vai 'prescola' ou entãu eli vai passia... ou será qui eli vai trabalhá?

S2: [taba'ja] ((*trabalhar?* questionando à mãe))

Mãe: sim...

S2: [mi'ninu] ((*menino?* ainda questionando à mãe))

Mãe: sim... u mininu vai trabalhá? ((mãe já manuseando com a criança carrinho com bonecos e lhe perguntando, aguardando sua opção de resposta))

S2: [ɛ:... taba'ja] ((é: trabalhar... respondendo, finalmente, à mãe))

Propõem-se, portanto, direções de análises gramaticais impulsionadas pelo acento, que interpretam, segmentam e organizam a representação mental de natureza linguística, traduzida em material linguístico acentuado, expresso, no início, por um enunciado reduzido a uma sílaba. Essa força de atração pela proeminência toma direção centrípeta, por considerar o enunciado – o todo – desencadeando a segmentação da sílaba acentuada, resultante da atração convergente centrípeta.

Esse material linguístico é potencializado pelas informações superpostas de sistemas que desempenham tarefas diferenciadas na língua: as relações semânticas e o léxico mental, a sintaxe organizadora do léxico na pauta fonológica e prosódica. As segmentações são tarefas prévias e necessárias para análises e interpretações em cada nível dessa hierarquia prosódica; são desencadeadas pelo acento de proeminência entonacional do enunciado, até identificar a sílaba mais proeminente, localizada na palavra mais significativa e relevante desse enunciado.

A sílaba mais proeminente possui um caráter representativo das unidades prosódicas superiores, assumindo esse valor representativo mediante a análise de enunciado do *input* que se processou na direção centrípeta, atraída por essa sílaba proeminente, desencadeando, por fim, a segmentação.

Nespor e Vogel (1986, p. 221) explicam a interação de fatores fonológicos, sintáticos e de natureza lógico-semântica na reestruturação do enunciado, considerando-os como associados para a organização da gramática em geral:

What is particularly interesting about U-level restructuring, however, is that it depends not only on phonological and syntactic factors, but also on factors of a logico-semantic nature. Thus, at the highest level of phonological analysis, we find an interaction among several components of the Grammar, an interaction which has implications not only for the organization of phonology, but also for the organization of the grammar in general (NESPOR; VOGEL, 1986, p. 221).<sup>3</sup>

---

<sup>3</sup> “O que é particularmente interessante sobre a reestruturação ao nível-U, entretanto, é que ela depende não somente de fatores sintáticos e fonológicos, mas também de fatores de natureza lógico-semântica. Assim, ao nível mais alto da análise fonológica, nós encontramos uma interação entre vários componentes da Gramática, uma interação que tem implicações não

É respaldando-se nessa interação entre os componentes gramaticais precognizada por essas autoras, que se postula que a sílaba proeminente, unidade da hierarquia prosódica, é segmentada para dar início aos processos fonológicos de preenchimento segmental. A prosódia, então, fornece as estruturas básicas para a percepção da fala e organização da sequência de sons em palavras.

### 3. A GEOMETRIA DE TRAÇOS E OS PREENCHIMENTOS SEGMENTAIS DA SÍLABA NA AQUISIÇÃO

No modelo proposto por Clements e Hume (1995), também de base gerativa, e mais recente, o fonema é autossegmentado, representado por meio de uma configuração arbórea que expressa a segmentação independente de partes do som. Há uma interdependência e sobreposição das constrições no trato vocal, refletindo seu caráter dinâmico. Esse modelo da fonologia autossegmental, denominado geometria de traços, mostra a organização interna dos traços distintivos.

Ao final de seis meses do primeiro ano, a criança incorpora características prosódicas e segmentais da língua do meio e seus sistemas perceptuais atuam para a língua específica e perdem a habilidade para discriminar diferenças não contrastivas por volta de oito a dez meses de idade. Há indícios de aprendizagem específica afetada por um estágio, supostamente, somente fonético do balbucio e de discriminação no primeiro ano (MACKEN, 1995, p. 675).

Nesse sentido, Macken (1995, p. 675) considera o termo ‘pré-linguístico’ como impreciso para descrever as habilidades desse período do primeiro ano da criança e que existe uma relação entre esse período e o segundo ano de vida que reflete a relação complexa entre a fonética e a fonologia.

Quanto à habilidade de produção, Menn e Stoel-Gammon (1997) comentam que a prática aumenta o controle e a precisão dos movimentos que os bebês realizam, moldando o trato vocal para a produção de sons e sequências de sons, tornando-os mais automáticos e fáceis de executar na produção inicial das palavras. No período de vocalizações, ocorrem os dois tipos de *input* vocal: a fala dos outros e suas próprias produções, melhorando a habilidade motora e ouvindo as próprias vocalizações, formando circuitos de *feedback* de impressões táteis, cinestésicas e auditivas. Quanto mais balbucia, mais circuitos são estabelecidos e que proporcionarão o monitoramento da própria fala (MENN; STOEL-GAMMON, 1997, p. 280).

---

somente para a organização da fonologia, mas também para a organização da gramática em geral.” (NESPOR; VOGEL, 1986, p. 221) (Traduzido do original pelas autoras deste texto).

Inicialmente, a fala das crianças contém menor número de contrastes, influenciando no vocabulário de reconhecimento e produção de palavras. As formas produzidas estão relacionadas com a prosódia ou formas canônicas. Essas formas, por sua vez, contêm um número menor de elementos especificados, ou seja, são constituídas com menos informação do que as formas adultas correspondentes. Ocorre um mapeamento das palavras adultas e redução dessas formas de acordo com as capacidades de produção da criança (MENN; STOEL-GAMMON, 1997, p. 280-282).

Essas autoras ressaltam que a criança aplica regras e processos fonológicos, substituindo ou reduzindo as formas adultas, realizando assimilações e harmonias consonantais. Essas regras fonológicas se aplicam na produção dos sons de uma sequência individual, mas não combinados entre si.

Menn e Stoel-Gammon (1997, p. 291) comentam as limitações tanto no padrão silábico como nas palavras utilizadas na fala da criança, que se constituem pelo predomínio de formas canônicas CV e pela harmonia consonantal no seu *output*. As autoras acrescentam que a fonologia autosegmental e a prosódica estão estreitamente relacionadas às exigências psicofisiológicas do controle motor articulatório sequencial e da decodificação do sinal de fala. Essas exigências psicofisiológicas e motoras limitam as sequências de *output* (MENN; STOEL-GAMMON, 1997, p. 291). Concluem, por fim, que a fonologia autosegmental estabelece o formalismo necessário para tentar explicar os eventos da aquisição fonológica no que se referem aos efeitos da frequência de palavras, às diferenças individuais entre as crianças, aos limites das regras fonológicas operantes, aos efeitos da fala não regidos por regras. Há a compreensão de uma lacuna na fonologia da criança, em relação ao que ela consegue entender e o que ela realmente produz (MENN; STOEL-GAMMON, 1997; p. 291-295).

Nesta perspectiva, considera-se que na fala inicial, então, há a produção de sílabas representativas de níveis prosódicos superiores e que vão sendo preenchidos os ‘pedaços’ de palavras – pés métricos – e assim sucessivamente na reconstrução das palavras que constituem os enunciados. Esse material fonológico-fonético é resultante do analisador gramatical interno, gerador desses movimentos de análise centrípeto-centrífuga guiados pelo acento.

Assim, por trás da produção da linguagem falada pela criança, evidencia-se uma série de operações mentais em níveis suprasegmentais e segmentais que determinam as seleções das sequências sonoras, os segmentos específicos e as condições e restrições fonotáticas da língua ao nível fonético.

As evidências de S1 com 1;0.4 de idade deste trabalho mostram enunciados realizados por meio de uma sílaba canônica, de padrão CV (consoante-vogal), já se delimitando e sustentando a proeminência relacionada a uma dada palavra, como exemplificam as sílabas [tɛ] [tɛ] [tɛ] Téó, chamando pelo nome o cachorro da avó, enquanto manipulava esse animal figurativo de brinquedo e ao mesmo tempo a mãe referia-se ao cachorro familiar à criança – ‘olha o Téó, S1, chama o Téó’, ‘Téó, Téó, vem cá, Téó...’ –, significando e contextualizando as produções da criança.

Além da vogal preenchendo o núcleo da sílaba, há sempre a associação do *onset*, ocupado por um segmento obstruinte, caracterizado por plosivas, ou um segmento sonorante nasal. Constata-se nessas primeiras combinações de fonemas para formar sílabas, com valor de ‘palavra-enunciado’, a estruturação silábica atendendo ao princípio de sonoridade entre os segmentos (GOLDSMITH, 1995, p. 111), selecionando para *onsets* preferencialmente as plosivas, que apresentam a menor sonoridade e contrastam com a vogal nuclear da rima, que possui a máxima sonoridade. As pistas acústicas contidas no acento entoacional concorrem na segmentação e seleção de material segmental que será produzido pela criança.

Os dois fatores que interferem na compreensão do sinal linguístico são a fala contínua do output e a falta do léxico na criança nos primeiros estágios da aquisição de linguagem. Daí, a necessidade de estratégias que facilitem a segmentação do sinal linguístico. A criança, então, apoia-se na estrutura rítmica prosódica para dar início à tarefa de segmentação e delimitação das fronteiras de palavras (FROMKIN et al., 2006).

A entrada no léxico se dá por meio das bordas de palavras, nas quais a criança parece reconhecer um padrão métrico ou prosódico não marcado, no qual a primeira sílaba do pé métrico de palavras de conteúdo é acentuada. Além de mostrar um esboço de organização métrica, essa sílaba acentuada de palavra de conteúdo demonstra uma entrada lexical dessa palavra na gramática da criança, por meio do primeiro segmento da referida sílaba. Assim, há prevalência da sílaba acentuada do alvo, excluindo as partes menos salientes perceptualmente da palavra, ou a criança pode acrescentar sílaba epentética para preencher a forma prosódica completa da palavra (FROMKIN et al., 2006, p. 675-678).

Quanto aos segmentos que constituem sílabas, Toro et al. (2008, p. 1515), em estudo experimental com palavras trissílabas sem sentido com a estrutura CVCVCV, afirmam que pistas diferentes podem ser usadas para adquirir partes específicas da língua. Comentam sobre os papéis diferentes de consoantes e

vogais durante o processamento da linguagem, afirmando que as consoantes são preferencialmente envolvidas no processamento lexical, e as vogais tendem a marcar a constituição sintática por meio das pistas prosódicas.

Esses mesmos autores, a partir desse estudo, deram duas interpretações sobre as assimetrias entre C e V; afirmam que podem ser devidas às restrições inatas no modo diferente de processar as vogais e as consoantes a partir de estágios muito precoces; ou as diferenças acústicas e distribucionais desses segmentos podem progressivamente predispor o processamento das representações fonológicas no sistema (TORO et al., 2008, p. 1524). Esses estudos experimentais na fonologia mostram a associação de aspectos fonéticos e fonológicos para o entendimento de como a linguagem é processada pela criança.

Scarpa (1997; 1999a; 1999b) e Santos (2001; 2003a; 2003b; 2007) defendem que a aquisição da linguagem nas primeiras etapas apoia-se na compreensão dos contornos prosódicos declarativo, exclamativo e interrogativo associados à sílaba nuclear e aos fragmentos segmentais. Os sons preenchedores de lugares prosódicos são considerados como sinais de subespecificação fonética e gramatical se estruturando. São fragmentos em sequências gestálticas como um molde prosódico que irá direcionar a recursividade sintática. A massa fônica inicial da criança é um princípio de organização prosódica ainda em processo de estruturação (SCARPA, 1999a, p. 257-279).

Nesse mesmo estudo, a autora comenta sobre o balbucio tardio e recorrente na criança como uma “padronização do balbucio”, classificado como “formas prosodicamente indissociáveis” num nível segmental e suprasegmental. Considera que nos fragmentos “semelhantes a palavras” (SCARPA, 1999a, p. 257) ocorre a continuidade dessa padronização segmental e suprasegmental como um todo prosódico, que já exhibe um trabalho de estruturação e de formação simbólica entre significante e significado, embora ainda não sejam estruturas linguísticas organizadas. Preconiza a atuação simultânea da sintaxe e da fonologia na aquisição, em contínua organização e reorganização de sistemas ou subsistemas.

## 4. DISCUSSÃO

Desses aspectos levantados em relação aos modelos fonológicos adotados no estudo, entende-se que dar conta e explicar como se organiza a aquisição de linguagem exige uma visão global e uníssona dos componentes que permeiam o acesso a uma língua. As vertentes autosegmental e métrica prosódi-

ca associadas permitem a compreensão do que acontece na aquisição de sons da linguagem. Além disso, esses modelos relacionam o princípio hierárquico nas tarefas de perceber e organizar os conteúdos linguísticos, provenientes de níveis diferenciados.

Baseando-se nesses aportes teóricos e na hipótese de movimentos de análises centrípeta e centrífuga aqui defendida, afirma-se que o enunciado é o começo e o fim da organização hierárquica que rege a estrutura linguística. Para que essa hierarquia de tarefas linguísticas analíticas se processe, é preciso que se desenvolvam esses movimentos de saliências perceptuais diferenciados quanto à direção de análise.

Esses movimentos centrípetos e centrífugos iniciam-se, respectivamente, no enunciado e na sílaba; eles são pontos de referência nos quais as saliências perceptuais, advindas de proeminências acentuais de sintagmas entonacionais, são identificadas e categorizadas na representação mental. Assim, a saliência perceptual resultante pode se originar da proeminência advinda do enunciado, gerada, então, pela análise centrípeta. Essa análise se caracteriza por ser global e receber diretamente o suporte de informações não fonológicas para serem mapeadas ou ‘traduzidas’ em combinações fonológicas básicas, ou seja, a delimitação de uma sílaba representante dos níveis prosódicos superiores a ela. Portanto, é uma sílaba que representa parte do enunciado na dinâmica e interação em que se desenvolve o ato comunicativo. As pistas semânticas, pragmáticas e morfossintáticas concorrem para legitimar essa sílaba representativa do enunciado.

Dessas primeiras segmentações, oriundas da análise centrípeta em níveis prosódicos superiores – enunciado e frase entonacional –, desenvolvem-se, então, as análises de preenchimento centrífugas, regidas pela sílaba proeminente das palavras que a criança em aquisição começa a manipular. Portanto, a tarefa de análise linguística inicial parece obedecer a uma hierarquia de prioridades.

A fonologia prosódica e a autosegmental sustentam as análises centrípeta e centrífuga por apresentarem as justificativas que integram os componentes gramaticais a fim de explicar a fonologia da língua. Os modelos teóricos preconizam as inter-relações entre esses componentes gramaticais, a hierarquia dos constituintes e a conexão entre os níveis prosódicos, que permitem a coerência na formação da sílaba representativa, ou seja, a sílaba proeminente captada pela criança e significada, com estatuto dos níveis prosódicos superiores.

Essa hierarquia e interdependência entre os componentes gramaticais e unidades fonológicas fornecem a estrutura esquelética que concorre favoravelmente



no preenchimento de material linguístico necessário à realização da linguagem falada pela criança.

## 5. CONCLUSÕES

A prosódia, então, desempenha papel de relevância no processo de aquisição da linguagem, desencadeando esses movimentos de análises centrípeto-centrífuga na estrutura fonológica da língua a partir de pistas acentuais. Nessa perspectiva, a aquisição envolve a interdependência das unidades prosódicas e segmentais. Essa interdependência justifica e respalda a inserção de práticas de avaliação e intervenções terapêuticas no âmbito da linguagem que consideram tanto as segmentações originárias de enunciados do input, como os preenchimentos segmentais na reconstrução da hierarquia prosódica pela criança em aquisição.

Nesse sentido, as orientações e intervenções fonoaudiológicas na clínica de linguagem que priorizam as manifestações não verbais e verbais espontâneas, em situações de caráter lúdico e afetivo, tendem a promover e ampliar as parcerias comunicativas entre a criança e o adulto. Essas parcerias também colocam em evidência os enunciados como a unidade desencadeadora das segmentações prosódicas.

É nesse contexto que se encontra o ambiente propício para a criança em aquisição desenvolver a atenção básica e as habilidades de análises linguísticas centrípeto-centrífuga interpretativas e organizadoras da fonologia da língua, integrando processos perceptuais e de produção de fala.

## REFERÊNCIAS

CLEMENTS, G. N.; HUME, E. The internal organization of speech sounds. *In*: GOLDSMITH, J. (ed.). **The handbook of phonological theory**. Cambridge: Blackwell, 1995, p. 245-306.

FROMKIN, V. A. et al. (ed.) **Linguistics – An Introduction to Linguistic Theory**. Oxford: Blackwell, 2006.

GOLDSMITH, J. A. **Autosegmental and metrical phonology**. London: Blackwell, 1995.

MACKEN, M. A. Phonological acquisition. *In*: GOLDSMITH, J. (ed.). **The handbook of phonological theory**. Cambridge: Blackwell, 1995, 671-696.

MENN, L.; STOEL-GAMMON, C. Desenvolvimento fonológico. *In*: FLETCHER, P.; MAC WHINNEY, B. **Compêndio da linguagem da criança**. Porto Alegre: Artes Médicas, 1997, p. 277-295.

NESPOR, M.; VOGEL, I. **Prosodic phonology**. Dordrecht: Foris Publications, 1986.

NESPOR, Marina; SHUKLA, Mohinish; VIJVER, Ruben van de; AVESANI, Cinzia; SCHRAUDOLF, Hanna; DONATI, Caterina. Different phrasal prominence realizations in VO and OV languages. **Lingue e Linguaggio**, VII. 2, p. 1-29, 2008.

PAYÃO, Luzia Miscow da Cruz. **Aquisição de fonologia**: a influência do acento e o preenchimento de unidades prosódicas em dados de fala de duas crianças entre 1;0.4 e 2;1.10 de idade, em contato com o português brasileiro falado em Alagoas e Pernambuco. 192 p. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Letras e Linguística. Faculdade de Letras, Universidade Federal de Alagoas, Maceió, 2010.

PAYÃO, Luzia Miscow da Cruz; COSTA, Januacele Francisca da. Preenchimento de unidades prosódicas na aquisição fonológica inicial do português brasileiro. Porto Alegre: **Letras de Hoje**, v. 51, n. 3, p. 433-441, 2016.

SANTOS, Raquel Santana. **A aquisição do acento primário no português brasileiro**. 316 p. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. Campinas – SP, 2001.

SANTOS, Raquel Santana. Estratégias para aquisição do acento primário em PB. **Letras de Hoje**, Porto Alegre, v. 38, nº1, p. 171-188, março, 2003a.

SANTOS, Raquel Santana. A aquisição da morfologia verbal e sua relação com o acento primário. **Letras de Hoje**, Porto Alegre, v. 38, nº 4, p. 249-260, dezembro, 2003b.

SANTOS, Raquel Santana. O acento e a aquisição da linguagem em português brasileiro. *In*: ARAÚJO, G. A. de. (org). **O acento em português** – abordagens fonológicas. São Paulo: Parábola Editorial, 2007, p. 226-258.

SCARPA, Ester. Learning external sandhi: evidence for a Top-down hypothesis of prosodic acquisition. *In*: SORACE, A.; HEYCOK, C.; SHILLCOCK, R. (org.). **Proceedings of Gala 1997 conference on language acquisition: knowledge representation and processing**, 1997, p. 272-277.

SCARPA, Ester. Sons preenchedores e guardadores de lugar: relações entre fatos sintáticos e prosódicos na aquisição de linguagem. *In*: SCARPA, E. (org.). **Estudos de prosódia**. Campinas: Editora da UNICAMP, 1999a, p. 253-284.

SCARPA, Ester. Interfaces entre componentes e representação na aquisição da prosódia. *In*: LAMPRECHT, R. R. (org.). **Aquisição da linguagem – questões e análises**. Porto Alegre: EDIPUCRS, 1999b, p. 17-38.

TORO, Juan M.; SHUKLA, Mohinish; NESPOR, Marina; ENDRESS, Ansgar D. The quest for generalizations over consonants: asymmetries between consonants and vowels are not the by-product of acoustic differences. **Perception & Psychophysics**, 70 (8), p. 1515-1525, 2008.

## A PALAVRA NOMINAL EM YAATHE (FULNI-Ô), LÍNGUA INDÍGENA BRASILEIRA

*Fábia Pereira da Silva<sup>1</sup>  
Januacele da Costa<sup>2</sup>*

### 1. INTRODUÇÃO

#### 1.1 A língua e seus falantes

A língua Yaathe é falada pelo povo indígena Fulni-ô, que assim se autodenomina porque a região, ancestralmente habitada pelo povo que anteriormente era denominado Carnijó, está localizada no vale do Rio Ipanema, um afluente do Rio São Francisco. Fulni-ô significa, literalmente, “o que tem rio”. No meio da terra indígena, encontra-se a cidade de Águas Belas, sede do município do mesmo nome, situado na área de transição entre as regiões agreste e sertão de Pernambuco.

A população indígena Fulni-ô é de 4.687 indivíduos, de acordo com dados do Siasi/Sesai (2012). Atualmente os Fulni-ô estão distribuídos, basicamente, em duas aldeias<sup>3</sup>: a aldeia sede, localizada junto à cidade de Águas Belas, e a

---

<sup>1</sup> Doutora em Linguística, professora da UFAL, campus Sertão.

<sup>2</sup> Doutora em Linguística, professora aposentada da UFAL, campus Maceió.

<sup>3</sup> Muitos índios Fulni-ô também moram na cidade de Águas Belas, em outras cidades e na zona rural do município.

comunidade Xixiakhla, localizada a poucos quilômetros da aldeia sede, no local denominado Supriano. Essas aldeias são os locais de residência principais, pois são onde os índios moram durante nove meses do ano, no período que vai de dezembro a agosto. Há uma terceira aldeia, o Ouricuri, que é a aldeia sagrada, a seis quilômetros da aldeia sede, na qual todos os índios passam um período de três meses, de setembro a novembro, em um retiro secreto, mantendo suas tradições culturais e religiosas.

Os povos indígenas que sobreviveram ao massacre, tanto étnico quanto físico, promovido por diferentes instituições e motivos – no caso dos Fulni-ô, principalmente pela exploração dos não índios que governavam e impunham suas vontades nessa região – perderam elementos importantes do seu equipamento cultural, o que os diferenciava das populações não índias vizinhas e entre si. Das perdas de marca de identidade, uma perda notável foi a extinção de línguas nativas. Atualmente, entre cerca de 29 etnias indígenas que vivem no Nordeste<sup>4</sup> e que tiveram sua identidade étnica reconhecida e suas terras legitimadas até o final do século passado, só os Fulni-ô preservaram a língua herdada dos ancestrais, a que eles denominam Yaathe, que podemos traduzir literalmente como “nossa fala”.

De acordo com Rodrigues (1986), a língua Yaathe está filiada remotamente ao tronco Macro-jê, sem relação direta atestada com nenhuma outra língua indígena brasileira conhecida. Niculin (2020, p. v), em uma tese que teve por objetivo fornecer uma reconstrução de aspectos da língua ancestral do tronco Macro-Jê, incluindo domínios tais como a fonologia, a morfossintaxe e o léxico, conclui que

(...) fazem parte do tronco Macro-Jê as famílias linguísticas Jê, Jaikó, Maxakalí, Krenák, Kamakã, Karajá, Ofayé, Rikbáktsa e Jabutí; a família Chiquitana estaria relacionada às famílias supracitadas em um nível mais profundo. **Contrariamente a ideias anteriores de alguns autores, excluimos as famílias Boróro, Yaathê, Purí, Guató, Karirí e Otí da proposta, mas não descartamos a possibilidade de uma relação genética distante entre o tronco Macro-Jê, por um lado, e as línguas Boróro, Karirí e Tupí, por outro lado (Grifo nosso).**

Atualmente, há uma alta porcentagem de falantes de Yaathe entre os Fulni-ô, mais de 90% da população total (COSTA, 1993). O uso da língua é muito difundido na comunidade. Nas famílias, de modo geral, os membros se comunicam em Yaathe. Por exemplo, dão ordens ou fazem perguntas aos filhos nessa língua, a despeito de estes, às vezes, responderem em português. Estudos

---

<sup>4</sup> Fonte: RICARDO, B. & RICARDO, F. [Editores gerais]. Povos Indígenas no Brasil: 2006-2010. São Paulo: Instituto Socioambiental, 2011.

recentes (COSTA; SILVA, 2010) indicam que crianças muito pequenas dominam a gramática da língua, bem como outros aspectos particulares do seu uso, como, por exemplo, as diferenças de gênero social, que são um aspecto da gramática e do léxico. Assim, podemos afirmar que os Fulni-ô fazem uso sistemático de sua língua internamente e têm demonstrado um esforço muito grande para manter vivas a sua língua e a sua cultura, através de diferentes iniciativas.

## 1.2 Estudos sobre a língua: o estado da arte

Embora tenham sido elaborados, relativamente, muitos trabalhos sobre a língua, há ainda muito pouco registro do Yaathe, se considerarmos que esses trabalhos representam uma fração muito pequena de conhecimento, tanto da língua quanto do conhecimento ancestral dos seus falantes, do qual ela é veículo e receptáculo. Até a década de 1990, vários estudos de descrição e análise linguística foram elaborados. Entre eles, podemos citar Meland e Meland (1967), Meland (1968a-b), Lapenda (1968) e Barbosa (1991). Tanto Meland e Meland (1967) quanto Meland (1968b) são descrições da fonologia, elaboradas sob o modelo tagmêmico (PIKE, 1947). Lapenda (1968) descreve a estrutura da língua de um ponto de vista tradicional e Barbosa (1991) é uma descrição fonética e fonológica, também apoiada no modelo tagmêmico. De 1993 para cá, vários trabalhos foram efetuados sobre a língua: Costa (1993); Costa (1999); Silva (2011); Cabral (2009); Silva (2008); Melo (2010); Dias (2014); Sousa (2014).

O fato de a língua Yaathe constituir-se como uma língua viva, preservando todas as funções que se acredita que uma língua precisa cumprir em uma comunidade, é por si só merecedor de atenção, pois, como sabemos, na região Nordeste, a maioria das etnias indígenas – e, conseqüentemente, suas línguas – foi esmagada pelo processo colonizador, perdendo todo ou parte do seu equipamento cultural de identificação étnica, sendo a língua um fator muito importante dessa identificação. Por isso, uma das características mais notáveis da situação dos índios Fulni-ô é a sobrevivência da língua, uma vez que todas as outras línguas indígenas faladas nessa parte do país já são consideradas extintas (OLIVEIRA JR., M.; COSTA, J. F. e FULNI-Ô, F., 2014).

## 1.3 Objetivos do trabalho

Este capítulo é um excerto da tese de doutorado da primeira autora, intitulada *A palavra em Yaathe*, na qual se propôs, como objetivo central, a definição e delimitação da palavra na língua. Para alcançarmos o objetivo geral do trabalho,

a delimitação da palavra fonológica, uma outra unidade prosódica, hierarquicamente inferior à palavra fonológica e para a qual não havia estudos prévios em Yaathe, o pé métrico, precisou também ser estudado e, conseqüentemente, o sistema acentual da língua. Por outro lado, a análise dos dados levou o estudo para unidades prosódicas de nível mais alto do que a palavra, de modo que um esboço de descrição do sintagma fonológico, do sintagma entonacional e do enunciado fonológico também foi efetuado e apresentado. Neste capítulo, trazemos uma discussão do que seria uma palavra fonológica nominal em Yaathe.

## 2. APORTE TEÓRICO

Como já dissemos, definir palavra não é algo muito fácil, pois há várias maneiras para definir uma palavra e nem sempre há uma definição completamente satisfatória dessa unidade do ponto de vista prosódico ou fonológico. Alguns critérios, porém, que nos ajudam no desenvolvimento dessa tarefa, já foram delineados pelos estudiosos, embora, muitas vezes, alguns desses sejam critérios problemáticos, se considerados isoladamente. Autores como Aronoff & Fudeman (2007), Booij (2007) e Dixon e Aikhenvald (2010) e Haspelmath (2010) tratam de critérios que definem e delimitam uma palavra e também mostram os possíveis problemas enfrentados no empreendimento de tal tarefa, pois eles apontam que há muitos fatores que tornam difícil a definição e delimitação de palavra em uma língua, sendo um fator muito importante a falta de correspondência entre os significados lexicais e palavras, como bem observa Haspelmath (2010, p. 668):

While simplistic approaches such as Swadesh-list-based comparison make the simplifying assumption of a one-to-one correspondence between lexical meanings and words, and thus between words across languages, reality is more complex: words in one language are often in semantic many-to-many relationships with words in another language.

Isso pode ser apreciado em toda a sua complexidade quando coletamos dados elicitados do tipo listas de palavras. O conceito solicitado pode estar presente – e geralmente está – mas a forma correspondente frequentemente não pode ser fornecida porque não é uma palavra, ou o que convenciamos chamar de palavra<sup>5</sup>, ou o falante nos dá uma forma complexa que pode ser traduzida em uma sentença e, na maior parte do tempo, é uma sentença também em Yaathe.

---

<sup>5</sup> Essa observação parece fornecer um argumento para defender a hipótese que palavra existe na consciência do falante, pois eles fazem hesitações quando o conceito solicitado em forma de palavra na língua de trabalho, no nosso caso o Português, não encontra correspondência com a palavra do Yaathe.

Nossa análise baseia-se principalmente na Fonologia Prosódica, proposta por Nespor e Vogel (1986), que é uma teoria de organização do enunciado em unidades fonológicas organizadas hierarquicamente e concebe a língua como sendo representada por um sistema em que cada constituinte da hierarquia atua como contexto de aplicação de regras e de processos fonológicos específicos. As considerações teóricas são aplicadas na análise dos dados, embora não estejam aí explicitadas.

No que diz respeito às questões acentuais, apoiamo-nos em Liberman e Prince (1977) e em Liberman (1985).

Neste capítulo, especificamente, apresentamos parte da análise sobre a definição da palavra. Outras unidades que foram investigadas serviram de base e evidência para essa definição.

### 3. PROCEDIMENTOS METODOLÓGICOS

De uma perspectiva mais ampla, a metodologia utilizada é a tradicional da linguística descritiva que vem sendo aplicada às línguas indígenas brasileiras desde que o estudo científico dessas línguas foi iniciado pelas universidades.

Os dados primários, utilizados para análise e formulação de hipóteses, são oriundos do Banco de Dados do Projeto Documentação da Língua Indígena Brasileira Yaathe (Fulni-ô) e foram coletados entre 2011 e 2013, na aldeia Fulni-ô, município de Águas Belas, Estado de Pernambuco, Brasil. Esse projeto foi financiado pelo CNPq, Edital MCT/CNPq N. 014/2010 – Universal, Processo nº 475763/2010-6 e teve seu prazo de vigência encerrado em 11/11/2013. O projeto foi aprovado pelo Comitê de Ética em Pesquisa (CEP-UFAL), em 20/09/2011, Processo nº 012672/2011-70, tendo os participantes da pesquisa assinado o Termo de Consentimento Livre e Esclarecido (TCLE). Todos os documentos e dados aqui citados encontram-se arquivados na coordenação do Projeto e no banco de dados *The language Archive* em *IMDI Corpora, Donated Corpora*.<sup>6</sup>

O *corpus* é formado de três tipos principais de dados: 1) fala espontânea e semiespontânea, representada por diferentes gêneros – tais como narrativas, textos procedimentais, conversação em díades – e fala elicitada, através de listas de palavras; 2) escrita, também em diferentes gêneros, desde narrativas até materiais didáticos, incluindo poesias e outros exemplares escritos; 3) dados elicitados gravados no início dos anos 1990 e ao longo dos anos 2000, incluindo listas

---

<sup>6</sup> <https://corpus1.mpi.nl/ds/asv/?4>.



de palavras, questionários gramaticais e narrativas orais; 4) eventos culturais, como músicas e danças típicas, gravados em áudio e vídeo.

Entre os dados que coletamos, incluem-se listas de palavras e frases, tendo como modelo as já clássicas listas Swadesh (Swadesh, 1955), LDQ (Comrie & Smith, 1977), e aquelas propostas por Healey (1975), em seu manual de trabalho de campo, e uma série de exemplares discursivos, entre os quais narrativas de experiência pessoal, mitos, narrativas procedimentais e conversas espontâneas. Considerando que informações visuais têm reconhecida importância para a compreensão de determinados fenômenos linguísticos, gravamos também em vídeo a maior parte das sessões de coleta de dados em campo. Dados não linguísticos – como vídeos de danças, fotografias – e produção do povo, como desenhos, literatura, material didático, entre outros, também foram coletados e armazenados.

Os dados de áudio e vídeo foram gravados e arquivados respeitando-se todas as medidas e indicações propostas pela *E-MELD School of Best Practice*,<sup>7</sup> que vêm sendo adotadas em projetos de documentação de línguas indígenas internacionalmente, pelo *Open Archival Information System (OAIS)*,<sup>8</sup> que é um modelo de referência, com padrão AS (14721:2003), adotado pelos bancos de dados linguísticos mais recentes, e anotados seguindo os preceitos do *Metadata Encoding and Transmission Standard (METS)*,<sup>9</sup> também adotados por bancos de dados internacionais. Os dados de áudio foram gravados com microfones tipo headset DPA Headband 4066 e um gravador digital de flash Marantz PMD661. Os dados de vídeo foram gravados com uma filmadora Digital Sony Handycam HDR-PJ10, em full-HD e altíssima qualidade.

As transcrições e anotações das listas de palavras foram feitas no programa PRAAT<sup>10</sup> (Boersma & Weenik, 2007), uma vez que este aplicativo nos permite uma segmentação precisa no nível dos sons e dá acesso a detalhes acústicos dos dados, o que não apenas facilita a transcrição, nos mais diferente níveis, mas também auxilia na realização de estudos acústicos os mais diversos. A anotação das listas de palavras no PRAAT apresenta cinco fiadas: i) palavra (palv); ii) segmento (segm); iii) fonológica (fonl); iv) ortográfica (ortg);

<sup>7</sup> E-MELD School of Best Practice (<http://www.emeld.org/school/>).

<sup>8</sup> Consultative Committee for Space Data Systems, Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-B-1 Blue Book January 2002 (Washington, DC: CCSDS Secretariat, 2002). Disponível online: <http://public.ccsds.org/publications/archive/650xObl.pdf>.

<sup>9</sup> Library of Congress, “METS: Metadata Encoding & Transmission Standard” (2007), <http://www.loc.gov/standards/mets/>.

<sup>10</sup> Aplicativo computacional usado para transcrição ([www.praat.org](http://www.praat.org)).

e v) tradução (trad). Os textos foram anotados no ELAN,<sup>11</sup> alinhando-se os arquivos de áudio e vídeo. A anotação foi feita por meio de cinco fiadas: i) texto (tx); ii) palavra (wd); iii) morfema (mb); iv) glossa (gl); e v) tradução livre (ft). Os dois aplicativos possuem interface, o que significa que os dados de um podem ser importados pelo outro, permitindo uma ampla gama de utilização em análises e estudos diferentes, uma vez que as transcrições feitas tanto no PRAAT quanto no ELAN são compatíveis entre si.

A transcrição e tradução foram feitas com o auxílio dos professores de Yaathe, o que garantiu maior acurácia e proporcionará uma discussão acerca de um modelo adequado de grafia a ser adotado, com aprovação da comunidade. A anotação dos dados para disponibilização foi feita pelos pesquisadores do Projeto, envolvendo ainda estudantes de Iniciação Científica.<sup>12</sup>

Para outra parte dos dados linguísticos, que não foi transcrita, anotada e traduzida, respeitando-se, assim, a vontade dos falantes nativos da língua, uma vez que se trata de discursos e textos tradicionais pertencentes a áreas fechadas da cultura, como a religião, por exemplo, fizemos ainda o trabalho de anotação de metadados, obedecendo às exigências do banco, para que também pudessem ser arquivados em formato original. Será uma forma de preservar informações sobre essa parte da cultura para as futuras gerações, caso seja necessário, tendo em vista que a língua pode ser incluída entre as que correm risco de extinção, de acordo com relatório da Unesco 2010.<sup>13</sup>

Foram gravadas 4 listas de palavras, com dois informantes; 08 textos procedimentais com informantes de ambos os sexos e de idades variadas, tendo também como tema diferentes produções artesanais, já que artesanato, sobretudo em palha, é uma marca forte dos Fulni-ô; 06 textos narrativos; e 04 textos conversacionais.

O banco de dados do Projeto Documentação da Língua Indígena Brasileira Yaathe (Fulni-ô) está, a princípio, organizado seguindo o modelo de bancos de dados existentes, com algumas adequações necessárias, no sentido de se considerar e atender às características do grupo e da língua em questão. Os dados referentes a cada sessão foram organizados dentro de uma pasta, contendo documento de áudio, vídeo e anotação, bem como as informações sobre a sessão (Metadata).

---

<sup>11</sup> ELAN (EUDICO Linguistic Annotator) é uma ferramenta de anotação multimídia desenvolvida pelo Max Planck Institute for Psycholinguistics.

<sup>12</sup> Pesquisadores: Profa. Dra. Januacele da Costa; Prof. Dr. Miguel Oliveira Jr.; Profa. Ms. Fábria Fulni-ô. Iniciação Científica: Mariana Sila Sousa; Jéssica Gonçalves da Silva.

<sup>13</sup> <http://www.unesco.org/languages-atlas/en/atlasmap.html>.

Feita a análise e levantadas hipóteses que não puderam ser testadas nos dados disponíveis, empreendemos outras gravações de dados, dessa vez através de elicitacões direcionadas para os fenômenos específicos.

## **4. A PALAVRA FONOLÓGICA**

### **4.1 A palavra como unidade linguística**

A palavra é uma unidade linguística usada de várias maneiras na fala cotidiana, em muitos e diversos tipos de discursos linguísticos. Nesse sentido, é importante fazer algumas distinções fundamentais, de acordo com o que propõem Dixon e Aikhenvald (2007): distinção entre um lexema e sua forma variada; entre uma palavra ortográfica e outros tipos de palavra; e entre uma unidade definida principalmente em termos de critérios gramaticais e uma unidade definida principalmente através de critérios fonológicos. Nem sempre as categorias gramaticais se comportam da mesma forma entre línguas, pois nem todas as línguas apresentam as mesmas categorias para expressar as mesmas funções.

A forma gramatical de uma palavra apresenta uma interface entre a morfologia e a sintaxe e entre a morfologia e a fonologia. Dado que a morfologia lida com a composição das palavras, a sintaxe lida com a combinação das palavras e a fonologia com o arranjo dos sons na composição da palavra, podemos definir palavras de três perspectivas diferentes, classificando-as como palavras morfológicas, sintáticas e fonológicas.

No domínio da Fonologia, alguns critérios, tais como o fato de a palavra fonológica ser tipicamente o domínio do acento, a não ocorrência de pausas dentro de uma palavra fonológica, a aplicação de regras fonotáticas, considerando os traços segmentais da língua, e a aplicação de regras fonológicas, são importantes evidências da sua existência.

Nesta seção, definimos a palavra fonológica em Yaathe, utilizando os critérios fonológicos mencionados anteriormente, os quais discutiremos a seguir.

#### **4.1.1 CRITÉRIOS FONOLÓGICOS PARA A DEFINIÇÃO DA PALAVRA**

Ao tentarmos conceituar e separar palavra gramatical e palavra fonológica e examinar as relações entre essas duas unidades, há uma confusão sobre o que exatamente é uma palavra. A palavra é primariamente uma unidade gramatical com algumas propriedades fonológicas ou é uma unidade fonológica com

algumas propriedades gramaticais ou é igualmente uma unidade gramatical e fonológica? Essas ideias variam. Uma coisa certamente podemos dizer: a palavra é o elemento central do sistema linguístico. Não é somente o objeto da lexicologia, como também é necessariamente importante na fonologia, para a análise da estrutura dos sons, e na sintaxe, para as delimitações do seu *status* nas mais complexas configurações.

Para delimitação de palavra em uma dada língua, Dixon e Aikhenvald (2007) sugerem alguns critérios que combinam elementos gramaticais e elementos fonológicos.

Não há um único critério para definir e delimitar a unidade palavra fonológica em todas as línguas. Cada língua organiza e seleciona elementos que irão caracterizar uma palavra fonológica. Há um conjunto de tipos de critérios que definem essa unidade, o que parece servir para todas as línguas até então descritas, dos quais cada língua particular utiliza alguns.

Conforme o modelo que adotamos (NESPOR e VOGEL, 1986), a palavra fonológica é uma unidade hierarquicamente maior que o pé. Do ponto de vista formal, em algumas línguas ela pode ser apenas uma sílaba. Algumas propriedades são selecionadas para podermos reconhecer e fazer essa delimitação. Tais propriedades são características de três áreas: traços segmentais; traços prosódicos e regras fonológicas.

Os traços segmentais dizem respeito à estrutura interna da sílaba e às propriedades do segmento e as suas realizações fonéticas em combinação com outros segmentos (fonotática), ao fenômeno da fronteira de palavra e à pausa. Os traços prosódicos referem-se ao acento (*stress*) e ou acento tonal e a outros traços prosódicos, tais como nasalização, retroflexão e harmonia vocálica. Em muitas línguas o acento é um critério indiscutível para delimitar a palavra fonológica, uma vez que essas línguas têm acento fixo. Algumas regras fonológicas são aplicadas no domínio da palavra fonológica, enquanto outras podem ser aplicadas além desse limite (sândi externo, por exemplo). É possível que ocorra uma interação entre esses tipos de traços porque muitas regras fonológicas operam considerando o acento da palavra, por exemplo. Isso significa que esses traços podem ocorrer em combinação no nível da palavra fonológica.

Discutimos a seguir três propriedades que nos ajudam a compreender e delimitar a unidade palavra fonológica na língua Yaathe. São elas a pausa, o acento e processos fonológicos.

Para delimitar a palavra fonológica em Yaathe, consideramos propriedades que podem ser utilizadas na compreensão e delimitação da unidade palavra

fonológica na língua, como a pausa e o acento, e observamos, também, quais os processos fonológicos que poderiam nos dar pistas desses limites. A princípio, assumimos a descrição da palavra gramatical de acordo com o que considera Costa (1999).

A pausa é um critério híbrido, uma vez que pode ser considerado tanto no nível da fonologia como no nível gramatical. De acordo com Nespor e Vogel (1986), no meio de uma palavra fonológica não ocorre pausa. No entanto, para delimitarmos uma palavra gramatical, um dos critérios adotados é a possibilidade de haver pausa dentro de um contínuo, uma vez que uma palavra gramatical pode ocorrer separadamente, diferentemente de morfemas, por exemplo. Portanto, enquanto estamos tratando do domínio da palavra, não estabelecemos a duração da pausa a ser considerada, uma vez que estamos observando apenas se há possibilidade ou não de pausa.<sup>14</sup>

No que diz respeito à proeminência acentual, as línguas do mundo podem ser classificadas em dois tipos: i) línguas *pitch-accent* e ii) línguas *stress-accent*. Em línguas *pitch-accent*, o indicador primário da proeminência relativa é o *pitch*. Os correlatos fonéticos da proeminência relativa são de difícil estabelecimento em línguas *stress-accent*, pois diferentes propriedades fonéticas, como duração, *pitch* e intensidade, podem ser utilizadas nessa função.

Não nos preocupamos, por enquanto, em definir as propriedades fonéticas que assinalam a proeminência relativa em Yaathe, conforme já observamos. Consequentemente, também não nos preocupamos aqui em estabelecer se a língua é de um ou outro tipo. Seguindo Ewen e van der Hulst (2001, p. 196-197), utilizaremos o termo acento para caracterizar a propriedade abstrata de proeminência conforme esses autores indicam<sup>15</sup>.

No que se refere aos processos fonológicos, podemos dizer que o número de processos que operam na linguagem humana é finito, mas suficiente para criar uma grande variabilidade linguística. Em Yaathe, alguns processos geram um número grande de alofonias. Os processos que foram encontrados na língua Yaathe são processos de assimilação (desvozeamento, nasalização, palatalização,

---

<sup>14</sup> Na tese, na seção em que tratamos da delimitação de unidades superiores à palavra fonológica, tanto para a delimitação do enunciado fonológico quanto para a delimitação do sintagma entonacional, consideramos as pausas silenciosas superiores a 150 ms.

<sup>15</sup> O correlato fonético do acento poderia ser encontrado observando-se a frequência fundamental. Uma evidência para essa hipótese vem do fato que foram observados picos de  $F_0$ , no domínio do enunciado fonológico, correspondentes às sílabas acentuadas, conforme podemos ver na seção 4.

labialização, harmonia vocálica), alongamento compensatório, debucalização, fusão, redução e elisão de vogais. Na seção 1 da tese, uma descrição desses processos foi fornecida, de acordo com Costa (1999).

Alguns processos fonológicos que consideramos evidência para a delimitação da palavra fonológica em Yaathe são apagamento de vogal, fusão de vogais, apagamento de consoante, silabificação e ressilabificação da consoante, desvozeamento de oclusiva, alongamento compensatório e nasalidade.

## 4.2 A palavra nominal fonológica em Yaathe

Nesta seção, de acordo com o que propomos para este artigo, definimos a palavra nominal fonológica. Antes, porém, precisamos tratar de uma classe de elementos que precedem algumas classes de nomes, e que podem ser definidos como clíticos, exercendo a função de índice de posse ou determinação. Esse elemento enquadra-se na categoria clítico, morfema dependente, conforme proposto por Câmara Jr. (1969).

### 4.2.1 Clíticos pronominais

No Quadro 1, apresentamos os clíticos do Yaathe e a seguir discutimos os critérios que justificam a sua categorização, uma vez que, em relação a uma série de critérios, eles se comportam de modo diferente tanto de palavra autônoma quanto de sílaba em uma palavra fonológica. Note-se que, aqui, o nosso quadro é ligeiramente diferente daquele apresentado por Costa (1999).

Quadro 1 – Clíticos pronominais

Singular	1ª pessoa	i
	2ª pessoa	a
	3ª pessoa	e
Plural	1ª pessoa	ya
	2ª pessoa	wa
	3ª pessoa	tha
Reflexivo/recíproco		sa
Determinante genérico		se

Um primeiro critério utilizado para a definição dessas formas como clíticos é o fato que elas não contam como elemento na formação do pé métrico e não sofrem processos fonológicos que operam entre sílabas de uma mesma palavra, conforme demonstrado em Silva (2016, Seção 2, p. 45 e ss.). Nessa seção, a autora

define o pé métrico em Yaathe, apontando para o fato que existe uma diferença na atribuição do acento entre palavras nominais e palavras verbais.

(...) o acento é predizível em termos de janela acentual. Entretanto, sabemos que fatores adicionais podem ser considerados na atribuição do acento. Nomes comportam-se diferentemente de verbos, por exemplo. Enquanto em nomes (e outras formas não verbais básicas) o acento é fixo e não olha a estrutura da sílaba, em verbos a sufixação pode alterar a posição do acento, de acordo com as propriedades do afixo e modificação na estrutura silábica (SILVA, 2016, p. 45).

As palavras nominais recebem acento em uma das duas últimas sílabas, formando-se um pé trocaico ou iâmbico. Nesse sentido, o acento é predizível, dado que há uma janela de duas sílabas do lado direito da palavra onde o acento deve ser colocado. Entretanto, propomos que a instrução sobre o tipo de pé a ser formado é lexical, ou seja, aprendida juntamente com a aquisição do léxico. Exemplo dos dois tipos de acentuação em palavras de duas sílabas são 1) e 2), a seguir. Em 1) o pé é trocaico, ao passo que em dois temos um pé iâmbico, sem que para isso seja observada a estrutura das sílabas.

- |  |                                    |
|--|------------------------------------|
| 1) a. ['ta.tʃ <sup>h</sup> a] <i>lenha</i> | 2) a. [to.'we] <i>fogo</i>         |
| b. ['k <sup>h</sup> ɔ.tsa] <i>porta</i>    | b. [sɛ.'tʃi] <i>pele de animal</i> |
| c. ['fe.he] <i>pé</i>                      | c. [sɛ:.'to] <i>pássaro</i>        |

Para utilizar a atribuição do acento como um critério na definição de palavra fonológica, interessa-nos olhar as palavras de três sílabas ou mais e observar o acento secundário, conforme o algoritmo de acentuação proposto por Silva (2016, p. 49): (...) *com palavras de três sílabas ou mais, observe a posição do acento lexical, forme um troqueu (...) ou um iambo (...), conforme o acento primário recaia na penúltima ou na última sílaba, respectivamente. Atribua acento secundário alternativa e iterativamente da direita para a esquerda.*

Comparemos 3) a 4):

- |                     |                                 |
|---------------------|---------------------------------|
| 3) [klɛ.tal.'kã.ne] | <i>cobra venenosa (espécie)</i> |
| 4) [sa# ,natsa'ka]  | <i>seu próprio feijão</i>       |

Em 3), temos uma palavra fonológica autônoma como um todo. Em 4), temos uma palavra, [natsa'ka] *feijão*, que pode ser enunciada autonomamente, sem o elemento [sa], que representa reflexivização da posse. Em ambos os enunciados, o acento principal é fixo. Na primeira, porém, o algoritmo de acentuação

se aplica e o acento secundário recai sobre a quarta sílaba a contar do lado direito da palavra. Na segunda, a aplicação do algoritmo de acentuação forma um pé degenerado, o que demonstra que a sílaba [sa] não interfere nas regras de atribuição do acento, permitindo considerar essa forma um clítico.

Um critério bastante sólido para a categorização dessas morfemas como clíticos, em oposição à palavra nominal gramatical, é o fato de que eles entram em alguns processos fonológicos, tais como fusão de vogais e alongamento compensatório, enquanto em fronteiras de nomes autônomos esses processos não se efetuam. Inversamente, processos fonológicos que ocorrem entre fronteiras de morfemas, dentro de uma palavra, não ocorrem em fronteiras com clíticos. Um processo fonológico que opera automaticamente em Yaathe e efetua-se dentro da palavra fonológica é assimilação do traço nasal por uma vogal a partir de uma consoante nasal na sílaba seguinte. A regra de operação desse processo é a seguinte:

$$V \rightarrow [+nasal] / \_ \_ \left[ \begin{array}{l} +consonantal \\ +nasal \end{array} \right]$$

Em palavras com quatro sílabas, como no exemplo 1), a seguir, podemos ver a regra sendo aplicada.

$$1) /,una'kesa/ \quad \rightarrow \quad [,\tilde{u}na'kesa] \quad \textit{grupos de cafurna}$$

Entretanto a mesma regra não se aplica quando a sílaba que precede a consoante nasal é um clítico, conforme 2), a seguir.

$$2) /sa= natsaka/ \quad \rightarrow \quad [sa,natsa'ka] \quad \textit{meu próprio feijão}$$

O critério mais importante, porém, é que os clíticos são formas discursivamente dependentes, do ponto de vista discursivo, mesmo se podem ser autônomas do ponto de vista fonológico. Em Yaathe, quando associadas a uma raiz iniciada por consoante, clíticos possuem autonomia fonológica,<sup>16</sup> conforme demonstramos pelo critério de acentuação e pelo critério de operação do processo fonológico, podendo constituir uma palavra fonológica por si mesmo. Contudo, essas formas não ocorrem jamais isoladas no discurso.

<sup>16</sup> Quando a palavra começa por uma vogal, ao clítico é associada uma posposição com a qual ele forma uma palavra fonológica (Ver Silva (2016, seção 3.2.1)).



### 4.2.2 A palavra nominal fonológica em Yaathe

Na seção precedente, discutimos um tipo de palavra, os clíticos, que são formas dependentes, para que pudéssemos justificar a exclusão desse elemento como sendo parte da palavra nominal em Yaathe. Argumentaremos aqui em favor de se considerar que a palavra nominal, do ponto de vista gramatical, é constituída por uma base lexical e um ou mais sufixos.

No quadro a seguir, apresentamos o esquema de constituição dessa palavra.

Quadro 2 - Palavra nominal gramatical

Palavra Nominal Gramatical				
(clítico)	raiz verbal	(agente) (paciente) (locativo)	(tema)	(gen) (dim) (exc)
	raiz nominal	(privação)		

Uma palavra nominal gramatical pode ser constituída por uma base lexical, que pode ser uma raiz verbal ou nominal. A essa base, associam-se sufixos derivativos.

A uma base verbal associa-se um sufixo derivacional (agente, paciente, locativo), conforme exemplos 3a) e 3b); e os exemplos 4a) e 4b).

3a) [ta ,kejni'ka]

/ta= kejni -ka/

3SGS= ensinar -IND

*Ele ensina.*

3b) [ ,sekej'niho]

/se= kejni -ho/

DET.GEN= ensinar -AG

*O que ensina.*

4a) [ta ,fej'tõ:kia]

/ta= fejtone -ka/

3SGS= trabalhar -IND

*Ele trabalha.*

4b) [ ,fej'tõ:ho]

/fejtone -ho/

trabalhar -AG

*Trabalhador (o que trabalha)*

A uma base nominal associa-se um sufixo derivacional de privação (negação do que é expresso pela raiz), como ilustram os exemplos 5) e os exemplos 6a) e 6b).

5. [e:dʒa'dʷa]  
 /e= edʒa -doa]  
 3SGS= erro -PRIV  
*Deus (Aquele que não erra)*
- 6a) [e 'tʃedʷa]  
 /e= tʃe -doa/  
 3SGS= pai -PRIV  
 Órfão (aquele que não tem pai)
- 6b) [e 'tʃe'dʷãne]  
 /e= tʃe -doa -ne/  
 3SGS= pai -PRIV -FEM  
 Órfã (aquela que não tem pai)

Os afixos flexionais que uma base ou um tema nominal podem receber são os morfemas de gênero, diminutivo e exclusividade, conforme esquematizado no Quadro 2.

A seguir, utilizando os critérios fonológicos acentuação e/ou operação de processos fonológicos, demonstramos o pertencimento desses elementos à palavra nominal fonológica e, conseqüentemente, propomos a delimitação da unidade.

#### 4.2.2.1 Sufixos derivacionais

O sufixo /-ho/, quando associado a uma raiz verbal da classe de verbos ativos, deriva um tema nominal, ou um nome, cujo significado é “o que faz X”, onde X representa o significado de qualquer raiz da classe em questão. Vejamos os exemplos a seguir.<sup>17</sup>

- 7a) [ta ,kejni'ka]  
 /ta= kejni -ka/  
 3SGS= ensinar -IND  
*Ele ensina.*
- 7b) [ ,sekej'niho]  
 /se= kejni -ho/  
 DET.GEN= ensinar -AG  
*O que ensina.*
- 8a) [ta ,fej'tõ:kia]  
 /ta= fejtone -ka/  
 3SGS= trabalhar -IND  
*Ele trabalha.*
- 8b) [ ,fej'tõ:ho]  
 /fejtone -ho/  
 trabalhar -AG  
*Trabalhador (o que trabalha)*

<sup>17</sup> Na transcrição fonológica, indicamos o acento em sufixos inerentemente acentuados apenas na transcrição fonética, não o fazendo na transcrição fonológica ou nas glossas.

O sufixo /-ho/ tem um papel na atribuição do acento, embora não seja um sufixo acentuado lexicalmente, pois nunca recebe acento quando é o último elemento da palavra ou quando o sufixo que o segue é lexicalmente acentuado.<sup>18</sup> Observe-se que em ambos os exemplos o acento principal recai sobre a raiz verbal que serve de base para a derivação.

Na sequência, damos as formas de feminino correspondentes aos dois nomes citados.

- |   |  |
|---|--|
| <p>9) [ ,sekej'ni:so]<br/>         /se= kejni -ho -so/<br/>         DET.GEN ensinar -AG -FE<br/> <i>A que ensina.</i></p> | <p>10) [ ,fej'tõ'hoso]<br/>         /fejtone -ho -so/<br/>         trabalhar -AG -FEM<br/> <i>(a) trabalhadora</i></p> |
|---|--|

Em ambos os casos, uma sílaba é apagada. No primeiro exemplo, o sufixo de formação de nome de agente, constituído de /-hV/, sofre apagamento e causa alongamento compensatório da vogal da sílaba precedente. No segundo caso, a sílaba apagada é a mesma que se apaga também na forma verbal. A última sílaba da raiz /sekej'ni/ é acentuada, enquanto que a última sílaba de /fej'tone/ não é, motivo pelo qual ela é apagada na forma verbal antes do morfema de indicativo /-'ka/, conforme já demonstramos.

Enquanto o apagamento do morfema /-ho/ em [ ,sekej'ni:so] mantém o acento na raiz, o apagamento de /-ne/ em [ ,fej'tõ'hoso] não permite que a sílaba /ho/ seja apagada. Por isso, ela recebe acento.

Evidência de que esse morfema se encontra dentro da palavra fonológica é seu papel na atribuição de acento, conforme vimos anteriormente. Uma outra evidência é o processo fonológico que apaga a fricativa aspirada e associa o tempo da vogal com a vogal precedente, criando alongamentos compensatórios. Esse processo opera apenas no domínio da palavra fonológica, como mostram os exemplos anteriores, em que /-ho/ é mantido quando se encontra em final da palavra, mas pode ser apagado se outro morfema for associado ao nome.

Também mostra que essa regra é de domínio da palavra fonológica o fato que, quando o apagamento da fricativa ocorre em final de palavra, o efeito do seu apagamento é a transformação da vogal em glide e sua reassociação com a *coda* da sílaba precedente, como mostra o exemplo 11), a seguir.

<sup>18</sup> A distinção entre elementos acentuados e elementos não acentuados lexicalmente encontra-se em Silva (2016, seção 2, p. 45 e ss.).

- 11) [ˈfej ˈke]  
/fehe ke/  
pé LOC.POSP  
*No pé.*

Em seguida, damos um exemplo em que o sufixo de agente é seguido pelo sufixo de privação, sobre o qual recai o acento principal da palavra.

- 12) [i ta ˌfnihoˈdʷa]  
/i= tafni -ho -doa/  
1SGS= mandar -AG -PRIV  
*Eu não sou o que manda.*

O sufixo derivacional /-do'a/, quando associado a uma raiz ou a um tema nominal, deriva um nome, cujo significado é “privação de X”, onde X representa qualquer raiz da classe em questão, conforme ilustrado pelos exemplos 13), 14a-b) e 15).

- 13) [ˌeːdʒaˈdʷa]  
/e= edʒa -doa/  
3SGS= erro -PRIV  
*Deus (O sem erro.)*

- 14a) [e ˈtʃedʷa]  
/e= tʃe -doa/  
3SGS= pai -PRIV  
*Órfão (O sem pai.)*

- 14b) [e ˌtʃeˈdʷãne]  
/e= tʃe -doa -ne/  
3SGS= pai -PRIV -FEM  
*Órfã (A sem pai.)*

- 15) [e foˈdʷãne]  
/e= fo -doa -ne/  
3SGS= marido -PRIV -FEM  
*Mulher solteira (A sem marido.)*

O fato de esses morfemas possuírem um papel na atribuição do acento, como vimos até agora, mostra que eles são parte da palavra nominal fonológica. Desses morfemas, aqueles que podem associar-se ao morfema de gênero feminino /-ne/ também mostram que a regra de assimilação de nasalidade se aplica entre eles, o que é mais uma evidência para se postular que todos fazem parte da palavra fonológica a que se agregam. Outro processo fonológico, apagamento

de fricativa aspirada com alongamento compensatório, demonstra que /-ho/ é também um sufixo associado à palavra fonológica.

#### 4.2.2.2. Sufixos flexionais

O morfema /-ne/ *gênero feminino*, que ocorre na maior parte das raízes nominais que se modificam para expressar o gênero, comporta-se como parte da palavra fonológica em relação ao acento. Assim, em uma palavra nominal que, lexicalmente, possui acento na última sílaba, o acento é aí mantido depois que o sufixo /-ne/ lhe é associado, conforme se observa quando comparamos 16a-b) e 17a-b).

<p>16a) [ja'li]</p> <p>/jali/</p> <p>camarada</p> <p><i>Camarada! (informal, masculino)</i></p>	<p>16b) [ja'lĩne]</p> <p>/jali -ne/</p> <p>camarada -FEM</p> <p><i>Camarada! (informal, feminino)</i></p>
<p>17a) [i 'fĩ]</p> <p>/i= fĩ/</p> <p>1SGPOS= irmão</p> <p><i>Meu irmão.</i></p>	<p>17b) [i 'fĩne]</p> <p>/i= fĩ-ne/</p> <p>1SGPOS= irmão -FEM</p> <p><i>Minha irmã.</i></p>

Se, porém, mais um sufixo monossilábico seguir o sufixo de gênero, o acento principal desloca-se para a primeira posição à direita disponível para a sua colocação no nível lexical, conforme podemos ver quando o sufixo /-wa/, diminutivo, é acrescentado às formas apresentadas em 16a-b), aqui repetidas como 18a) e 19a), respectivamente.

<p>18a) [ja'lĩne]</p> <p>/jali -ne/</p> <p>camarada -FEM</p> <p><i>Camarada! (informal, feminino)</i></p>	<p>18b) [jalĩ'nẽwa]</p> <p>/jali -ne -wa/</p> <p>camarada -FEM -DIM</p> <p><i>Fulaninha!</i></p>
---	--

19a) [i 'ʃĩne]

/i= ʃĩ-ne/

1SGPOS= irmão -FEM

*Minha irmã.*

19b) [i ʃĩ'nẽwa]

/i= ʃĩ -ne -wa/

1SGPOS= irmão -FEM -DIM

*Minha irmãzinha.*

Comparando 18a) a 18b) e 19a) a 19b), verificamos que o acento principal, após a afixação de /-wa/ recai sobre a sílaba /-ne/, mantendo-se na penúltima sílaba da palavra.

No conjunto de exemplos a seguir, 20a-b-c), podemos observar a movimentação do acento em direção ao lado direito da palavra, à medida que um desses sufixos flexionais são associados ao complexo nominal.

20) [ˈtilʃi]

/tilʃi/

bonito

*Bonito*

20b) [ˌtilʃine]

/tilʃi -ne/

bonito -FEM

*Bonita*

20c) [ˌtilʃi'nẽwa]

/tilʃi -ne -wa/

bonito -FEM -DIM

*Bonitinha*

Uma classe de nomes forma o feminino pela afixação de um morfema com a forma subjacente /-'sV/, onde V corresponde a uma vogal não especificada subjacentemente, que harmoniza com a última vogal da raiz. Quando comparamos os exemplos 21a) a 21b), vemos que o acento principal, que na palavra base recai sobre a última sílaba, ao ser acrescentado o sufixo de gênero feminino, passa a recair sobre esse sufixo, agora ocupando o lugar da última sílaba da palavra.

21a) [i 'ka]

/i= ka/

1SGPOS= filho

*Meu filho*

21b) [i ka'sa]

/i= ka -sa/

1SGPOS= filho -FEM

*Minha filha*

Os exemplos 22a) e 22b) também pode ser comparados. Eles mostram o morfema de gênero feminino provocando a mudança do acento da primeira sílaba da palavra base para a segunda sílaba da palavra criada pelo seu acréscimo.

- |  |  |
|--|--|
| 22a) ['sa:who]<br>/sa:w -ho<br>briga -AG<br><i>Guerreiro</i> | 22b) [,sa:w'hoso]<br>/sa:w -ho -so<br>briga -AG -FEM<br><i>Guerreira</i> |
|--|--|

Um terceiro sufixo de gênero sempre ocorre com a forma [˜kʲa]<sup>19</sup> – onde “ : ” representa alongamento da vogal da sílaba precedente e “ ˜ ” representa nasalidade dessa vogal. Esse morfema se associa a algumas formas de nomes e, do mesmo modo que os demais morfemas de gênero, ele não recebe acento se ocorrer na última posição da palavra, mas modifica a posição do acento em palavras paroxítonas, conforme exemplos 23a-b).

- |   |   |
|---|---|
| 23a) [kla'i]<br>/klai/<br><i>homem branco</i> | 23b) [,kla'ĩ:kʲa]<br>/klai -neka/<br>homem branco – FEM<br><i>mulher branca</i> |
|---|---|

O morfema de diminutivo /-wa/ é outro sufixo não acentuado lexicalmente<sup>20</sup>. Ao agregar-se a um tema nominal, seja esse tema uma raiz primitiva ou derivada, ele pode tanto receber acento principal, quanto exercer influência sobre a atribuição do acento, como podemos ver na sequência de formação de palavras a seguir.

---

<sup>19</sup> Costa (1999, p. 103) atribui a esse sufixo à forma de base /ne'ka/, mas não fornece argumentos linguísticos que justifiquem a sua decisão. Nos nossos dados, a forma encontrada foi sempre [˜kʲa], com o acento recaindo sobre a sílaba longa precedente. Por isso, o consideramos um sufixo não acentuado lexicalmente. Consideramos que a forma fonológica é /neka/, embora não tenhamos encontrado evidências explícitas, do ponto de vista sincrônico, para sua recuperação. Entretanto, podemos apresentar algumas evidências encontradas em outros pontos da estrutura da língua, como o fato de a vogal nasal longa ser o resultado, em todos os casos em que ocorre, de um processo de alongamento compensatório causado pelo apagamento de /n/ em uma sílaba subsequente, ao mesmo tempo que a palatalização de /k/ também é o resultado de um processo de apagamento da vogal /e/ na sílaba precedente.

<sup>20</sup> Ver Silva (2016).

- |   |  |
|---|--|
| <p>24a) [i 'ka]<br/>/i= ka/<br/>1SGPOS= filho<br/><i>Meu filho</i></p>                              | <p>24b) [i 'kawa]<br/>/i= ka -wa/<br/>1SGPOS= filho -DIM<br/><i>Meu filhinho</i></p> |
| <p>24c) [i ,ka'wãna]<br/>/i= ka -wa -na/<br/>1SGPOS= filho - DIM-EXC<br/><i>Só meu filhinho</i></p> |  |

No exemplo 24a), o acento principal recai sobre a última sílaba da palavra e é aí mantido em 24b), em que é acrescido o morfema /-wa/. Essas duas formas respeitam a posição do acento das palavras no nível lexical, que, como vimos, recai sobre uma das duas últimas sílabas da palavra. Em 24c), porém, a manutenção do acento na posição original violaria essa regra e, por isso, o acento é deslocado uma sílaba à direita, recaindo sobre o morfema /-wa/, uma vez que o morfema seguinte é também lexicalmente não acentuado.

Nos casos examinados em 25a-c), observamos o seguinte: a) o acento de ['k<sup>h</sup>ɔf'kia] *palha* é alterado após a junção do sufixo /-wa/ *diminutivo*, passando para a sílaba imediatamente seguinte. Ao associar-se à palavra mais um sufixo /-na/ *exclusividade*, o acento não avança para esse sufixo, mas um processo fonológico apaga a vogal de /-wa/, associando o glide à sílaba precedente.

- |  |   |
|--|---|
| <p>25a) [ja 'k<sup>h</sup>ɔf'kia]<br/>/ja= k<sup>h</sup>ɔf'kia/<br/>1PLPOS= palha<br/><i>Nossa palha.</i></p>                            | <p>25b) [ja ,k<sup>h</sup>ɔf'kiawa]<br/>/ja= k<sup>h</sup>ɔf'kia -wa/<br/>1PLPOS= palha -DIM<br/><i>Nossa palhinha.</i></p> |
| <p>25c) [ja ,k<sup>h</sup>ɔf'kiawna]<br/>/ja= k<sup>h</sup>ɔf'kia -wa -na/<br/>1PLPOS= palha -DIM -EXC<br/><i>Só nossa palhinha.</i></p> |   |



Em 26a-b), temos uma palavra da Língua Portuguesa usada como empréstimo pela falante de Yaathe. Este é um exemplo interessante, onde podemos ver que a palavra em Português é [ta'peti] com acento na penúltima sílaba. Com a agregação do morfema diminutivo, o acento avança para a sílaba seguinte. A vogal [i] adquire qualidade de vogal plena porque passa a ser acentuada.

26a) [ta'peti]	26b) [,tape'tiwa]
<i>Tapete</i>	tapeti -wa
	tapete -DIM
	<i>Tapetinho</i>

O morfema de exclusividade /-na/, como podemos ver no exemplo 24a-c), que aqui repetimos como 27a-c) para demonstrar o fato descrito neste ponto, bem como no exemplo 28a-b) a seguir, ocupa a última posição na palavra.

27a) [i 'ka]	27b) [i 'kawa]
/i= ka/	/i= ka -wa/
1SGPOS= filho	1SGPOS= filho -DIM
<i>Meu filho</i>	<i>Meu filhinho</i>
27c) [i ,ka'wãna]	
/i= ka -wa -na/	
1SGPOS= filho - DIM-EXC	
<i>Só meu filhinho</i>	
28a) [wa'pela]	28b) [,wape'lãna]
livro	/wapela -na/
	livro -EXC
	<i>Só livro</i>

Esse morfema não recebe acento, dado que ele é um morfema não acentuado lexicalmente e por ocupar sempre a última posição na palavra. Como vimos, no entanto, ele tem um papel na atribuição do acento, causando o avanço do acento lexical para uma das posições permitidas no nível da palavra.

Além disso, como é um sufixo iniciado por consoante nasal, a vogal que o precede é nasalizada pela regra de assimilação de nasal, o que nos garante que ele pode ser interpretado como um sufixo, formado por uma única sílaba que se junta a uma outra sílaba para formar um pé de uma palavra fonológica.

Podemos, então, definir uma palavra nominal fonológica como um complexo de elementos, conforme mostramos no quadro a seguir.

Quadro 3 - Palavra fonológica nominal

Palavra Nominal Fonológica			
raiz verbal	(agente)	(tema)	(gen) (dim) (exc)
	(paciente) (locativo)		
raiz nominal	(privação)		

Em resumo, uma palavra nominal fonológica em Yaathe é formada pelos mesmos elementos que a constituem gramaticalmente, excetuando-se os clíticos pronominais. Ressaltamos que incluímos os clíticos na palavra gramatical porque palavras da classe que exigem um morfema de posse ou determinação são agramaticais na língua se enunciadas sem esse elemento. Ou seja, uma raiz dessa classe não se atualiza na fala sem o clítico.

## 5. CONCLUSÃO

Utilizamos neste trabalho, que teve como objetivo central definir e delimitar o que é uma palavra nominal fonológica em Yaathe, dois dos principais critérios sugeridos pela literatura da área.

O primeiro deles, a atribuição de acento, mostrou-se relativamente eficaz. Consideramos, porém, que existem muitos problemas com a sua utilização, dado que o sistema acentual de qualquer língua, sobretudo as pouco estudadas a esse respeito, como é o caso da língua Yaathe, é difícil de ser apreendido. Mesmo se na tese da qual este capítulo é oriunda buscamos descrever o acento na língua, o fizemos de um ponto de vista fonológico, sem identificarmos os seus correlatos acústicos. Isso, provavelmente, é um ponto problemático que precisa ser ainda explicitado. O critério operação de processos fonológicos através de fronteiras das unidades gramaticais (e não prosódicas) mostrou-se mais efetivo. Através dele, pudemos ver com mais clareza a integração dos elementos em um todo que podemos denominar palavra fonológica.

Definir palavra em qualquer língua não é uma tarefa simples, motivo pelo qual no nosso trabalho deixamos evidente que a proposta de definição e delimitação da palavra fonológica na língua é uma que pode e deve ser aperfeiçoada em trabalhos posteriores.

## REFERÊNCIAS

- ARONOF, M. and FUDERMAN, K. *What is morphology?* Oxford: Blackwell, 2007.
- BARBOSA, E. A. *Aspectos fonológicos da língua Yatê*. (Dissertação de Mestrado). Brasília: UnB, 1991.
- BOERSMA, P. e D. WEENIK. *Praat*. www.praat.org, 2007.
- BOOIJ, G. *The grammar of words*. 2ª Edição. New York: Oxford University Press, 2007.
- CABRAL, D. F. *O acento lexical em Yaathe*. 110 f. Dissertação de Mestrado – Universidade Federal de Alagoas, Maceió, 2009.
- CÂMARA JR., J. M. *Problemas de linguística descritiva*. Petrópolis: Vozes, 1969.
- COMRIE, B e N. SMITH. *Lingua Descriptive Studies: Questionnaire (= Lingua 42.1)*. Amsterdam: North-Holland. 72 pp, 1977.
- COSTA, J. F. *Bilinguismo e atitudes linguísticas interétnicas*. Aspectos do contato Português-Ya:the. (Dissertação de Mestrado). Recife: UFPE, 1993.
- COSTA, J. F. *Ya:thê, a última língua nativa do nordeste do Brasil: aspectos morfofonológicos e morfossintáticos*. (Tese de Doutorado). Recife: UFPE, 1999.
- COSTA; SILVA, 2010).
- DIAS, C. S. *O comportamento do traço nasal em Yaathe, língua indígena brasileira*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2014.
- DIXON, R. M. W. e AIKHENVALD, A. Y. (eds.). *Word. A Cross-Linguistic Typology*. Cambridge: Cambridge University Press, 2007.
- EWEN, C. J. e van der HULST, H. *The phonological structure of word*. An introduction. Cambridge: Cambridge University Press, 2001.

HASPELMATH, M. Comparative concepts and descriptive categories in crosslinguistic studies. *Language*, Volume 86, Number 3, September 2010, pp. 663-687 (Article), disponível em [http://scholar.google.com.br/scholar?q=haspelmath+comparative&btnG=&hl=pt-BR&as\\_sdt=0%2C5](http://scholar.google.com.br/scholar?q=haspelmath+comparative&btnG=&hl=pt-BR&as_sdt=0%2C5), acesso em 03/11/2014, às 08:51.

HASPELMATH, MARTIN; MATTHEW S. DRYER; DAVID GIL; and BERNARD COMRIE (eds.). *The world atlas of language structures*. Oxford: Oxford University Press, 2005.

HEALEY, A. (ed.) *Language learner's field guide*. Ukarumpa, EHD, Papua New Guinea: Summer Institute of Linguistics Printing Department, 1975.

LAPENDA, G. *Estrutura da língua Yatê, falada pelos índios Fulni-ôs em Pernambuco*. Recife: UFPE, Imprensa Universitária, 1968.

LIBERMAN, M. e PRINCE, A. S. *On stress and linguistic rhythm*. *Linguistic Inquiry* 8. 1977, p. 249-336.

LIBERMAN, M. *The intonational system of English*. (Dissertação de Mestrado). Massachusetts Instituto of Technology, 1985.

MELAND, D. e MELAND, D. *Word and morpheme list of the Fulni-ô Indian language*. Dallas, Texas: Summer Institute os Linguistics, 1968a.

MELAND, D. e MELAND D. Fulni-ô (Yathe) phonology statement. *Arquivo lingüístico n. 025*. Brasília, D. F: Summer Institute os Linguistics, 1967.

MELAND, D. Fulni-ô grammar. *Arquivo lingüístico n. 026*. Brasília, D. F: Summer Institute os Linguistics, 1968b.

MELO, J. A. *Gênero gramatical em Yaathe*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2010.

NESPOR, M. E VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.

NICULIN, A. Proto-Macro-Jê: Um estudo reconstrutivo. (Tese). Universidade de Brasília: Brasília, 2020.

OLIVEIRA, M., COSTA, J. F. e FULNI-Ô, F. Ética em documentação de línguas. In: FRANCISCO, D. J. e SANTANA, L. *Problematizações éticas em pesquisa*. Maceió: Edufal, 2014, p. 103-124.

RODRIGUES, A. D. 1886. *Línguas brasileiras*. Para o conhecimento das línguas indígenas. São Paulo: Loyola.

SILVA, F. *A sílaba em Yaathe*. Dissertação de Mestrado – Universidade Federal de Alagoas, Maceió, 2011.

SILVA, F. P. *Revisão da fonologia do Yaathe para uma proposta de uniformização da escrita na língua*. (Trabalho de Conclusão de Curso de Graduação). Maceió: UFAL, 2008.

SWADESH, M. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics*, 21, 121-137, 1955.

UNESCO, in: <http://www.unesco.org/languages-atlas/en/atlasmap.html>.

## E AGORA, JOSÉ? VOGAIS NASAIS FONOLÓGICAS EM YAATHE EXISTEM?

*Crislaini da Silva Dias<sup>1</sup>  
Januacele da Costa<sup>2</sup>*

### 1. INTRODUÇÃO

Nesses pouco mais de 500 anos de contato, ou melhor, de política de extermínio das populações indígenas e/ou dos seus modos de vida, mais de 1.000 línguas indígenas brasileiras foram extintas (RODRIGUES, 2003, p. 11). Atualmente, no Brasil, estima-se que são faladas aproximadamente 180 línguas indígenas, enquanto no nordeste brasileiro, entre a Bahia e o Piauí, apenas os Fulni-ô, cuja aldeia está situada no município de Águas Belas, Pernambuco, conseguiram preservar sua língua nativa, o Yaathe, classificada como pertencente ao tronco linguístico Macro-Jê (RODRIGUES, 1986).

Os trabalhos já realizados sobre essa língua versam sobre diversos aspectos, a saber: Costa (1999), trata da descrição de aspectos morfofonológicos e morfosintáticos; Cabral (2009), faz uma descrição do acento lexical no Yaathe; Silva (2011), faz uma análise da estrutura silábica do Yaathe; Silva (2016), apresenta

---

<sup>1</sup> Doutoranda em Linguística pelo Programa de Pós-Graduação em Linguística e Literatura da Universidade Federal de Alagoas.

<sup>2</sup> Professora de Linguística do Programa de Pós-Graduação em Linguística e Literatura da Universidade Federal de Alagoas.

aspectos da organização prosódica no Yaathe; Dias (2017), faz análise e descrição do comportamento do traço nasal no Yaathe; Sá (2017) faz um trabalho de documentação da língua Yaathe; e Sousa (2017), apresenta um estudo acústico-experimental da duração de vogais em Yaathe. Há também trabalhos com diversos enfoques na gramática da língua desenvolvidos por estudantes da graduação, como Cabral (2007); Silva (2008); Melo (2010); Dias (2014); Sousa (2014).

Através da revisão dos trabalhos anteriores, pode-se observar que a nasalização no Yaathe, em todos eles, era considerada apenas como sendo uma realização fonética, causada por regras fonológicas e morfofonológicas. Entretanto, nos dados que analisamos para este trabalho, encontramos ocorrências de nasalização da vogal sem que se possa atestar aplicação de regras do ponto de vista sincrônico. Isso nos permite levantar algumas hipóteses sobre a derivação de vogais nasais, baseadas em alternâncias, principalmente, e propor que algumas dessas vogais devem ser consideradas fonemas no estágio atual da língua. De modo geral, aqui<sup>3</sup> buscamos descrever um aspecto fonológico do Yaathe, o traço nasal, que ainda precisa de melhor conhecimento. Assim, nos dedicamos a estudar o comportamento desse traço na língua, isto é, procuramos verificar qual a sua função, em termos de distintividade, bem como o seu comportamento em relação às regras de assimilação.

Subjacente à descrição dos dados e à interpretação dos resultados da análise, estão *insights* de diferentes abordagens da fonologia e da fonética. Da fonética, observamos as noções básicas, tanto do ponto de vista articulatório quanto acústico, ainda que análises acústicas não estejam visíveis no corpo deste trabalho. Do ponto de vista da Fonologia, tiramos proveito das abordagens clássicas estruturais para a descrição dos dados, mas também procuramos nos apoiar em teorias mais modernas, como a Teoria Autossegmental (GOLDSMITH, 1976) e a Fonologia Prosódica (NESPOR e VOGEL, 1986), para uma melhor compreensão e explicação dos fenômenos descritos. Os modelos fonológicos apontam para diferentes modos de tratar os dados. Do ponto de vista estrutural, procuramos suporte para a descrição, organizando os dados de modo a entender sua estrutura, para, desse modo, podermos observar a sua distribuição na língua, distinguindo entre o que é fonológico e o que não é. A fonologia gerativa permite observar as regras que criam as alternâncias. Os modelos lineares nos dão uma perspectiva

---

<sup>3</sup> Este texto representa a tese central de dissertação de Mestrado intitulada “A Função e o Comportamento do Traço Nasal em Yaathe, Língua Indígena Brasileira” apresentada ao Curso de Linguística do Programa de Pós-Graduação em Linguística e Literatura (PPGLL), UFAL, sob a orientação da Profa. Dra. Januacele da Costa.

de observação do comportamento do traço em relação a espriamento, flutuação, estabilidade e outras das suas propriedades. Os pressupostos da fonologia prosódica auxiliam na explicação de fenômenos de espriamento ou de restrição à aplicação de regras nos domínios da estrutura prosódica.

## 2. PROCEDIMENTOS METODOLÓGICOS

O *corpus* utilizado neste trabalho compõe-se de dados oriundos de duas fontes distintas.

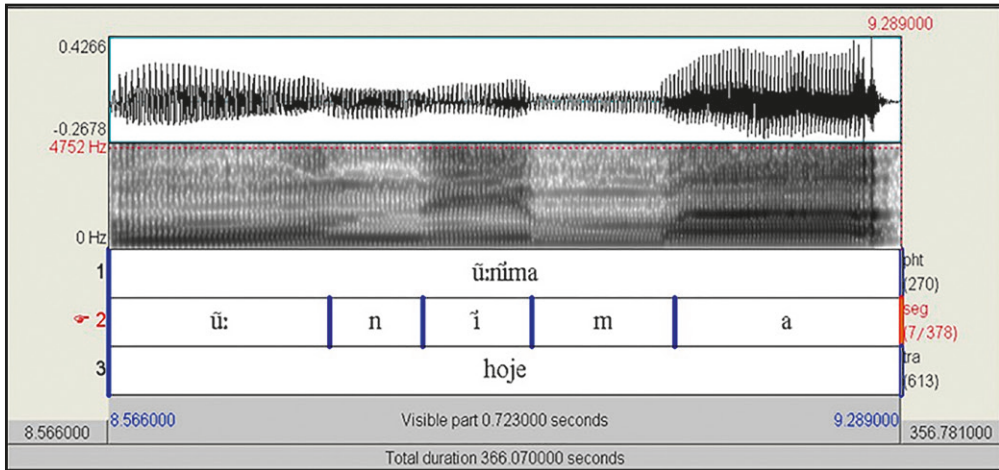
Uma parte está constituída por uma lista de palavras previamente selecionadas, de modo a apresentar todos os ambientes em que o traço nasal foi identificado a partir de *corpora* anteriormente compilados, basicamente os utilizados nos trabalhos de Costa (1999), Cabral (2009) e Silva (2011). Um *corpus* constituído por dados de fala espontânea foi coletado durante a disciplina “Seminários Temáticos em Teoria e Análise Linguística: Descrição de Línguas Indígenas”, ofertada pelo Programa de Pós- Graduação em Letras e Linguística da Universidade Federal de Alagoas (PPGLL-UFAL), em janeiro de 2016. Nessa ocasião, dois falantes nativos de Yaathe estiveram durante uma semana na UFAL para que os alunos da disciplina pudessem efetuar trabalhos práticos de coleta, transcrição, descrição, elicitación, análise e documentação de dados linguísticos. Os dados coletados para a análise do traço nasal no nosso trabalho são de um desses informantes, um falante adulto, do sexo masculino. A gravação dos dados de fala foi realizada mediante consentimento oral do informante e assinatura do Termo de Consentimento Livre e Esclarecido (TCLE), conforme orientações do Conselho de Ética da Universidade Federal de Alagoas. Nessa ocasião, gravamos uma lista contendo 148 palavras previamente selecionadas, buscando-se observar a ocorrência do traço nasal. Foi solicitado ao informante que nos fornecesse as palavras traduzidas para sua língua nativa, o Yaathe. Pedimos que ele repetisse a palavra duas vezes. As falas espontâneas versaram sobre temas escolhidos pelo próprio informante e, assim, referem-se a assuntos do cotidiano do nosso consultor, como natureza e trabalho. Os dados foram gravados em Yaathe e em seguida traduzidos para Português pelo mesmo indivíduo.

Para a gravação, utilizamos microfones tipo headset DPA Headband 4066 e um gravador digital flash Marantz PMD661 com frequência de amostragem de 24bit/96Hz. Para que pudéssemos ter uma melhor captura do sinal acústico, a gravação foi realizada em cabine acústica, no Laboratório de Fonética da Faculdade de Letras/UFAL. Durante as gravações também fizemos anotações manuais referentes à identificação dos consultores e informações extras de apontamentos



que eles enfatizavam e que julgávamos que seriam relevantes para nosso trabalho. Ainda no curso da disciplina, os dados foram codificados e transcritos no Praat,<sup>4</sup> como podemos ver na Figura 1, a seguir:

Figura 1. Espectrograma da palavra [ũ: 'nĩma] 'hoje'.



Fonte: Dias, 2017, p.10.

A transcrição foi realizada em três fiadas (*tiers*): (i) transcrição fonética da palavra; (ii) transcrição por segmentos; e (iii) tradução para o Português.

Feita a análise desses dados, consideramos que poderíamos encontrar outras ocorrências do traço nasal em dados de discurso natural, visto que a elicitación de palavras nos fornece dados em construções específicas a serem respondidas nas traduções apresentadas pelo informante, de modo que a informação que se obtém é a que foi solicitada. O problema de se trabalhar apenas com esse tipo de dado é que perdemos outros tipos de construções da língua (CHELLIAN e REUSE, 2011). Levando em conta esses problemas, procuramos também trabalhar com dados diversificados, tais como discurso natural, e explorar o corpus Projeto de Documentação da Língua Indígena Brasileira Yaathe (Fulni-ô),<sup>5</sup> um banco de dados coletados no período entre 2011 e 2013, na aldeia Fulni-ô, município de Águas Belas, Estado de Pernambuco, Brasil. Esse *corpus* é bastante

<sup>4</sup> Software capaz de, entre outras funções, medir ondas sonoras, formantes, níveis de frequência e intensidade dos dados que armazenam.

<sup>5</sup> Projeto financiado pelo CNPq, Edital MCT/CNPq N. 014/2010-Universal. (Processo nº 435763/2010-6) Que tem por objetivo formar um banco de dados da língua Yaathe, bem como de aspectos sociais e culturais, para a preservação e manutenção da língua.

extenso e composto por listas de palavras, fala espontânea de diversos gêneros, textos escritos e manifestações culturais. Os consultores do banco de dados do Projeto são homens e mulheres falantes nativos do Yaathe que também sempre viveram na aldeia. Segundo consta dos metadados, os ambientes para a gravação foram os mais silenciosos possíveis, dado que em situações naturais de fala, a coleta sendo realizada com o auxílio de microfones tipo headset DPA Headband 4066 e um gravador digital de flash Marantz PMD661. Todos os dados já se encontram nos arquivos do projeto transcritos e anotados. Para as listas de palavras, os realizadores do projeto utilizaram o aplicativo *PRAAT*, enquanto textos de fala espontânea, gravados em áudio e vídeo, estão transcritos e anotados no *ELAN*.<sup>6</sup> Para fins de análise, em nosso trabalho, exportamos o arquivo do *ELAN* para o *PRAAT*.

Sendo o banco de dados do Projeto bem abrangente, nos detivemos apenas nas listas de palavras, oito listas ao todo, e nos dados de fala espontânea, especificamente textos procedimentais e narrativas. Procuramos ocorrências de segmentos nasais em todos os textos desse banco, encontrando muito poucos casos de vogais nasais diferentes das já apresentadas nos trabalhos sobre a língua que citamos neste texto.

De modo geral, acessamos os arquivos e, para melhor identificar os ambientes de ocorrência do fenômeno que estamos estudando, fizemos um levantamento das palavras em que aparece a nasalização e também das palavras em que a consoante nasal não engatilha a nasalização da vogal. Para isso, separamos as ocorrências em início, meio e final de palavra. Na seção a seguir, apresentamos a descrição e análise dos dados.

### **3. O COMPORTAMENTO DO TRAÇO NASAL EM YAATHE**

#### **3.1 Informações preliminares**

De acordo com Costa (1999), o Yaathe apresenta um inventário de 33 fonemas, dos quais 21 são consonantais e 12 são vocálicos. Os únicos segmentos consonantais nasais na língua em estudo são /m/ e /n/. Para a autora (Idem, 1999), apenas são consideradas fonológicas as vogais do Quadro 1, a seguir.

---

<sup>6</sup> Software usado principalmente para sincronizar dados de áudio e vídeo e fazer anotações de diferentes tipos.

Quadro 1. Inventário fonológico das vogais do Yaathe

	Anteriores		Centrais		Posteriores	
Altas	i	i:			u	u:
Médias altas	e	e:			o	o:
Médias baixas	ɛ				ɔ	
Baixas			a	a:		

Fonte: SILVA, 2016, p. 13.

Foneticamente, podemos encontrar vogais com duração breve [u i o e ɔ ɛ a] e suas correspondentes nasalizadas, exceto para as médias baixas [ɔ ɛ], ou longas [u: i: o: e: ɔ: ɛ: a:] e suas correspondentes nasalizadas, exceto, novamente, para as vogais médias baixas. A seguir, apresentamos o inventário fonético desses segmentos.

Quadro 2. Inventário fonético das vogais do Yaathe

	Anteriores				Centrais				Posteriores			
	orais		nasais		orais		nasais		orais		nasais	
Altas	i	i:	ĩ	ĩ:					u	u:	ũ	ũ:
Médias altas	e	e:	ẽ	ẽ:					o	o:	õ	õ:
Médias baixas	ɛ	ɛ:							ɔ	ɔ:		
Baixas					a	a:	ã:	ã:				

Fonte: SILVA, 2011, p. 25.

Assim, para Costa (1999) e Silva (2016), o Yaathe apresenta 24 realizações de vogais, das quais, como vimos, apenas 12 são consideradas fonemas na língua. Para Costa (1999), todas as ocorrências de vogais com o traço nasal seriam consideradas realizações de superfície, causadas por processos fonológicos como nasalização automática e alongamento compensatório. Entretanto, em nossos dados encontramos vogais nasais que, a princípio, não podem ser consideradas como resultado qualquer desses processos, se abordado do ponto de vista sincrônico e em uma análise linear.

Na seção seguinte, apresentamos a descrição dos segmentos nasais, observando o comportamento do traço nasal no sistema fonológico do Yaathe, a fim de argumentar a favor da existência de dois tipos de vogais portadoras desse traço: vogais nasalizadas, não fonológicas; e vogais nasais, fonológicas.

Na apresentação dos exemplos, damos a forma fonética, a forma fonológica e a tradução. A anotação de morfemas é feita apenas quando a palavra apresenta estrutura interna depreensível. Embora o acento seja considerado fonológico (SILVA, 2016), não foi marcado na transcrição fonológica, dada a sua complexidade em termos de morfemas que são acentuados e morfemas que não são acentuados lexicalmente e, ainda, devido à interação desses fenômenos com regras de atribuição de acento que consideram também aspectos da estrutura silábica. Para melhor compreensão desses aspectos, ver Silva (2016).

## 3.2 Vogais nasalizadas

### 3.2.1 Assimilação

Como exposto anteriormente, há duas consoantes nasais na língua Yaathe: /m/ e /n/.

Enquanto a consoante nasal bilabial vozeada /m/ pode ocorrer tanto em *onset* como em coda silábica, a consoante nasal alveolar vozeada /n/ ocorre apenas em *onset*, ambas apenas em início e meio de palavra, como mostram os exemplos a seguir.

1) ['mũ.mja]	/mumja/	‘vara, graveto’
2) [t <sup>h</sup> o.mã.mã.kja]	/t <sup>h</sup> o.mamane -ka/ espantar -IND	‘espantar’
3) [se.tsõ.ne]	/se.tso.-ne/ índio-FEM	‘índia’
4) ['t.kã.no]	/tkano/	‘dois’
5) [mum.ni.'ka]	/mum.ni. -ka/ esfregar -IND	‘esfregar’

No que diz respeito à assimilação, as duas consoantes, quando em *onset* silábico seguinte, espriam o traço nasal para a vogal da sílaba precedente, nasalizando-a, como podemos observar nos exemplos (1) a (4). Desse modo, as vogais nasais que ocorrem nesse ambiente não são fonológicas. Um fato a ser observado é que o espraimento do traço não se efetua quando a consoante nasal – [m] – ocorre tautossilabicamente, conforme exemplo (5).

O primeiro argumento que temos para defender o *status* não fonológico dessas vogais é a inexistência de pares mínimos que demonstrem o contraste vogal oral/vogal nasal. Além disso, outro argumento forte que nos permite

sustentar essa hipótese, ou seja, que essas vogais não são subjacentemente nasais, mas resultado de assimilação, vem da análise de exemplos como os apresentados a seguir. Nesses casos, comparando (6a) com (6b) e (7a) com (7b), podemos observar que o traço [nasal] espraia de uma consoante nasal sobre uma vogal precedente quando um sufixo iniciado por consoante nasal é anexado a uma base terminada em vogal oral.

6a) [e'so]	/e= so/ 3SGO= outro	'outro'
6b) [e'sõne]	/e=so-ne 3SGO= outro-FEM	"outra"
7a) [i'fĩ]	/i= fĩ/ 1SGP= irmão	'meu irmão'
7b) [i'fĩne]	/i= fĩ -ne/ 1SGP= irmão -FEM	'minha irmã'

Isto posto, podemos propor uma regra de espraçamento do traço nasal de uma consoante à direita sobre uma vogal à esquerda em sílabas distintas. Ou seja, vogal oral passa a nasal antes de uma consoante nasal na sílaba seguinte. Essa regra, porém, apresenta restrições: vogais longas e vogais médias abertas não se nasalizam nos contextos favorecedores já descritos. Nos exemplos a seguir, vogais que precedem uma consoante nasal realizam-se sem o traço nasal.

8) [wene'ka]	'abrir'
9) [mo:mo]	'pimenta'
10) [to:na]	'coisa'

Um outro tipo de processo morfofonológico, que discutiremos a seguir, cria uma vogal nasal longa, que é também de superfície.

### 3.2.2 Alongamento compensatório

Vogais longas nasais em Yaathe são sempre, de acordo com os dados analisados neste trabalho, o resultado de processos fonológicos. O processo fonológico mais produtivo em termos de criação de vogais nasais longas é o denominado alongamento compensatório. Dois ambientes favorecem alongamento compensatório: i) apagamento do morfema /-ne/; ii) apagamento de uma morfema monossilábico constituído por /hV/, seguido de uma sílaba /ne/.

No primeiro caso, ocorre apagamento de uma sílaba /ne/. Essa sílaba pode ser um morfema derivativo, que é classificado morfologicamente como factivo ou causativo (COSTA, 1999), uma vez que cria verbos a partir de raízes nominais, ou pode ser apenas a sílaba final de uma raiz a que se associam às desinências das flexões verbais, sendo atestado mais amplamente com o morfema de indicativo [-ka/].

11)	[ 'kĩ:kja]	/kine -ka/ sentar -IND	‘sentar’
12)	[ 'kfẽ:kja]	/kfene -ka/ acreditar -IND	‘acreditar’
13)	[ 'kã:kja]	/kane -ka/ botar -IND	‘botar’
14)	[e' tʃõ:kja]	/etʃone -ka/ trazer -IND	‘trazer’
15)	[tʰlu'tʰlũ:kja]	/tʰlutʰlune -ka/ atrapalhar -IND	‘atrapalhar’

Todas as vogais, a princípio, podem sofrer essa alteração.

Alongamento compensatório do tipo apresentado aqui é um caso em que a nasalização da vogal ocorre de forma complexa, visto que se dá a partir do apagamento do sufixo /-ne/ nesses contextos. Em Costa (1999) encontramos a seguinte descrição em prosa das vogais longas criadas por alongamento compensatório:

- o segmento [+nasal +coronal] cai, depois de nasalizar a vogal da sílaba precedente;
- a vogal da sílaba precedente alonga-se por incorporar a unidade de tempo da nasal elidida.

[naha+ne+ka] → [nã:kja]

(COSTA, 1999, p. 69)

Outra complexidade desse processo deve-se ao fato de que nem todas as formas assim constituídas sofrem o processo de alongamento. Segundo Silva (2016, p. 50), as raízes que sofrem alteração diante do morfema /-ka/ são consideradas fracas: “Diferentemente dos nomes, no verbo, em uma forma atualizada, que é a forma de citação, conforme descrita acima, o acento é fixo sobre a última sílaba. Entretanto, há uma divisão nas raízes verbais, que podem ser fortes, com acento na última sílaba da raiz, ou fracas, com acento na penúltima

sílaba da raiz”. Nos exemplos (16) e (17) a seguir, temos casos em que o processo não se efetua, apesar de, aparentemente, todas as condições necessárias estarem presentes.

16) [eʃĩne'ka]	/e-ʃĩne-ka/ 3PSG -contar -IND	‘ele conta’
17) [wɛne'ka]	/wɛne-ka/ abrir -IND	‘abrir’

Além do morfema /-ka/, outros elementos que favorecem o disparo da regra são os morfemas de agentividade [-ho] (18); o morfema de particípio locativo [-se] (19); o morfema de particípio paciente [do'wa] (20).

18) [e'tʰlõ:ho]	/e= tʰlõne -ho/ 3SGS= segurar -AG	‘firme,seguro’
19) ['kĩ:se]	/kine-se/ sentar -PART.LOC	‘cadeira’ (onde senta)
20) [kĩ:do'wa]	/kine-dowa/ sentar -PART.PAC	‘sentado’

Também podemos observar que essa vogal longa e nasal ocorre em sílaba acentuada e em sílaba não acentuada. Silva (2016, p. 76) afirma que:

Nesses verbos, ocorre um processo de alongamento compensatório: o morfema /ne/, causativizador, é apagado com os traços flutuantes dos seus segmentos – nasal e coronal – espalhando regressiva e progressivamente para os segmentos adjacentes. A vogal longa criada por esse processo recebe o acento principal. Os fatores que motivam esses processos são: a) apagamento de vogal fraca no final da raiz devido à distinção entre raízes fracas e raízes fortes; b) restrição de nasal em coda silábica (estrutura silábica). A sílaba precedente, com um núcleo nasal ramificado, é vista como pesada para as regras de atribuição de acento na palavra verbal.

Apesar de esse tipo de sílaba ter sido considerada por Silva (2016) como sendo uma sílaba pesada em Yaathe, ela não é obrigatoriamente acentuada, pois só podem ser acentuadas sílabas que estão dentro da janela acentual da língua, que é de duas sílabas a contar do lado direito da palavra, conforme explicado por Silva (2016, p. 63): “Embora o acento recaia sempre sobre uma das duas últimas sílabas da palavra, ele não pode ser dito predizível, pois não é fixo. O acento em Yaathe é fonêmico, mas limitado em sua colocação, apresentando uma janela dissilábica do lado direito da palavra”.

Vale notar que a vogal nasal longa criada pela aplicação desses processos fonológicos também ocorre em palavras que não são verbos – do ponto de vista sincrônico e semântico, dado que as palavras podem ser resultado de mudança temporal não mais acessada pelos falantes – mas nomes, o que se dá quando um sufixo de gênero feminino é associado a uma raiz nominal, como mostram os exemplos a seguir.

21a) [se'tso]	/setso/	‘índio’
21b) [se'tsõ:kʲa]	/setso -neka/ índio -FEM	‘índia’

O alongamento compensatório pode ocorrer através do apagamento da sílaba /ha/ em verbos, quando se associa à forma básica qualquer um dos morfemas já descritos, além do morfema factivo [-ne].

22) [nã:ne'ka]	/naha -ne-ka/ ver-FAC-IND	‘mostrar’ (Lit.: fazer ver)
----------------	------------------------------	-----------------------------

No exemplo apresentado em (22), ocorre o apagamento da sílaba /ha/, que é parte da raiz do verbo. A nasalização da vogal da primeira sílaba se realiza porque essa vogal é agora a fusão de duas vogais, sendo que a segunda já era nasalizada automaticamente pelo espraiamento do traço nasal do morfema factivo. Podemos ter a seguinte formalização, representando a derivação da forma fonética em que se tem uma vogal nasal longa não acentuada.

Forma subjacente	/naha -ne -ka/
nasalização de vogal	nahãneka
apagamento de /n/	naãneka
alongamento compensatório	nã:neka
Forma de superfície	[nã:neka]

Embora não tenhamos submetido esse dado, especificamente, a experimento acústico a fim de medir a duração da vogal alongada, é possível perceber que, em relação aos demais casos, aqui a duração da vogal é maior, ocupa um período de tempo mais longo, uma consequência possível do apagamento de duas sílabas CV consecutivas.



### 3.3 Vogais nasais

Em todos os trabalhos anteriores, já citados, aos quais recorreremos para observar o tratamento descritivo que tem sido dado às vogais nasais em Yaathe, concluiu-se que vogais nasais ocorrem apenas antes de consoante nasal na sílaba seguinte, seja por assimilação simples, seja por alongamento compensatório, conforme já apresentamos. Isso quer dizer que, do ponto de vista da análise fonológica tradicional, que observa contrastes e distribuição complementar para identificar fonemas em uma língua, não pode ficar provado que existam vogais nasais fonológicas em Yaathe, uma vez que pares mínimos, pares análogos e alternantes complementares não foram encontrados. Entretanto, em nossos dados verificamos casos em que vogais nasais breves ocorrem fora de contextos nasais assimilantes, conforme descrito. É importante assinalar que ocorrências desse tipo não são raras na língua.

Nos exemplos de (23) a (27), apresentamos casos observados nos dados que analisamos. Propomos hipóteses que poderiam explicar a presença de algumas dessas vogais nasais e, contudo, afirmamos que, sincronicamente, elas podem ser consideradas nasais fonologicamente.

- |                            |                      |
|----------------------------|----------------------|
| 23) [i'fmã]                | ‘para mim’           |
| 24) [ã'hã]                 | ‘sim’                |
| 25) [k <sup>h</sup> ofe'ã] | ‘para’ (Benefactivo) |
| 26) [ẽ'hẽ]                 | ‘olhe’               |
| 27) ['nũfa]                | ‘daqui pra frente’   |

Observando esses dados, podemos propor, em caráter hipotético e com base em algumas evidências, que as vogais finais nasais em Yaathe são reflexos de uma antiga sílaba nasal [nV] que sofreu apagamento.

No caso de [i'fmã] ‘para mim’, expressões semelhantes, como a descrita no exemplo (28), parecem ser uma pista para confirmação da hipótese.

- |                |                    |               |
|----------------|--------------------|---------------|
| 28) [sak'mãna] | /sa ke -ma -na/    | ‘só para ele’ |
|                | ele LOC -FIN -EXCL |               |

Sugerimos, então, que [i'fmã] poderia ser uma forma reduzida de [i'fmãna] “só para mim”, cuja sílaba final teria sido apagada com estabilidade do traço [nasal], desde que este é um traço flutuante, sobre a vogal precedente. Estabilidade do traço nasal é bastante produtiva na língua, como vemos nos casos

de vogais longas nasalizadas que são resultado de apagamento de sílaba nasal seguinte, constituindo alongamentos compensatórios. Esses exemplos confirmam que ocorrem vogais nasais breves em Yaathe que não são, sincronicamente, resultado de operações fonológicas e que, assim, devem ser arroladas como fonemas da língua.

Em alguns casos, há uma alternância entre a pronúncia da vogal em sílaba final entre vogal oral/vogal nasal. O contexto seguinte, nesse caso, é [Ø], pelo menos no domínio da palavra. Observemos, a princípio, os exemplos (29) a (31).

- |                       |               |               |
|-----------------------|---------------|---------------|
| 29) [to:'na]~[to:'nã] | /to:na/       | ‘coisa’       |
| 30)[to:'nãwa]         | /to:na-wa/    | ‘coisinha’    |
| 31) [to:'nãwna]       | /to:na-wa-na/ | ‘só coisinha’ |

Podemos observar, à medida que novas palavras vão sendo derivadas, pelo acréscimo de sufixos, que elementos nasais uma vez existentes podem contribuir para a realização nasal da vogal. Sobretudo, esses dados constituem-se como uma evidência da estabilidade da proposta teórica do traço nasal como um traço flutuante, conforme Clements e Hume (1995).

Em 32) temos uma evidência forte desses fatos.

- |             |         |            |
|-------------|---------|------------|
| 32) ['nẽwa] | /'nẽwa/ | ‘e, então’ |
|-------------|---------|------------|

Nos dados analisados, a forma ['nẽwa] ocorre em alternância livre com a forma ['nẽma]. Sobre esse aspecto, Sá (2017, p. 75) faz a seguinte observação: “A realização desse morfema como duas diferentes formas parece apontar para uma evolução histórica em que /m/ → [w] deixando a nasalização sobre a vogal precedente, o que não é esperado antes de [w]”.

Vogais nasais ocorrem, sem que possamos explicar suas origens em diversas palavras de uso amplo, uma parte delas classificadas como partículas, como podemos observar em (33) e (34), a seguir.

- |               |                               |
|---------------|-------------------------------|
| 33) [e'hẽ:há] | ‘em prol, benefício, bondade’ |
| 34) [ʼftã:hã] | ‘sozinho’                     |

Como essas vogais são longas, pode-se propor que sejam resultado do processo já conhecido de alongamento compensatório, sem que, no entanto, do ponto de vista sincrônico, tal processo possa ser identificado. As formas estão já cristalizadas, como também ocorre nos exemplos (35) a (37).



As formas em (40) mostram que um /m/ em uma sílaba seguinte nasaliza a vogal precedente antes de [w], possivelmente, devido à queda da vogal átona, o traço nasal associa-se à coda de uma sílaba precedente, realizando-se aí como /w̃/ e criando um ditongo nasal. A vogal da segunda sílaba é mais fraca do que todas as demais vogais, apontando novamente para um processo em que a vogal depois de [m] em uma sílaba não acentuada reduz-se até apagar totalmente. Esses processos alimentam a passagem de [m] a [w], com a nasalidade já espriada de [m] para a vogal precedente estabilizando-se após a perda da nasal no *onset* seguinte. Um fato a mais a observar-se aqui é que a língua não permite consoantes, com exceção dos glides, em codas finais, o que nos permite explicar a realização de /m/ como o glide foneticamente semelhante.

#### 4. CONCLUSÃO

De modo abrangente, no Yaathe (COSTA, 1999) todas as ocorrências de vogais nasais seriam consideradas fonéticas, causadas por processos de ordem fonológica. Em nossos dados, como pode ser visto, não encontramos pares mínimos que demonstrem o contraste entre vogal oral e nasal, o que evidenciaria, de acordo com os modelos de descrição clássicos, que [+nasal] não é um traço fonológico da vogal. Entretanto, encontramos vogais nasais que, a princípio, não podem ser consideradas como resultado de processos de assimilação, do ponto de vista sincrônico e em uma análise linear.

Sabemos que, do ponto de vista da análise fonológica tradicional, que se baseia em contrastes e complementação para identificar fonemas em uma língua, não pode ficar provado que existam vogais nasais fonológicas nasais em Yaathe, uma vez que pares mínimos, pares análogos e distribuição complementar não foram encontrados. Entretanto, em nossos dados aparecem vogais nasais breves fora de contextos nasais assimilantes como em [khofe'ã] 'para' (Benefactivo). Em alguns casos, pode-se falar de cristalização de formas que foram criadas por processos fonológicos e cujas formas subjacentes não mais são recuperáveis.

A partir das análises que efetuamos, embora possamos sugerir explicações não sincrônicas para algumas ocorrências de vogais nasais, propomos que essas vogais sejam consideradas fonemas, visto que não são resultado de operações fonológicas atuais.

## REFERÊNCIAS

- CABRAL, D. F. *Descrição fonética de pitch e intensidade no nível da palavra em Yaathe (Fulni-ô)*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2007.
- CABRAL, D. F. *O acento lexical em Yaathe*. (Dissertação de Mestrado). Maceió: PPGLL/Universidade Federal de Alagoas, 2009.
- CHELLIAH, Shobhana L., de REUSE, Willem J. *Handbook of Descriptive Linguistic Fieldwork*. Springer. 1th Edition, 2011
- CLEMENTS, G. N.; HUME, E. V. *The internal organization of speech sounds*. In: GOLDSMITH, J. A. *The handbook of Phonological Theory*. London: Blackwell, 1995.
- COSTA, J. F. *Ya:thê, a última língua nativa no nordeste do Brasil: aspectos morfofonológicos e morfossintáticos*. (Tese de Doutorado). Recife: UFPE, 1999.
- DIAS, C. S. *A Função e o Comportamento do Traço Nasal em Yaathe, Língua Indígena Brasileira*. (Dissertação de Mestrado). Maceió: PPGLL/Universidade Federal de Alagoas, 2017.
- DIAS, C. S. *O comportamento do traço nasal em Yaathe, língua indígena brasileira*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2014.
- GOLDSMITH, John. *Autosegmental Phonology*. (Tese de Doutorado). Cambridge, Mass.: MIT Press. 1976.
- MELO, J. A. *Gênero gramatical em Yaathe*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2010.
- NESPOR, M. E VOGEL, I. *Prosodic Phonology*. Dordrecht – Holland: Foris Publications, 1986.
- RODRIGUES, A. D. *Línguas brasileiras: Para o conhecimento das línguas indígenas*. São Paulo: Loyola, 1986.
- RODRIGUES, A. D. *Silêncio, nasalidade e laringalidade em línguas indígenas brasileiras*. Letras de Hoje. v. 38, n. 4, p. 11-24. Porto Alegre. 2003. Disponível em: <http://www.etnolinguistica.org/biblio:rodrigues-2003-silencio>. Acesso em: 03/2016.

SÁ, H. F. *Documentação de narrativas de anciãos Fulni-ô*. (Dissertação de Mestrado). PPGLL/UFAL, 2017.

SILVA, F. *A organização prosódica do Yaathe, a língua do povo Fulni-ô*. (Tese de Doutorado). Maceió: PPGLL/Universidade Federal de Alagoas, 2016.

SILVA, F. *A sílaba em Yaathe*. (Dissertação de Mestrado). Maceió: PPGLL/Universidade Federal de Alagoas, 2011.

SOUSA, M. S. *Marcação fonética do acento em duas classes de palavras da língua indígena brasileira Yaathe: nome e verbo*. (Trabalho de Conclusão de Curso). Maceió: UFAL/FALE, 2014.

SOUSA, M. S. *Análise acústica experimental da duração de vogais em Yaathe*. (Dissertação de Mestrado). Maceió: PPGLL/UFAL, 2017.





Estudos em Fonética e Fonologia é uma coletânea de textos científicos que se situam dentro da grande área da Linguística e se ocupam do componente expressivo das línguas humanas, sons e gestos. Alguns textos apresentam o resumo de uma pesquisa completa, outros, aspectos de uma determinada investigação, outros ainda, uma discussão da metodologia utilizada. Estudiosos do grupo de pesquisa Fonufal – Estudos de Fonética e Fonologia, Ufal, os autores desenvolveram seus trabalhos em campos diversos, como: determinação de diagnósticos; detecção de atitudes; processamento de fala e compreensão; entoação de agrupamentos numéricos; rastreamento ocular; fonoaudiologia; prosódia da Libras; fonética forense; ensino de segunda língua; variação; e língua indígena. Trata-se, pois, de uma obra indispensável para todos que tenham interesse pela área.



[openaccess.blucher.com.br](http://openaccess.blucher.com.br)



**Blucher** Open Access