

SERGIO AUGUSTO MAUAD

**ANÁLISE DAS CARACTERÍSTICAS RÍTMICAS E MELÓDICAS DA
LEITURA DE UM DIÁLOGO EM INGLÊS POR APRENDIZES
BRASILEIROS**

DOUTORADO EM LINGUÍSTICA APLICADA E ESTUDOS DA LINGUAGEM

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

2014

SERGIO AUGUSTO MAUAD

**ANÁLISE DAS CARACTERÍSTICAS RÍTMICAS E MELÓDICAS DA
LEITURA DE UM DIÁLOGO EM INGLÊS POR APRENDIZES
BRASILEIROS**

Tese apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de DOUTOR em Linguística Aplicada e Estudos da Linguagem, sob orientação da Prof^a Dr^a Sandra Madureira.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

2014

MAUAD, Sergio Augusto. Análise das características rítmicas e melódicas da leitura de um diálogo em inglês por aprendizes brasileiros. São Paulo: p.188 2014.

Tese (Doutorado): Pontifícia Universidade Católica de São Paulo.

Área de Concentração: Linguística Aplicada e Estudos da Linguagem.

Orientadora: Prof^a Dr^a Sandra Madureira.

Banca Examinadora

“There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle.”

(Albert Einstein)

DEDICATÓRIA

Augusto

Friozinho da manhã. Cheirinho de café coado. Os pássaros, sinfonia de mil sons. A grama ainda úmida. A piscina pontilhada de incautos insetos.

De calças de sarja, baldinho na mão, lá vem ele com os restos de comida: raspas de prato, ossos de galinha. Seu fiel amigo, Tuco, latindo freneticamente embaixo da mangueira, aguarda.

O nobre senhor cruza o gramado. Aquele quintal é sua casa. O gramado, seu tapete. Altivo na nobre missão de alimentar seu velho amigo.

Setenta e sete anos de vida. Missão ainda não cumprida. De camisa de linho já surrada, meias curtas cinzas enfiadas num *Rider* azul, barba a fazer. Ele segue seu curso.

(Prosa poética – Sérgio Augusto Mauad – 20/02/2002)

Ao meu pai, *in memoriam*

AGRADECIMENTOS

Aos meus pais, Augusto e Lúdia, pelos princípios e valores que nortearam meu caráter.

À minha esposa, Lenita, pelo amor, carinho e compreensão.

Às minhas filhas, Giovanna e Júlia, meu *joie de vivre*, pelo prazer do convívio diário.

Aos meus irmãos, Marcos e Claudio, pelo exemplo de vida e por estarem presentes nos momentos significativos de nossa família.

À minha orientadora, Prof^a Dr^a Sandra Madureira, pela amizade sincera, confiança e pelo genuíno interesse em meu crescimento profissional e pessoal.

À Prof^a Dr^a Zuleica Camargo, pelo estímulo constante e contribuições.

À Prof^a Yara de Castro, pela assessoria na análise estatística dos dados.

À Prof^a Dr^a Maria Aparecida Caltabiano, cuja simplicidade na resolução dos problemas e no “levar a vida” me servem de inspiração.

Aos Profs Drs Alexandro Meireles e Luciana Lucente, pelas orientações construtivas durante a qualificação.

Aos professores doutores da Banca Examinadora, por procederem à análise deste trabalho de pesquisa.

Aos professores doutores do LAEL, pelas contribuições que deram à minha formação.

Às pesquisadoras Paulina Rocca, Camila Dixo Lieff e Zaina Abdalla Nunes (*in memoriam*), pelos trabalhos.

Aos colegas do LIAAC – Aline Pessoa, Amaury Silva, Andrea Sacco, Cristina Canhetti, Fabiana Gregio, Isaías Santos, Jae Kohson, Lilian Cristina Kuhn Pereira, Layla Penha, Luciana Oliveira, Marcelo Henrique Barbosa, Maria Augusta Svicero, Marilea Fontana, Mario Augusto de Souza Fontes, Paulo Menegon, Roberta Isolan Cury e Vanessa Medina – por compartilharem seus trabalhos nos seminários de pesquisa e pelo *feedback* construtivo.

Aos voluntários, brasileiros e norte-americanos, que, generosamente, se deslocaram até os estúdios de rádio e TV da PUC-SP para a leitura e gravação do *corpus* de pesquisa.

À Adriana Pucci, por contatar os voluntários norte-americanos em todas as etapas da gravação do *corpus* da pesquisa.

A Ernesto Luis Foschi, Alípio Foschi e Ronaldo Barbosa da Silva, técnicos de áudio dos estúdios de gravação da PUC-SP, pelo cuidado durante as gravações e pela forma amistosa com que receberam os sujeitos da pesquisa.

Ao Prof Ms Mario Madureira Fontes, pelo incentivo e auxílio na elaboração de meu *website*.

À Fátima de Albuquerque, secretária do LIAAC, por me manter a par dos eventos acadêmicos e pela prontidão na disponibilização das teses e dissertações.

À Maria Lucia dos Reis, secretária do LAEL, que, desde os tempos de mestrado, me auxilia quanto aos procedimentos acadêmicos.

À Marcia Martins, secretária do CEPRIL, por me orientar sobre a submissão de artigo para a revista *The ESpecialist* e por proporcionar a todos um ambiente musicalmente agradável no LAEL.

Aos tutores do curso Inglês Online da EVESP, professores da Rede Pública de Ensino do Estado de São Paulo, pelo respeito e seriedade com que exerceram suas funções.

Aos colegas coordenadores de tutoria do curso Inglês Online da EVESP, pelo engajamento no projeto e pelo espírito de *teamwork*.

À CAPES, pelo apoio financeiro.

A todos meus alunos, de ontem, hoje e amanhã, que, em algum momento, me permitiram fazer parte de suas vidas.

RESUMO

A prosódia, pela interação que estabelece com os segmentos fônicos, é um dos elementos a ser considerado na investigação do sotaque estrangeiro. Neste estudo, tenho como objetivo investigar, com apoio na análise fonético-acústica, dois elementos prosódicos: o ritmo e a entoação, contrastando as produções de fala de duas brasileiras e de duas norte-americanas na leitura de um diálogo em inglês. O ritmo, que implica em regularidade e estruturação ao longo do tempo, exerce papel relevante no discurso oral. A entoação, a melodia da fala, organiza o fluxo sonoro e expressa modalidades, emoções e atitudes. No caso de falantes aprendizes de uma língua estrangeira, são grandes as possibilidades de transferência de padrões prosódicos da língua materna. Neste trabalho, os padrões rítmicos e entoacionais das sentenças de um diálogo são contrastados com o apoio da análise fonético-acústica, utilizando-se o *software* PRAAT. A extração de medidas contempla a duração, em milissegundos, das unidades V-V, sílabas, pés rítmicos e vogais plenas em adjacência, bem como marcações de *f0* em alvos específicos do contorno entoacional: os *pitch accents*. A fundamentação teórica para o ritmo do inglês norte-americano baseia-se na Teoria de *Pitch Accent*, proposta por Bolinger. Para a notação dos padrões entoacionais, é utilizado o sistema ToBI (*Tones and Break Indices*) e, para o tratamento estatístico dos dados, o teste ANOVA e o índice de variabilidade. Foram constatadas diferenças significativas ($p < ,05$) entre as produções de fala dos sujeitos norte-americanos e brasileiros em relação à duração média relativa de sílabas e unidades V-V e diferenças em relação ao índice de variabilidade dos pés silábicos. Também foram constatadas nas produções das brasileiras interferências de L1 em relação à tonicidade, junturas, qualidade vocálica e processos de redução e elisão segmental que afetaram a estrutura rítmica dos enunciados. Em relação à entoação, observou-se, nas emissões das norte-americanas, a ocorrência, de forma mais sistemática, do fenômeno de *downstep*, e, nas produções das brasileiras, a ocorrência produtiva dos tons L^*+H e $L+H^*$, o que afeta a produção de sentidos. Com o desenvolvimento deste trabalho, pretendo contribuir para aprofundar os conhecimentos sobre o ritmo e a entoação no contexto de ensino de inglês para brasileiros.

Palavras-chave: prosódia, ritmo, inglês, entoação, ensino de línguas, Fonética

ABSTRACT

Due to its interaction with the sound segments, prosody is one of the elements to be considered when investigating foreign accent. In this study, the speech patterns of a dialogue read by two Brazilians and two Americans are contrasted, and two prosodic elements are investigated: rhythm and intonation. Speech rhythm, which refers to a temporal patterning of elements in the flow of speech, plays a relevant role in discourse. Intonation, the melody of speech, organizes the flow of speech and expresses modalities, emotions and attitudes. In the case of L2 speakers, it is very likely that certain prosodic patterns of L1 will be transferred to L2. In this study, the rhythmic and intonational patterns of a read dialogue are contrasted through acoustic-phonetic analysis by means of the PRAAT software. The acoustic data comprises duration measurements of V-V units, syllables, feet and adjacent full vowels, as well as pitch accent markings. The analysis of the rhythm of American English is mainly based on Bolinger's Pitch Accent Theory. ToBI (*Tones and break indices*) is used for the tonal labelling, and ANOVA and the variability index for the statistical treatment of the data. Results point to significant differences ($p < .05$) in syllable and V-V unit duration means, and differences in the variability index of the syllabic feet. The rhythmic structure of the utterances produced by the Brazilian subjects was affected by L1 transfer of tonicity, juncture, vowel quality as well as processes of segmental reduction and elision. As regards intonation, it was observed, in the productions of the American subjects, a systematic occurrence of the phenomenon of downstep, and, in the productions of the Brazilian subjects, a productive occurrence of L*+H e L+H* tones, which affected the production of meaning. With this study, I wish to contribute to the research on the teaching of rhythm and intonation to Brazilian speakers of English.

Keywords: prosody, rhythm, English, intonation, language teaching, Phonetics

SÍMBOLOS FONÉTICOS

Os símbolos fonéticos, utilizados para a transcrição dos enunciados deste trabalho foram extraídos do IPA, *International Phonetic Alphabet*, e encontram-se listados abaixo, seguidos de um exemplo de palavra na qual ocorrem.

Consoantes:

[p]	<u>p</u> air	[f]	<u>f</u> eel	[j]	<u>y</u> ou
[b]	<u>B</u> ill	[v]	<u>v</u> ictory	[h]	<u>h</u> ouse
[t]	<u>t</u> ea	[θ]	<u>th</u> ink	[m]	<u>m</u> atrix
[d]	<u>d</u> octor	[ð]	<u>th</u> e	[n]	<u>n</u> ose
[k]	<u>c</u> at	[s]	<u>s</u> oup	[ŋ]	<u>ng</u>
[g]	<u>g</u> ate	[z]	<u>z</u> oo	[l]	<u>l</u> ook
[tʃ]	<u>ch</u> ip	[ʃ]	<u>sh</u> ip	[r]	<u>r</u> ose
[dʒ]	<u>g</u> eorge	[ʒ]	be <u>ig</u> e	[w]	<u>w</u> ash

Vogais:

[ɪ]	f <u>i</u> ll	[i]	fe <u>e</u> l	[aʊ]	hou <u>se</u>
[ɛ]	<u>a</u> ny	[eɪ]	fa <u>ce</u>	[ɔ]	hor <u>se</u>
[æ]	m <u>a</u> n	[aɪ]	fi <u>r</u> e	[ə]	<u>a</u> bout
[ɑ]	ca <u>r</u>	[ɔɪ]	bo <u>y</u>	[ə]	con <u>cern</u>
[ʌ]	cu <u>t</u>	[u]	fo <u>o</u> l	[i]	ros <u>e</u> s
[ʊ]	bo <u>o</u> k	[oʊ]	bo <u>a</u> t	[ə]	willow <u>o</u>

LISTA DE FIGURAS

Figura 1 – Exercício para a prática do ritmo. Fonte: Underhill (1994, p. 71)	38
Figura 2 – Exercício para percepção de proeminências. Fonte: Lujan (2004)	40
Figura 3 – Exercício para a percepção de <i>content</i> e <i>function words</i> . Fonte: Grant (2004)	41
Figura 4 – Exercício para a percepção das proeminências acentuais a partir da visualização da curva entoacional. Fonte: Mauad <i>et al</i> (2011).....	42
Figura 5 – Forma da onda e espectrograma de banda larga da emissão <i>Let me see</i> , produzida por um falante norte-americano. Fonte: Silva (2009, p. 79)	49
Figura 6 – Forma da onda e espectrograma de banda larga da emissão <i>Let me see</i> , produzida por um falante brasileiro. Fonte: Silva (2009, p. 81)	49
Figura 7 – Forma da onda, curva entoacional e segmentação ortográfica do enunciado <i>Posso deixar minha bagagem na recepção?</i> , produzido por uma falante brasileira. Fonte: Mauad (2010b).....	52
Figura 8 – Forma da onda, curva entoacional e segmentação ortográfica do enunciado <i>Can I leave my bag at the front desk?</i> , produzido por um falante americano. Fonte: Mauad (2010b)	53
Figura 9 – Representação gráfica de stress-timing e syllable-timing. Fonte: http://www.phonetiklabor.de	58
Figura 10 – Quadro das vogais do inglês	71
Figura 11 – Anúncio do sabonete Boraxo	79
Figura 12 – Forma da onda, contorno entoacional, pés silábicos, número de sílabas por pé, duração do pé silábico, porcentagem da duração do pé em relação ao todo e <i>pitch accents</i> do enunciado <i>We're just dropping by to let you know we're having a Halloween party</i>	88
Figura 13 – Forma da onda, curva entoacional, espectrograma de banda larga e segmentação com transcrição ortográfica da palavra <i>daughter</i> no enunciado <i>My daughter is in Switzerland</i> , ilustrando episódios de <i>breathiness</i> como marcador de proeminência acentual	101
Figura 14 – Duas fotos com intervalo de 1 segundo, ilustrando a qualidade de voz de <i>breathiness</i> que acompanha a atitude de surpresa. Fonte: série <i>Desperate Housewives</i>	102

Figura 15 – Frase composta de duas palavras exemplificando duas análises para a regra rítmica: a <i>reversal analysis</i> (RA) e a <i>deletion analysis</i> (RR). Fonte: Vogel, Bunnell e Hoskins (1995)	107
Figura 16 – Forma da onda, espectrograma, curva entoacional, pés rítmicos acentuais, duração do pé, número de sílabas por pé e <i>pitch accents</i> do enunciado <i>Watching the videos is fun</i> , produzido por Norah Jones. Fonte: revista <i>Speak-up</i>	109
Figura 17 – Forma da onda, curva entoacional, pés rítmicos acentuais, número de sílabas por pé, duração do pé e <i>pitch accents</i> do enunciado <i>Where did you get the funny suit?</i> , ilustrando o “encaixe” entre o ritmo silábico e o ritmo acentual. Fonte de áudio: filme <i>The Purple Rose of Cairo</i>	113
Figura 18 – Forma da onda, curva entoacional, pés rítmicos silábicos, número de sílabas por pé rítmico silábico, pé rítmico acentual, número de sílabas por pé rítmico acentual e <i>pitch accent</i> do enunciado <i>I’m up for new experiences</i> , ilustrando a independência entre o ritmo silábico e o ritmo acentual. Fonte de áudio: filme <i>The Purple Rose of Cairo</i>	114
Figura 19 – Frequência fundamental de 100 Hz e subsequentes harmônicos múltiplos.....	119
Figura 20 – A estrutura de um grupo tonal em inglês. Fonte: Underhill (1994).....	128
Figura 21 – Contorno de f_0 , palavras e notação entoacional de três configurações de <i>pitch</i> descritas por Bolinger (1986) (<i>profile AC</i> , <i>profile C</i> e <i>profile A</i>) do enunciado <i>Anna married Lenny</i> . Fonte de áudio: Silverman <i>et al</i> (1992).....	130
Figura 22 – Contorno de f_0 , segmentação em 2 grupos tonais e notação entoacional, segundo Brazil (1995), do enunciado <i>And usually for lunch, I’ll have a salad</i> , extraído da série <i>Seinfeld</i>	131
Figura 23 – Contorno de f_0 , grupos tonais, segmentação em 2 grupos tonais e notação entoacional segundo Pierrehumbert do enunciado <i>Anna married Lenny</i> . Fonte: Silverman <i>et al</i> (1992)	132
Figura 24 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI: (1) ortográfica, (2) de tons, (3) grau de junção e (4) ocorrências para o primeiro enunciado, do arquivo <i>Anna</i> . Fonte: Silverman <i>et al</i> (1992).....	136
Figura 25 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI: (1) ortográfica, (2) de tons, (3) de grau de junção e (4) das ocorrências, para o segundo enunciado do arquivo <i>Anna</i> . Fonte: Silverman <i>et al</i> (1992)	137
Figura 26 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI (1) ortográfica, (2) de tons, (3) grau de junção e (4) ocorrências para o terceiro enunciado do arquivo <i>Anna</i> . Fonte: Silverman <i>et al</i> (1992).....	137

Figura 27 – Forma da onda, contorno de <i>f0</i> e transcrição de tons da primeira repetição de S4 do enunciado: <i>Let's get Cris to organize the activities</i>	145
Figura 28 – Forma da onda, espectrograma de banda larga e segmentação da palavra <i>above</i>	146
Figura 29 – Forma da onda, espectrograma de banda larga e camadas de segmentação das sílabas da palavra <i>difference</i> com transcrição ortográfica, na camada superior, e fonética, na camada inferior. Fonte: Silverman (1992).....	148
Figura 30 – Linhas verticais no espectrograma, representando os pulsos glotais.....	148
Figura 31 – Medidas de duração e marcação dos eventos tonais do primeiro enunciado produzido pelo sujeito brasileiro S1: <i>tier 1</i> : sílabas; <i>tier 2</i> : pés silábicos; <i>tier 3</i> : pés acentuais; <i>tier 4</i> : vogais plenas em adjacência; <i>tier 5</i> : unidade V-V; <i>tier 6</i> : eventos tonais.....	149
Figura 32 – “Encaixe” entre o ritmo silábico e o ritmo acentual na primeira repetição da emissão <i>No, not Sandra</i> , produzida pelo sujeito brasileiro S1.....	158

LISTA DE TABELAS

Tabela 1 – Porcentagens de tipos silábicos em inglês e em espanhol. Fonte: Dauer (1983, p. 56)	66
Tabela 2 – Vogais plenas do inglês. Fonte: Bolinger (1986)	74
Tabela 3 – Vogais reduzidas do inglês. Fonte: Bolinger (1986)	75
Tabela 4 – A relação <i>stress</i> e <i>accent</i> na marcação de proeminências em inglês. Fonte: Bolinger (1964, p. 22)	98
Tabela 5 – Enunciado composto de 7 pés silábicos, exemplificando o ritmo silábico do inglês. Fonte: Bolinger (1986).....	112
Tabela 6 – Enunciado composto de 5 pés acentuais, exemplificando o ritmo acentual do inglês. Fonte: Bolinger (1986)	112
Tabela 7 – Rótulos dos 5 tons para o acento de <i>pitch</i> para o inglês, segundo o sistema de notação entoacional ToBI.....	133
Tabela 8 – Rótulos para os acentos frasais para o inglês, segundo o sistema de notação entoacional ToBI	134
Tabela 9 – Rótulos para os três tons de fronteira, segundo o sistema de notação entoacional ToBI.....	134
Tabela 10 – Rótulos para os quatro tons de fronteira final, segundo o sistema de notação entoacional ToBI	135
Tabela 11 – Quatro camadas, ou <i>tiers</i> , na transcrição dos eventos tonais, segundo o sistema de notação entoacional ToBI.....	135
Tabela 12 – Índice the variabilidade dos pés silábicos.....	160
Tabela 13 – Número de pés acentuais nas emissões dos 4 falantes.....	162
Tabela 14 – Vogais plenas em adjacência nos seis enunciados, sujeitas à regra de <i>extra lengthening</i> (BOLINGER, 1986).....	163
Tabela 15 – Eventos tonais nas 3 repetições do enunciado 1, produzidas pelos sujeitos brasileiros (S1 e S2)	164
Tabela 16 – Eventos tonais nas 3 repetições do enunciado 1, produzidas pelos sujeitos norte-americanos (S3 e S4).....	165
Tabela 17 – Eventos tonais nas 3 repetições do enunciado 2, produzidas pelos sujeitos brasileiros (S1 e S2)	166

Tabela 18 – Eventos tonais nas 3 repetições do enunciado 2, produzidas pelos sujeitos norte-americanos (S3 e S4).....	166
Tabela 19 – Eventos tonais nas 3 repetições do enunciado 3, produzidas pelos sujeitos brasileiros (S1 e S2)	167
Tabela 20 – Eventos tonais nas 3 repetições do enunciado 3, produzidas pelos sujeitos norte-americanos (S3 e S4).....	168
Tabela 21 – Eventos tonais nas 3 repetições do enunciado 4, produzidas pelos sujeitos brasileiros (S1 e S2)	168
Tabela 22 – Eventos tonais nas 3 repetições do enunciado 4, produzidas pelos sujeitos norte-americanos (S3 e S4).....	169
Tabela 23 – Eventos tonais nas 3 repetições do enunciado 5, produzidas pelos sujeitos brasileiros (S1 e S2)	170
Tabela 24 – Eventos tonais nas 3 repetições do enunciado 5, produzidas pelos sujeitos norte-americanos (S3 e S4).....	170
Tabela 25 – Eventos tonais nas 3 repetições do enunciado 6, produzidas pelos sujeitos brasileiros (S1 e S2)	171
Tabela 26 – Eventos tonais nas 3 repetições do enunciado 6, produzidas pelos sujeitos norte-americanos (S3 e S4).....	172

LISTA DE GRÁFICOS

Gráfico 1 – Isocronismo acentual ideal. Fonte: Barbosa, 2006	63
Gráfico 2 – Isocronismo silábico ideal. Fonte: Barbosa, 2006.....	64
Gráfico 3 – Ausência de isocronismo: pé métrico com duração variável. Fonte: Barbosa, 2006.....	65
Gráfico 4 – Duração do pé rítmico vs número de sílabas no pé. Fonte: Kim e Cole (2007).....	84
Gráfico 5 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 2.....	154
Gráfico 6 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 3.....	155
Gráfico 7 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 4.....	156
Gráfico 8 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 5.....	157

SUMÁRIO

INTRODUÇÃO	23
Objetivo	26
Organização geral do estudo	28
Capítulo 1 QUESTÕES SOBRE PRONÚNCIA NO CONTEXTO DE INGLÊS COMO L2	30
1.1 Sotaque estrangeiro	31
1.2 Questões de sotaque e o modelo de aprendizagem de fala de uma segunda língua (Flege)	31
1.3 Variáveis no grau de sotaque estrangeiro em inglês	32
1.4 Inteligibilidade vs compreensibilidade	33
1.5 influência da instrução formal no grau de sotaque estrangeiro	35
1.6 hipótese do período crítico para o aprendizado de uma língua	37
1.7 ensino-aprendizagem de ritmo do inglês e os materiais didáticos para ensino de língua inglesa	38
1.8 Transferências da Língua materna (L1)	43
1.8.1 Transferências relacionadas à tonicidade.....	43
1.8.2 Transferências relacionadas às vogais reduzidas	45
1.8.3 Transferências da L1 relacionadas à fala coarticulada.....	48
1.8.4 Transferências da L1 relacionadas à fonotaxe	50
1.8.5 Transferências relacionadas à entoação	51
Capítulo 2 QUESTÕES SOBRE O RITMO DA FALA	55
2.1 Conceituação do ritmo	56
2.2 Estudos sobre o ritmo das Línguas	56
2.3 Ritmo do inglês norte americano.....	70
2.3.1 Vogais do inglês	71
2.3.1.1 Vogais plenas e vogais reduzidas.....	73
2.3.2 Ritmo silábico (<i>full vowel timing</i>)	77
2.3.2.1 Alongamento silábico	78
2.3.2.2 <i>The Borrowing Rule</i>	82
2.3.2.3 Pés rítmicos silábicos.....	87
2.3.3 Ritmo acentual (<i>pitch accent timing</i>)	91
2.3.3.1 <i>Stress</i>	92
2.3.3.2 <i>Pitch accent</i>	96
2.3.3.3 <i>Content words vs function words</i>	99
2.3.3.4 Outras pistas para o <i>accent</i>	100
2.3.3.4.1 <i>Breathiness</i>	100
2.3.3.4.2 Intensidade.....	103
2.3.3.5 Regra rítmica e colisão acentual (<i>stress clash</i>).....	104
2.3.3.6 Pés ritmos acentuais.....	108
2.3.4 Interação entre os ritmos silábico e acentual no inglês	111

Capítulo 3 QUESTÕES SOBRE A ENTOAÇÃO DA FALA	117
3.1 Conceituação de entoação.....	118
3.2 Funções da entoação.....	120
3.2.1 Função gramatical.....	121
3.2.2 Função discursiva.....	121
3.2.3 Função atitudinal	123
3.2.4 Função sociolinguística	124
3.2.5 Função acentual.....	125
3.2.5.1 Foco.....	126
3.3 Modelos de descrição da entoação.....	128
3.3.1 Escola Britânica.....	128
3.3.2 Teoria de <i>Pitch Accent</i> , de Bolinger.....	129
3.3.3 <i>Discourse Intonation</i>	130
3.3.4 Modelos gerativos	131
3.3.4.1 Sistema de notação entoacional TOBI	132
3.3.5 Modelo dinâmico-entoacional.....	138
Capítulo 4 METODOLOGIA DE PESQUISA	140
4.1 Escolha da metodologia de pesquisa.....	141
4.2 Sujeitos de pesquisa	142
4.3 <i>Corpus</i> de pesquisa	142
4.4 Gravação do <i>corpus</i>	143
4.5 Notação dos padrões entoacionais	144
4.6 Análise acústica	145
4.6.1 Instrumentos, procedimentos e técnicas de análise.....	145
4.6.2 Procedimentos de segmentação e medição das unidades linguísticas	149
4.7 Análise estatística	151
Capítulo 5 RESULTADOS, ANÁLISE E INTERPRETAÇÃO DOS DADOS	152
5.1 Análise dos padrões rítmicos: resultados e discussão.....	153
5.1.1 Unidades V-V	153
5.1.2 Sílabas.....	158
5.1.3 Pés silábicos.....	159
5.1.4 Pés acentuais.....	162
5.1.5 Vogais plenas em adjacência.....	162
5.2 Análise dos padrões entoacionais: resultados e discussão	164
CONSIDERAÇÕES FINAIS	174
REFERÊNCIAS BIBLIOGRÁFICAS	178

INTRODUÇÃO

Um dos elementos importantes para a comunicação oral é a prosódia, devido a seu papel pragmático na expressão de modalidades, atitudes e emoções.

Segundo Nootboom (1997), a palavra *prosódia* tem sua origem no grego antigo e significava uma “canção cantada com música instrumental”. No passado, o termo prosódia referia-se à “ciência da versificação” e ocupava-se do estudo da modulação da voz humana na leitura de poesia em voz alta.

Na fonética moderna, a palavra *prosódia* e seu adjetivo *prosódico* são mais comumente utilizados para referir-se às propriedades da fala que perpassam ou interagem, de forma relativamente livre e independente, com os segmentos fônicos.

Conforme afirma Kent (1992, p. 152):

Prosody might be regarded as the fabric of speech, within which segments are the individual stitches or fibers. Prosodic patterns span the linguistic levels, holding together the many influences that make up the rich tapestry of language in context. Prosody serves essential, if sometimes subtle, functions in communication, and its acoustic foundations are no less important in the speech signal than those that distinguish segments.

A prosódia, ao englobar os fenômenos que interagem com segmentos vocálicos e consonantais, tem papel fundamental no processamento da fala. Entre os elementos prosódicos, podemos citar a entoação, o acento, a taxa de elocução, a qualidade de voz e o ritmo. Todos esses elementos têm relevância na medida em que organizam o fluxo da fala de forma que ela possa ser apreendida pelo falante e seu ouvinte.

No ensino de pronúncia da língua inglesa, esse aspecto funcional da prosódia é geralmente negligenciado. Os exercícios de prosódia destinados a aprendizes

brasileiros geralmente têm forte viés impressionístico e muitas vezes são escritos por estrangeiros que, não tendo conhecimento do português, não conseguem dar conta das dificuldades dos aprendizes brasileiros em produzir padrões fonéticos do inglês. Por exemplo, fenômenos de redução silábica, imprescindíveis ao ritmo, são muitas vezes aprendidos no nível da palavra e, portanto, fora de seu contexto discursivo. No tocante à entoação propriamente dita, tende-se a privilegiar as relações entre a entoação e a expressão de modalidades, negligenciando a expressão de atitudes e emoções e seu papel na estrutura rítmica dos enunciados.

Por conta do desenvolvimento de cursos de pronúncia inglesa no programa de extensão da PUC-SP e de meus trabalhos acadêmicos, passei a voltar minha atenção ao modo como a prosódia interage com os segmentos. Descartei a ideia de que a prosódia se sobrepõe aos segmentos ao perceber que, na realidade, ela os afeta. Chamou-me atenção, por exemplo, o fato de que a pronúncia de um mesmo segmento possa variar em função de sua posição na frase entoacional. Mas, acima de tudo, pude perceber como são vários os pontos de intersecção entre o ritmo e a melodia na produção de fala encadeada. Conforme lembra Agüena (2006, p. 5):

A fala é composta por uma organização de sons sob uma melodia (entoação ou curva entoacional – variação entre tons graves e agudos) e um ritmo (alternância entre acentos fracos ou fortes), que podem ser modificados pela taxa de elocução (fala rápida ou lenta) adotada pelo falante, a qual, por sua vez, condiciona o grau de sobreposição entre os sons (coarticulação), fazendo com que eles alterem-se, o que provoca, então, ajustes melódicos e rítmicos. Isso demonstra que não é possível lidar com a produção de sons isolados sem considerar sua interação com a entoação, o ritmo, a taxa de elocução, entre outros. A isso se dá o nome de fala “corrente ou encadeada”.

Acredito que a dificuldade maior do aprendiz não esteja na produção e percepção dos sons em palavras isoladas, mas, sim, na produção e percepção dos sons em contextos mais amplos. Meus alunos, por exemplo, não têm muita dificuldade em produzir a variante *flape* do /t/ em uma mesma palavra (como em *later*), mas têm muita dificuldade em implementá-la quando ela ocorre *entre* palavras (como em *I go to work*).

OBJETIVO

Este trabalho tem como objetivo analisar, com apoio na análise fonético-acústica, as características rítmicas e melódicas das produções de um diálogo em língua inglesa por duas falantes brasileiras e por duas norte-americanas.

O recurso à Fonética Acústica é relevante, pois, como aponta Madureira (2006), “a fonética acústica serve como instrumento para se ouvir, inferir os gestos envolvidos na produção dos sons, integrar o produzido ao percebido e lidar com a variação que caracteriza a fala: o detalhe fonético”.

Madureira (2006) afirma serem inúmeras as vantagens de se usar os recursos dos instrumentais de análise acústica. Primeiramente, os dados acústicos dão informações que seriam impossíveis de se perceber apenas auditivamente, ensejando, assim, o conhecimento. Além disso, diferentes estratégias para a produção de um mesmo som podem ser avaliadas, ampliando as formas de abordagem didática e avaliação. Finalmente, a visualização pode ser utilizada como

guia para aprendizes como forma de estabelecer vínculo entre a produção e a percepção dos sons¹.

Na fala encadeada, não há limites precisos entre os segmentos fônicos. Na verdade, os segmentos são coarticulados. O fluxo de fala é composto de elementos gradientes e não discretos.

A fonética acústica permite a visualização de vários ajustes impostos por componentes que atuam no fluxo de fala, tais como o ritmo e a taxa de elocução. Detalhes fonéticos relevantes relacionados às consoantes plosivas, tais como a aspiração (como em *take*) e o *flapping* (como em *later*) ou mesmo fenômenos de *sandhi*, como a produção de *want to* como *wanna*, podem ser visualizados na forma da onda e no espectrograma.

Neste trabalho, a análise acústica é feita por meio do programa PRAAT. São extraídas medidas de duração das sílabas, das vogais em adjacência, dos pés rítmicos e das unidades V-V ou GIPC (*Group Inter-perceptual Center*)².

Em relação à entoação, o evento tonal destacado para análise é o acento de *pitch*. A transcrição entoacional é realizada por meio do sistema de notação ToBI.

Embora a fundamentação teórica para este estudo do ritmo do inglês norte-americano privilegie a análise fonético-descritiva para o ritmo proposta por

¹ Pesquisa elaborada por Rocca (2003) demonstra que aprendizes brasileiros de inglês que têm acesso ao sistema computacional para o aprendizado da entoação apresentam bons resultados, porque o *feedback* visual é imediato e permite autoavaliação de seus desempenhos, além de orientá-los para novas tentativas de autocorreção.

² O GIPC é uma unidade fonético-acústica que compreende o espaço delimitado pela vogal de uma sílaba (rima) e a vogal da próxima sílaba, incluindo as consoantes (entre as sílabas), mas excluindo a vogal da segunda sílaba. Sua importância nos estudos do ritmo das línguas decorre do fato de que o *onset* vocálico é considerado ponto de saliência perceptual (teoria do *p-center*), pois a transição C-V exerce um papel relevante no processamento periférico auditivo, uma vez que há predominância de neurônios primários respondendo a eventos do tipo *onset* (BARBOSA, 2000).

Bolinger (1958, 1964, 1978, 1981, 1986, 1998), ao longo de minha exposição, serão também feitas referências a estudos relacionados aos aspectos rítmicos e entoacionais, tais como os trabalhos experimentais de Roach (1982), Dauer (1983, 1987, 1993), Madureira (2002, 2004, 2006) e Barbosa (2000, 2006). Em relação ao sotaque estrangeiro, a referência principal será Flege (1986, 1995).

Este trabalho insere-se em um conjunto de outros realizados no Laboratório Integrado de Fonética Acústica e Cognição (LIAAC/PUC-SP) relacionados à aquisição de um conhecimento fonético-acústico sobre os processos de compreensão e produção do inglês por falantes brasileiros³.

ORGANIZAÇÃO GERAL DO ESTUDO

Esta tese compreende cinco capítulos seguidos de considerações finais. No capítulo um, considero resultados de pesquisas sobre diversas variáveis que influenciam o grau de sotaque estrangeiro. A seguir, abordo aspectos envolvidos no processo de ensino-aprendizagem de prosódia do inglês como L2, particularmente em relação ao ritmo, e que se refletem em atividades propostas em livros de pronúncia do inglês como L2. Por fim, focalizo processos de fala de aprendizes brasileiros de inglês com base em resultados de pesquisa e em vivências de ensino de inglês para brasileiros.

O capítulo dois compreende vários aspectos relacionados ao ritmo das línguas, focalizando a teoria para o ritmo do inglês norte-americano proposta por Bolinger (1986).

3 Rocca (2003), Polaczek (2003), Castilho (2004), Agüena (2006), Mauad (2007), Cristino (2007), Silva (2009), Andrade (2009), Peleias (2009), Barbosa M. (2013).

No capítulo três, concentro-me na entoação das línguas. Apresento as funções da entoação e os modelos de descrição da entoação, bem como o sistema de notação entoacional ToBI, que será utilizado na análise dos dados.

No capítulo quatro, apresento a metodologia adotada, incluindo os sujeitos de pesquisa, a construção do *corpus* e os procedimentos de análise acústica.

O capítulo cinco apresenta os resultados dos dados coletados, sua discussão e interpretação e, na sequência, as considerações finais focalizarão a síntese dos resultados apontados pelo estudo e sua relevância para o ensino de pronúncia da língua inglesa para brasileiros. Por fim, apresento as referências bibliográficas.

CAPÍTULO 1 QUESTÕES SOBRE PRONÚNCIA
NO CONTEXTO DE INGLÊS COMO L2

1.1 SOTAQUE ESTRANGEIRO

Segundo Laver (2002), o sotaque estrangeiro corresponde a uma maneira de pronunciar que reflete diretamente as influências dos padrões fonéticos das línguas materna (L1) e estrangeira (L2). Ele é percebido quando há influência forte dos padrões da L1 na L2.

A questão dos fatores desencadeantes do sotaque estrangeiro tem sido, ao longo de minha experiência como professor (e aprendiz) de inglês, um assunto que muito me intriga. Embora meus alunos consigam perceber uma fala com sotaque estrangeiro, eles geralmente têm dificuldade em identificar os fatores que levaram a esse sotaque.

Acredito que uma maior aproximação aos sons do inglês, tanto nos aspectos segmentais como prosódicos, e a conseqüente redução do sotaque estrangeiro, surge justamente quando o aprendiz consegue identificar detalhes fonéticos do inglês e incorporá-los em sua fala. Caso ele não crie essas metas perceptuais, ocorrerão falhas na produção. É sobre essas propriedades das realizações fonéticas em contextos específicos que discorrerei a seguir, no modelo de aprendizagem de fala proposto por Flege.

1.2 QUESTÕES DE SOTAQUE E O MODELO DE APRENDIZAGEM DE FALA DE UMA SEGUNDA LÍNGUA (FLEGE)

O *Second-language Speech Learning Model* (Modelo de Aprendizagem de Fala de uma Segunda Língua) desenvolvido por Flege (1995), defende a ideia de que, embora os mecanismos de produção de fala continuem ativos ao longo de toda a vida, fatores como a ação do sistema fonético de L1 sobre L2 (e vice-versa), bem como a idade na qual o falante de L2 teve seu primeiro contato com L2, contribuem

para que haja diferenças entre as produções de falantes nativos e não nativos aprendizes de L2.

Segue-se que a produção imprecisa de vogais e consoantes por parte do aprendiz de L2 decorre do declínio na sua capacidade de reconhecer certas características fonéticas relevantes para diferenciar sons da L1 dos da L2. Quando o aprendiz possui essa capacidade, ele fica mais propenso a formar novas categorias de sons. Desse modo, ele consegue estabelecer novas categorias fonéticas para esses sons e isso tende a se refletir na sua produção também, contribuindo para minimizar o grau de sotaque estrangeiro.

Flege (1995) afirma que os sons mais difíceis de serem incorporados pelos aprendizes são justamente aqueles mais próximos ou similares aos sons de L1 e que, quando o aprendiz ouve um som novo, ele automaticamente forma uma nova categoria fonética. Por outro lado, se o som for muito similar, ele o coloca em uma categoria fonética de L1. Isso é o que acontece com as vogais /æ/ e /ɛ/, frequentemente percebidas por brasileiros como equivalentes, portanto não diferenciadas em pares mínimos como *bad/bed* e *dad/dead*.

1.3 VARIÁVEIS NO GRAU DE SOTAQUE ESTRANGEIRO EM INGLÊS

Flege e Liu (2001) examinaram a influência de algumas variáveis no grau de sotaque estrangeiro em inglês de falantes italianos, bilíngues de italiano e do inglês, residentes no Canadá. Essas variáveis foram: (1) idade em que o sujeito começou a

falar inglês; (2) tempo de residência no Canadá; (3) sexo do falante; (4) habilidade para o aprendizado de L1; e (5) frequência de uso contínuo de L1 (italiano).

Os autores constataram que as variáveis *idades em que o falante começou a falar inglês e frequência de uso contínuo de L1* afetam o grau de sotaque estrangeiro. Variáveis como *sexo do falante, tempo de residência no Canadá* e autoavaliação sobre *habilidade para o aprendizado de L1* não foram significativas.

1.4 INTELIGIBILIDADE VS COMPREENSIBILIDADE

É necessário que se diferencie os três termos presentes na literatura: sotaque, inteligibilidade e compreensibilidade (*comprehensibility*). Todos eles contribuem para a percepção de proficiência oral, representando conceitos relacionados, porém distintos.

Segundo McGregor (2007), a percepção do *sotaque* estrangeiro ocorre quando o ouvinte nativo encontra aspectos na fala em L2 que destoam das normas de sua fala.

Inteligibilidade refere-se ao processamento acústico da fala e envolve, portanto, o reconhecimento de palavras. De modo mais simplista, uma fala inteligível é aquela cujas palavras podem ser reproduzidas por escrito. O termo *compreensibilidade*, por sua vez, está relacionado à capacidade de o ouvinte entender o significado de um enunciado em seu contexto. Ser compreensível significa, portanto, poder ser compreendido.

Derwing, Rossiter, Munro e Thomson (2004) constataram que, em testes perceptivos de fluência, juízes consideraram como mais fluentes os falantes de L2 cuja fala foi mais compreensível, sendo que a questão do sotaque não se mostrou tão relevante.

Vale ressaltar que o contexto frasal auxilia o ouvinte a identificar as palavras. Peterson e Barney (1952), por exemplo, verificaram que mesmo os falantes nativos do inglês nem sempre conseguem identificar as vogais de sua língua em contextos de palavras isoladas. Em um estudo envolvendo 70 ouvintes nativos do inglês e testes de discriminação perceptiva para 10 monotongos em um contexto /h_d/, /ɪ/ (*hid*) foi erroneamente identificado como /ɛ/ (*head*) em 7% dos casos e /ɛ/ (*head*) foi identificado como /æ/ (*had*) em 9% das vezes⁴.

O contexto discursivo e as intenções comunicativas do ouvinte também influenciam como ele “entende” aquilo que lhe é dito. Em recente viagem à Nova Iorque, ao fazer o *check-in*, a recepcionista do hotel no qual me hospedei fez-me a seguinte pergunta: *Do you have kids?*. Automaticamente respondi: *Yes, I have two kids*. Um tanto surpresa, ela repetiu a pergunta e percebi, então, que não a havia entendido. Na realidade, ela havia me perguntado: *Do you have keys?*

Acredito que minha compreensibilidade (capacidade de compreensão) mostrou-se imprecisa, pois estava, na realidade, procurando um quarto maior que acomodasse minhas filhas, de modo que a palavra *kids* era aquela que inconscientemente eu esperava ouvir. Além disso, não havia, em minha memória, nenhum *script* ou esquema construído pela minha vivência sócio-cultural para contextos similares. Em suma, havia construído o *contexto* baseado em minhas próprias intenções comunicativas e não a partir da situação real da língua em uso.

⁴ É importante salientar que a distribuição desses sons (em palavras) está sujeita a variações dialetais.

1.5 INFLUÊNCIA DA INSTRUÇÃO FORMAL NO GRAU DE SOTAQUE ESTRANGEIRO

Outra variável relevante no contexto de ensino de inglês como L2 é aquela relacionada ao ensino formal do idioma. Segundo Flege e Lui (2001), vários estudos sobre a influência da instrução formal no grau de sotaque estrangeiro trazem resultados pouco animadores para professores de inglês. Variáveis relacionadas à instrução formal parecem ter pouco efeito em testes de percepção que medem o grau de sotaque estrangeiro. Isso provavelmente decorre do fato de que a maioria dos cursos de inglês não dão muita atenção ao ensino da pronúncia.

Bongaerts *et al* (1997) identificaram cinco aprendizes tardios de inglês cujas avaliações de fala eram comparáveis àsquelas obtidas por falantes nativos. Segundo os autores, esses aprendizes foram bem sucedidos porque receberam treinamento intensivo na percepção e produção dos sons do inglês.

Moyer (1999), por sua vez, analisou a fala de aprendizes ingleses de alemão. Ela pôde constatar que aqueles que haviam recebido treinamento sobre os aspectos segmentais e prosódicos do alemão foram avaliados positivamente pelos juízes nativos, e a fala desses falantes aproximou-se da fala nativa.

Resultado semelhante foi obtido por Barbosa M. (2013) em experimento sobre o grau de sotaque estrangeiro de aprendizes brasileiros de inglês. Em teste de percepção, juizes nativos de inglês ouviram gravações da leitura em inglês de sentenças aleatorizadas, produzidas por sujeitos brasileiros e norte-americanos. Os juizes classificaram apenas um sujeito brasileiro na mesma categoria dos norte-

americanos, justamente aquele que possuía prática e conhecimento linguístico extensivos sobre os aspectos prosódicos do inglês.

Atualmente, o professor de inglês dispõe de uma vasta gama de materiais didáticos direcionados à pronúncia, mas são escassos os materiais baseados na investigação fonético-acústica. Mauad (2010a) argumenta a favor do papel da fonética acústica para auxiliar ambos, professor e aprendiz de Inglês, na percepção e produção dos aspectos prosódicos relacionados ao sotaque estrangeiro:

A pedagogic approach to the teaching of L2 prosody based on acoustic phonetic tools can be of utmost value in assisting both teacher and learner. On the one hand, it allows the teacher-researcher to identify language specific prosodic strategies used in the construction of meaning. On the other hand, it gives the learner immediate feedback on his/her oral production... When language learners are given the opportunity to hear, “see” and interpret the L2 speech signal, they activate those prosodic cues responsible for perceiving and understanding spoken language and ultimately become more competent language users (MAUAD, 2010a, p. 163).

Em experimento com aprendizes brasileiros de inglês, Rocca (2003) constatou, por meio de medições de parâmetros acústicos e avaliações em testes perceptivos, melhora significativa na produção das *yes/no questions* no grupo experimental. Esse grupo havia recebido treinamento mediado por *software* de análise de fala sobre os aspectos entoacionais do inglês, o que lhes permitiu visualizar sua própria fala e, assim, aprimorar sua produção. Para a autora, “a conscientização *daquilo que o ouvido não ouve e daquilo que o aparato vocal não produz* é essencial para o desenvolvimento da aquisição de línguas” (ROCCA, 2003, p. 175).

1.6 HIPÓTESE DO PERÍODO CRÍTICO PARA O APRENDIZADO DE UMA LÍNGUA

Muitos teóricos defendem a existência de um período crítico para o aprendizado da fala humana e a ideia de que o domínio total de uma segunda língua não seria possível após o término desse período. Lenneberg (1971 *apud* Moyer 2004) afirma que a base neurológica para o período crítico para o aprendizado de uma língua pode estar relacionada a mudanças eletroquímicas no cérebro, muitas das quais atingem um estado estável por volta dos 10-12 anos de idade.

Acredita-se que isso se deva à lateralização, ou seja, ao enquadramento de funções de linguagem específicas para cada hemisfério cerebral (esquerdo ou direito). Quando o processo de lateralização completa-se (por volta do início da puberdade), a aquisição de uma língua passa a ser menos completa e menos eficaz, visto que o cérebro já se encontra menos maleável.

Cruz-Ferreira (2010), por sua vez, faz reservas à ideia de que haveria um período crítico para o aprendizado de uma segunda língua. Ela afirma ser muito difícil isolar a variável idade das demais, visto que o ambiente em que adultos e crianças aprendem uma segunda língua são muito diferentes. A autora afirma que, de fato, com o passar dos anos, podemos perder uma habilidade que antes tínhamos. Mas, diz a autora, afirmar que o cérebro humano “fecha-se” para o aprendizado de uma língua após certa idade não específica (“por volta dos 10-12 anos”) é uma afirmação no mínimo vaga. Aceitá-la pressupõe que, independentemente do que o aprendiz faça, ele nunca será bem sucedido no aprendizado de uma nova língua, posto que seu cérebro já teria passado de um estado cognitivo receptivo ao aprendizado de L2.

Tal opinião parece ser compartilhada por Flege (1986) quando o autor afirma que a fala não se torna imutável nem se cristaliza em algum ponto do desenvolvimento do indivíduo nem com a sua maturidade. Por outro lado, o autor reconhece que crianças e adultos diferem no aprendizado de uma língua estrangeira, pois os adultos já criaram uma representação fonética para sons da língua materna e, assim, podem interpretar os sons de L2 de acordo com categorias preestabelecidas.

1.7 ENSINO-APRENDIZAGEM DE RITMO DO INGLÊS E OS MATERIAIS DIDÁTICOS PARA ENSINO DE LÍNGUA INGLESA

Os materiais didáticos voltados ao ritmo do inglês estão repletos de exercícios em que há defesa implícita da isocronia acentual⁵. Cauldwell (2002) credita a longevidade dessa visão reducionista de isocronia ao fato de essa teoria ser aparentemente fácil de ser apresentada. O autor cita um exemplo de exercício extraído de Underhill (1994, p. 71), direcionado a professores de inglês como língua estrangeira, no qual o professor é instruído a pronunciar as sílabas proeminentes (em letra maiúscula) com o mesmo espaçamento de tempo, independente do número de sílabas entre as proeminências, conforme ilustra a Figura 1:

a.	YOU	ME	HIM	HER
b.	YOU and	ME and	HIM and	HER
c.	YOU and THEN	ME and then	HIM and then	HER
d.	YOU and then it's	ME and then it's	HIM and then it's	HER

Figura 1 – Exercício para a prática do ritmo. Fonte: Underhill (1994, p. 71)

⁵ Intervalos de tempos iguais entre as proeminências marcadas por acento de *pitch*.

Buuren (2000) também critica o material didático voltado à pronúncia de inglês para estrangeiros por difundir erroneamente a ideia de que os intervalos entre as proeminências (ou *stresses*) é constante:

The only approach available to English teachers, to my knowledge, is the so-called stress-timing approach. In my talk I tried to show this does not work by saying The North Wind and the Sun⁶ with “isochronous stresses”. It makes one sound rather like a square-dancing master intoning “take your partner, and walk along, hold her hand and step to the left, now turn around and ...”, etc. Any such attempt at a regular “isochronic” rhythm would indeed be quite ridiculous in ordinary speech. So perhaps we should try a different “allochronic” approach.

Creio ser mais produtivo sensibilizar os aprendizes apenas ao fato de que palavras de conteúdo, tais como substantivos, verbos, adjetivos e advérbios, geralmente apresentam uma sílaba proeminente, que as vogais em sílabas átonas estão sujeitas à redução e que, embora a regularidade entre as proeminências seja *percebida* na fala, os grupos acentuais (pés rítmicos) não têm a mesma duração.

Conforme alerta Dauer (1987, p. 449), “non-native speakers of English can improve their rhythm enormously by reducing unstressed syllables, and this is usually more successful than trying to get them to equalize stress beats”. O exercício a seguir, na Figura 2, extraído de Lujan (2004), cumpre bem essa finalidade:

⁶ The North Wind and the Sun were disputing which was the stronger, when a traveler came along and wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler taking his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

Directions: Practice using reduced forms. Identify the matching rhythm for the items in each group of exchanges. The first one has been done for you.

<p>Did you agree? Did you apply? Where had they gone? Would you have known?</p> <p>Wasn't that plenty? Wouldn't it matter? Didn't she make it? Hadn't he noticed?</p> <p>When should we meet? When could we go? When will you call? Where should we meet?</p>	<p>○ ○ ● ○ ● I could not agree. I applied today. They had gone to town. I might not have guessed.</p> <p>It isn't merely enough. It wouldn't matter at all. She didn't make it on time. She might've noticed it too.</p> <p>We can meet after lunch. I could leave in an hour. We will call you at ten. We should meet at the park.</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figura 2 – Exercício para percepção de proeminências. Fonte: Lujan (2004)

Na atividade da Figura 2, todas as palavras que recebem proeminência por acento de *pitch* são palavras de conteúdo (*content words*). Após ter praticado a redução silábica nas *function words* em exercício anterior, o aluno lê a série de 3 enunciados com características rítmicas distintas e, ao enfatizar as sílabas tônicas e reduzir as átonas, é levado a desvendar o padrão rítmico subjacente à cada bloco (no segundo bloco de enunciados da Figura 2, o padrão rítmico é ○ ● ● ○ ○ ●, e no terceiro é ○ ○ ● ○ ○ ●).

As *content words*, tais como os verbos, os substantivos e os adjetivos, têm potencial para receber proeminência acentual, enquanto as *function words*, tais como preposições e artigos, são geralmente não proeminentes. Dada a sua importância na

estruturação rítmica dos enunciados, é bastante louvável e desejável que livros voltados à pronúncia do inglês incorporem essa distinção.

Conforme salienta Dauer (1993):

In order to achieve a good rhythm in English, you need to slow down, stretch out, and very clearly pronounce one-syllable content words and the stressed syllables of longer words. And you must reduce unstressed words and other unstressed syllables. The two most common mistakes made by non-native speakers are pronouncing one syllable content words too quickly, by rushing them and dropping final consonants, and not reducing function words and unstressed syllables enough (DAUER, 1993, p. 95).

O exercício da Figura 3, extraído de Grant (2004), é uma forma simples e prática de sensibilizar o aluno sobre as diferenças semânticas entre as *function words* e as *content words* e sobre o fato que as últimas estão mais propensas a receber proeminência acentual.

<p>Listen to the incomplete dialogue below. Even though more than half the words are missing, can you understand the dialogue?</p> <p>CUSTOMER: _____ possible _____ fly _____ Los Angeles _____ Sunday?</p> <p>AGENT: Yes, _____ couple _____ flights. One _____ 9:30 _____ other _____ 3:15.</p> <p>CUSTOMER: What _____ fare _____ coach?</p> <p>AGENT: _____ round-trip fare _____ \$318.00 plus tax. _____ want _____ make _____ reservation?</p>

Figura 3 – Exercício para a percepção de *content* e *function words*. Fonte: Grant (2004)

No momento em que os alunos percebem que as *function words* podem ser facilmente deduzidas pelo contexto, tanto no texto escrito quanto oral, eles compreendem melhor as razões que levam o falante nativo a reduzi-las.

Finalmente, considerando-se que o *pitch accent* é a pista acústica mais relevante nas proeminências acentuais em inglês (BOLINGER, 1986), é importante que as atividades para a prática do ritmo contemplem a relação entre a entoação e o ritmo.

O exercício seguinte (Figura 4), extraído de Mauad *et al* (2011), requer que o aluno ouça um enunciado e identifique as palavras proeminentes enquanto observa seu contorno entoacional.

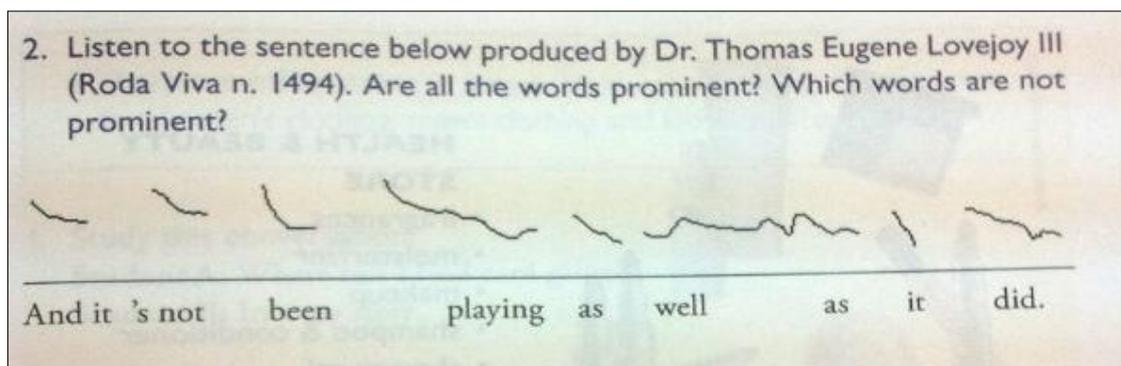


Figura 4 – Exercício para a percepção das proeminências acentuais a partir da visualização da curva entoacional. Fonte: Mauad *et al* (2011)

Nos exercícios anteriores a este, o aluno havia sido apresentado às diferenças semânticas entre as *content words* e as *function words* e, em um segundo momento, identificado as *content words* em frases isoladas. Nesta terceira etapa, o aluno pôde, por meio de *software* de análise de fala, visualizar a curva entoacional e, assim, ter apoio extra para a realização das proeminências acentuais.

1.8 TRANSFERÊNCIAS DA LÍNGUA MATERNA (L1)

1.8.1 Transferências relacionadas à tonicidade

Em minha prática didática, percebo alguns desvios recorrentes na fala de aprendizes brasileiros de inglês relacionados à tonicidade. Nos próximos parágrafos, listarei alguns deles.

Algumas palavras polissilábicas como *development*, *obstacle*, *vegetable* e *catholic* tendem a ser pronunciadas como paroxítonas (*deve'lopment*, *osb'tacle*, *vege'table* e *ca'tholic*).

Verbos terminados em *-ate*, tais como *e'labo,rate*, *'gradu,ate* e *'separ,ate*, tendem a ser pronunciados como oxítonos: *elabor'ate*, *gradu'ate* e *separ'ate*. Essas mesmas palavras, quando na forma de adjetivos ou substantivos (*e'laborate*, *'graduate* e *'separate*, com [ɪt] na última sílaba), tendem a ser pronunciadas de maneira análoga aos verbos (como oxítonas terminadas em [ɛɪt]).

Além disso, aprendizes brasileiros apresentam dificuldade em implementar a acentuação tônica em substantivos cujos verbos correspondentes têm a mesma grafia (ex: *'conflict (n.) / to con'flict (v.)*; *'increase (n.) / to incr 'ease (v.)*; *'suspect (n.) / to sus'pect (v.)*; *'insult (n.) / to ins'ult (v.)*).

A tonicidade nos substantivos compostos também ocorre de forma distinta em português e inglês. Tomemos um termo recente utilizado nas ciências da

computação: *mother board* /'mʌðər ,bɔrd/. O primeiro substantivo recebe o acento primário e o segundo o acento secundário, como era o esperado dado o predomínio à esquerda nos processos linguísticos em inglês.

Em português, o mesmo não ocorre. Na palavra *placa-mãe* /,plaka 'mɛj/ é o segundo substantivo que recebe o acento primário. Não surpreende, portanto, a forte tendência de o aprendiz brasileiro dar mais proeminência ao elemento à direita em vários *compound nouns* do inglês, tais como *cable car*, *ice cream* e *weekend*. Para um ouvinte nativo do inglês, essa pronúncia iria automaticamente remeter a uma frase descritiva e não a uma palavra. Assim, *weekend* seria percebida como *weak end*, como na sentença *Do not place the box on its weak end*⁷.

Hirst e Di Cristo (1998), por sua vez, também relacionam esse parâmetro de dominância linguística (*right/left headness*) com a estrutura rítmica ao afirmar:

We also suggested that if such structuring is in fact part of the mental representation of speech, we should expect this to give rise to different durational patternings in these languages (HIRST e DI CRISTO, 1998, p. 41).

A dominância à direita no português em contraste com à esquerda no inglês nos casos acima apontados pode ser fonte de interferência nas produções em inglês por falantes nativos do português.

⁷ Tal pronúncia foi também observada por Garcia e Otheguy (1989) para o *Chicano English*, dialeto do inglês norte-americano falado por descendentes mexicanos.

1.8.2 Transferências relacionadas às vogais reduzidas

Outra dificuldade do aprendiz brasileiro que pode ocasionar desvios de transferência em inglês está relacionada às particularidades do sistema vocálico dessa língua. Conforme afirma Malmberg (1988), no aprendizado de uma língua estrangeira, é necessário que o aluno aprenda e domine distinções dessa língua inexistentes em sua língua ou sem valor funcional. A palavra *information* é ideal para ilustrar esse ponto:

in	for	ma	tion
ɪn	fə	meɪ	ʃən
sílaba plena	sílaba reduzida	sílaba plena	sílaba reduzida

Ela é composta de sílabas com duas vogais plenas: /ɪ/ em *in* e /eɪ/ em *ma*, e duas vogais reduzidas: /ə/ em *for* e /ə/ em *tion*. Por conta da interferência da grafia, os brasileiros tendem a pronunciar o primeiro /ə/, em *for*, como [or]; e o segundo /ə/, em *tion*, como [õ:]. Mesmo aqueles que produzem o /ə/ tendem a pronunciá-lo com maior duração do que o esperado. Esse processo fonológico do português é apontado por Dauer (1987, p. 448), quando a autora afirma que, em português, “the unaccented vowel system is smaller than that of accented vowels, but unaccented vowels are not necessarily centralized”.

O *schwa* /ə/ é a vogal reduzida de maior incidência na língua inglesa. No nível lexical, ela aparece com bastante frequência em posições pré e pós-tônica, como

em *approximate* [ə'prɒksəmət] e *apartment* [ə'pɑ:tmənt]. Como no português brasileiro a distribuição de vogais reduzidas é mais restrita, os aprendizes brasileiros de inglês tendem a pronunciar essas palavras com um número maior de sílabas plenas. Em outras palavras, a percepção do sotaque prosódico dá-se por conta da substituição de vogais reduzidas por vogais plenas.

Lieff e Nunes (1993), em trabalho sobre as características da pronúncia de aprendizes brasileiros de inglês, apontam que muitos dos desvios de transferência ocorrem, porque várias palavras de função gramatical em português (como *poder, dever, deveria, ele, ela, e você*) têm mais de uma sílaba, enquanto que em inglês essas mesmas palavras são monossilábicas e apresentam formas reduzidas e formas não reduzidas.

Conseqüentemente, ao transferir esse aspecto do português para o inglês, os aprendizes brasileiros tendem a sempre pronunciar essas palavras (*can, she, would* etc) com vogais plenas. O mesmo acontece com as vogais pré-tônicas de palavras de conteúdo, como ocorre, por exemplo, na palavra *apartment*, em que a primeira vogal /ə/ é pronunciada como /a/.

Na posição pós-tônica em português, ocorrem vogais reduzidas, caracterizadas acusticamente por diferenças relativas à frequência de formantes, duração e intensidade. As sete vogais orais em posição tônica reduzem-se a três em posição pós-tônica /i/, /u/ /a/ como em *fale, falo e fala*. Também é comum o desvozeamento de segmentos vocálicos não acentuados em posição final de palavra, como aponta Cristófar-Silva (2001) e, acrescentamos, em contexto fonético desvozeado ou antes de pausa silenciosa.

A transferência desse padrão de redução na vogal pós-tônica em português talvez

explique o porquê do enfraquecimento da última sílaba das palavras *city* e *busy* na pronúncia de alguns aprendizes brasileiros de inglês, fazendo com que essas palavras sejam às vezes percebidas como [sɪt] e [bɪz]⁸.

Além disso, aprendizes de inglês como L2 tendem a ignorar as diferenças de duração entre as sílabas plenas e reduzidas, o *vowel timing*, que desempenha papel relevante no ritmo do inglês.

Em estudo desenvolvido por Fokes, Bond e Steinberg (1984), os padrões de acento de falantes não nativos de cinco línguas diferentes foram comparados com aqueles de falantes nativos do inglês norte-americano. Foram analisadas as características espectrais e duracionais das vogais, e constatou-se que as produções de fala dos não nativos de inglês apresentavam, de forma sistemática, vogais reduzidas mais longas e vogais acentuadas mais curtas do que a fala nativa.

Em relação à percepção de sotaque estrangeiro, Koo (1997) aponta que juízes nativos de inglês consideraram a não implementação dos padrões rítmicos, incluindo aqui a redução vocálica, como o problema mais sério enfrentado pelos aprendizes coreanos de inglês como L2.

Gut (2003) investigou o ritmo da fala de aprendizes de alemão no que concerne à transferência dos padrões de L1. Os dados, registros de leitura de uma história, foram produzidos por três grupos distintos de sujeitos, com L1 distintas: polonês, italiano e chinês. Os resultados apontaram diferenças significativas entre os falantes nativos e não nativos em relação à redução e elisão vocálica e às medidas de

⁸ Segundo relato de uma norte-americana, a pronúncia da palavra *faca* por brasileiros causa certo embaraço, posto que ela lhe soa como *fuck*.

duração das sílabas reduzidas e plenas.

Os falantes não nativos reduziram menos as vogais, e a elisão da vogal mostrou-se bastante rara, principalmente para os sujeitos chineses. Além disso, para todos eles, as sílabas adjacentes não reduzidas foram, em média, de igual duração. Já no caso dos falantes nativos do alemão, as sílabas plenas precedendo sílabas com vogais reduzidas ou totalmente elididas foram em média 1,87 vezes mais longas.

1.8.3 Transferências da L1 relacionadas à fala coarticulada

Na fala encadeada, não há limites precisos entre os segmentos fônicos. Na verdade, os segmentos se sobrepõem, ou seja, são coarticulados, uma vez que o fluxo de fala é composto de elementos gradientes e não discretos. Outra dificuldade enfrentada pelo aprendiz brasileiro de inglês está relacionada ao modo como ocorre a implementação desses fenômenos coarticulatórios.

Em Silva (2009), foram comparadas produções de americanos e brasileiros de enunciados de um monólogo, entre eles o enunciado *Let me see*. Ao analisar as produções desse enunciado, Silva constatou que os brasileiros produziram o [t] seguido de [m] enquanto que os americanos produziram apenas o [m].

Na parte direita das Figuras 5 e 6, apresentamos, respectivamente, uma emissão de um americano e de um brasileiro do enunciado *Let me see* do *corpus* analisado por Silva (2009).

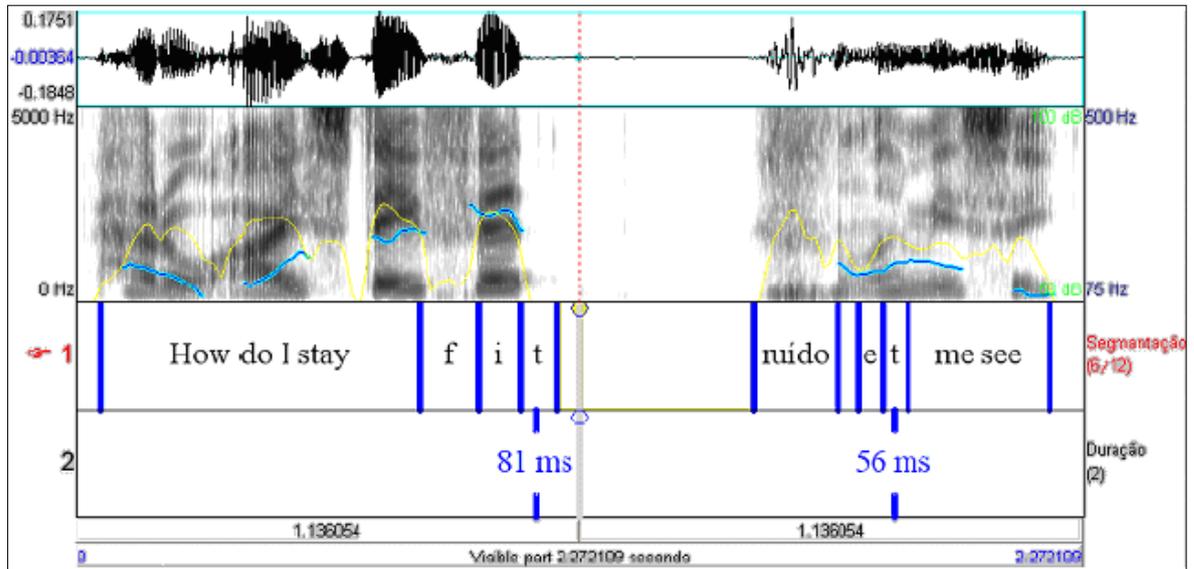


Figura 5 – Forma da onda e espectrograma de banda larga da emissão *Let me see*, produzida por um falante norte-americano. Fonte: Silva (2009, p. 79)

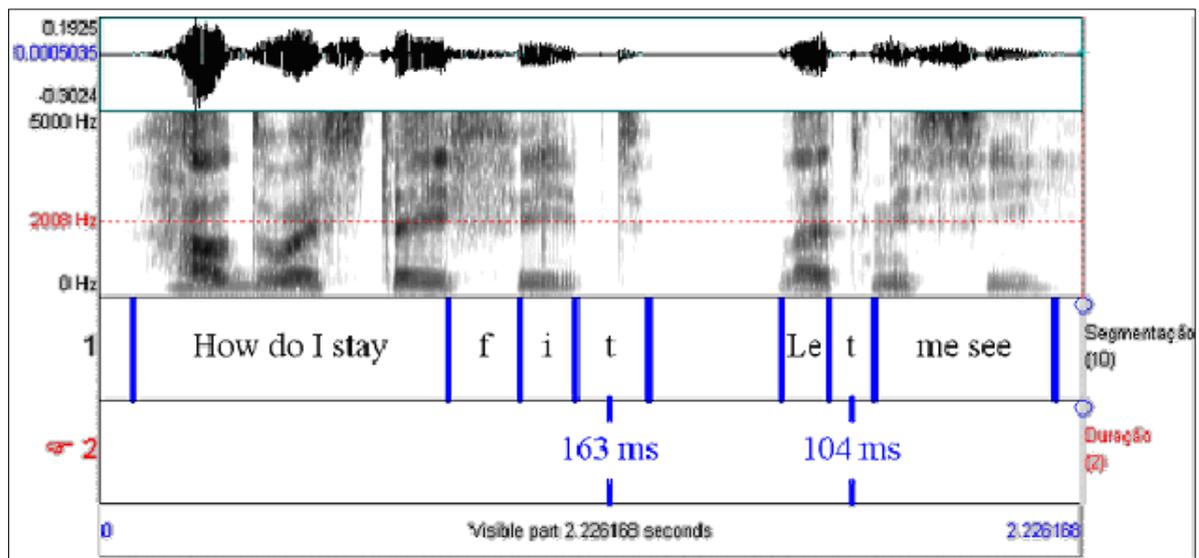


Figura 6 – Forma da onda e espectrograma de banda larga da emissão *Let me see*, produzida por um falante brasileiro. Fonte: Silva (2009, p. 81)

Segundo Silva (*opus cit*), nas produções dos americanos houve sobreposição dos gestos articulatórios do /m/ sobre o /t/, ou seja, esses sons foram coarticulados em um fenômeno coarticulatório denominado *hiding* (Browman e Goldstein, 1990), e, nas produções dos brasileiros, a consoante /t/ foi produzida de modo pleno pelo fato de sua relação espaço-temporal não ter sido truncada pelos segmentos adjacentes.

1.8.4 Transferências da L1 relacionadas à fonotaxe

O sotaque estrangeiro de aprendizes brasileiros de inglês pode, também, estar relacionado a diferenças fonotáticas.

Tomemos como exemplo a palavra *stray*. Como a língua portuguesa não aceita a combinação CCCV, a tendência é que a palavra seja pronunciada como uma dissílaba VC CCV [is'treɪ].

Outra transferência bastante comum na fala de aprendizes brasileiros de inglês ocorre por conta da predominância, em português, de estruturas silábicas do tipo CV e pela restrição a certos tipos de consoantes em posição final de palavra. Assim, palavras como *have* /hæv/ e *big* /bɪg/ tendem a ser pronunciadas como [hævi] e *big* [bɪgi]⁹.

A epêntese de [ɪ], processo fonológico que ocorre no português brasileiro entre sequência de consoantes, também pode causar desvios de transferência. Ela ocorre, por exemplo, na palavra *atmosfera* (at[i]mosfera). Assim, ao pronunciar *atmosphere*, o aprendiz brasileiro pode inadvertidamente pronunciar o [i] de forma a romper a sequência de consoantes [tm].

É importante ressaltar, por fim, que essas palavras, quando inseridas no discurso, acabam por afetar a estrutura rítmica, contribuindo, assim, para o sotaque prosódico nas produções dos aprendizes brasileiros de inglês.

⁹ Norte-americanos falantes de português brasileiro também sofrem influência da fonotaxe de L1. Eles tendem, por exemplo, a pronunciar a palavra “Pelé” como [peɪ'leɪ] ou [p'eɪleɪ]. A substituição da vogal /ɛ/ por /eɪ/ na sílaba “lé” ocorre por conta da regra fonotática do inglês que não admite as vogais /ɛ æ ʊ/ em posição final de sílaba.

1.8.5 Transferências relacionadas à entoação

Além das transferências relacionadas a tonicidade, vogais reduzidas, fonotaxe e fala coarticulada, a fala de aprendizes de inglês como L2 pode apresentar características entoacionais transferidas de L1.

Em capítulo intitulado *In Teaching English to Latin Americans*, Pike (1945) reporta a experiência dos professores do English Language Institute na Universidade de Michigan. Os professores norte-americanos identificaram duas características no contorno melódico desses aprendizes: “excessive rise in pitch of unstressed syllables before a pause in the middle of sentences” e “early fall in pitch of sentences”, esta última a mais problemática (PIKE, 1945, p. 107).

Essa característica da entoação de falantes brasileiros de inglês, o tom *low-fall* nas frases declarativas, foi também observada por Lief e Nunes (1993). As autoras afirmam que esse tom, quando transferido para o inglês, resulta em “a flat final pitch movement or in an inaudibility of the last word”. Cruz-Ferreira (2003) acrescenta que no inglês o *low-fall* em sentenças declarativas remete a atitudes de aspereza ou, na melhor das hipóteses, indiferença por parte do falante.

Outro tom do inglês que os aprendizes brasileiros tendem a ter dificuldade de produzir é o acendente na fronteira final de certos pedidos (*requests*). Mauad (2010b) comparou duas versões do mesmo enunciado, uma em português, produzida por uma brasileira, e outra em inglês, produzida por um norte-americano. O autor aponta que o enunciado em português apresenta um tom mais alto no início,

contorno entoacional global descendente¹⁰ e declinação ascendente-descendente na fronteira final (L+H* L%), como pode ser observado na curva entoacional (em azul) na Figura 7:

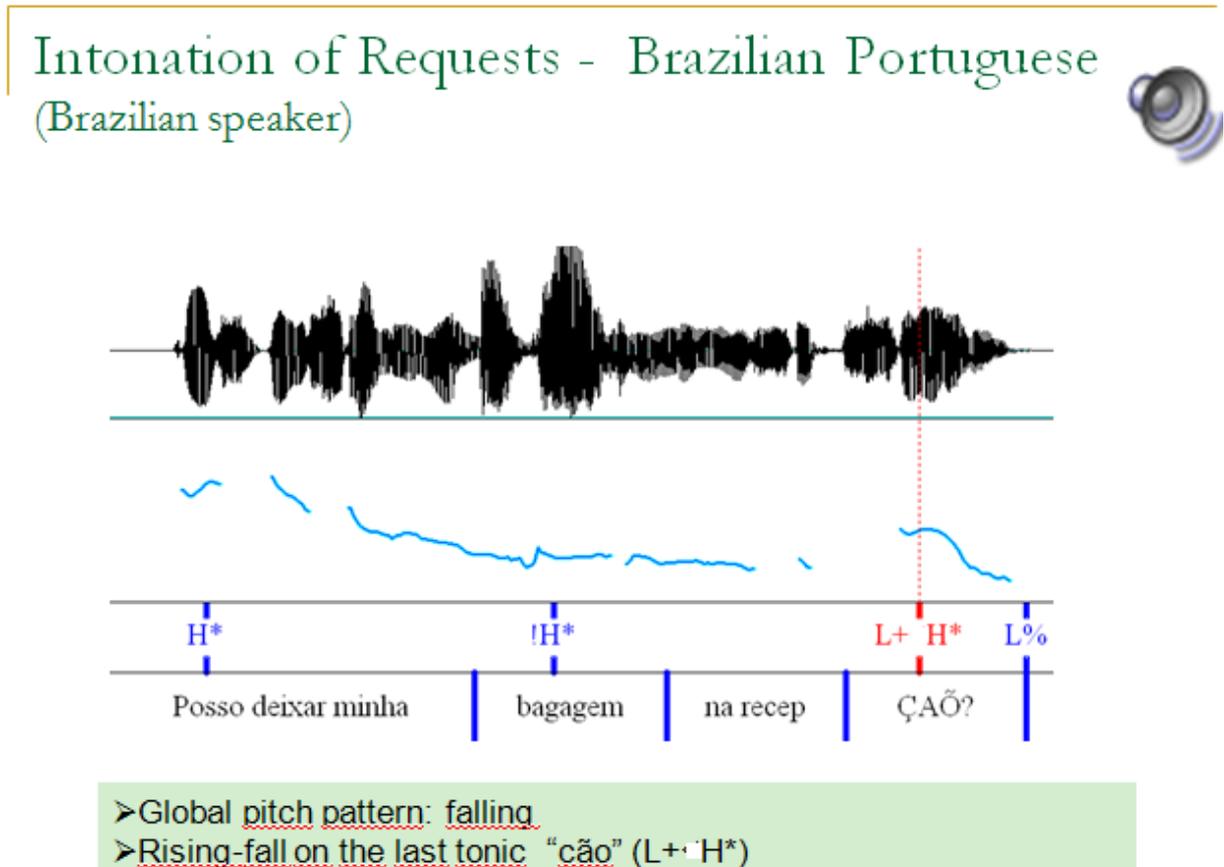


Figura 7 – Forma da onda, curva entoacional e segmentação ortográfica do enunciado *Posso deixar minha bagagem na recepção?*, produzido por uma falante brasileira. Fonte: Mauad (2010b)

Na versão em inglês (Figura 8), por sua vez, o pesquisador observou contorno entoacional global ascendente (como geralmente ocorre nas questões totais ou *yes/no questions* do inglês) e declinação ascendente na fronteira final (L* H-H%).

¹⁰ O contorno entoacional global descendente do enunciado ocorre quando o valor de f_0 , o correlato acústico de *pitch*, no final do enunciado é inferior ao valor apresentado pela sílaba inicial.

Intonation of Requests (North-American speaker)

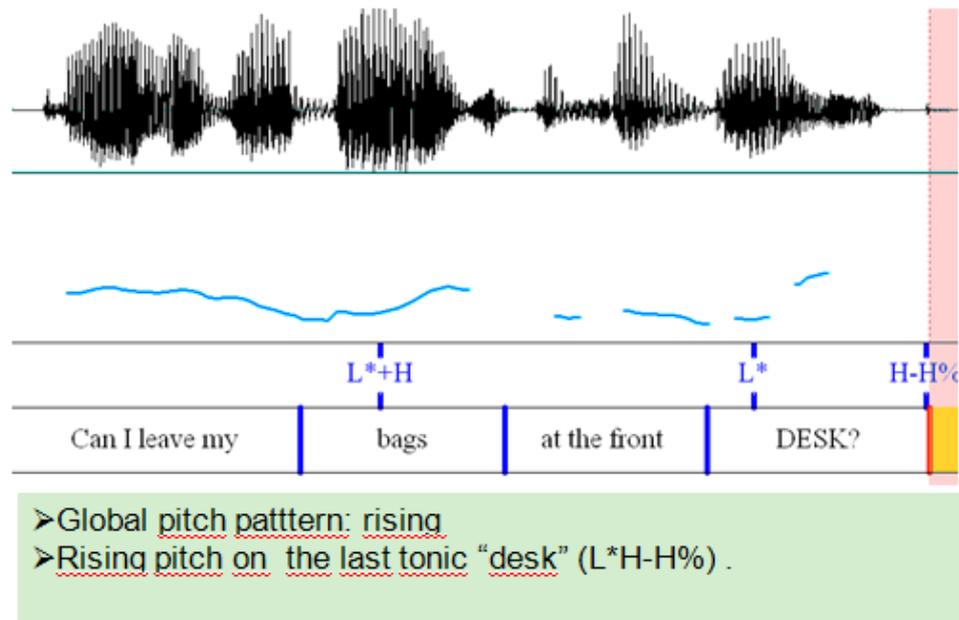


Figura 8 – Forma da onda, curva entoacional e segmentação ortográfica do enunciado *Can I leave my bag at the front desk?*, produzido por um falante americano. Fonte: Mauad (2010b)

O pesquisador afirma que, nesse tipo de *request*, é muito comum os falantes brasileiros de inglês transferirem o padrão entoacional final descendente, fazendo com que o enunciado não seja percebido como um pedido, mas como uma ordem.

O experimento acima descrito evidencia a relevância da entoação na apreensão dos sentidos de ordem pragmática.

Ademais, sabe-se que, em inglês, a frequência fundamental é correlato importante do acento lexical, enquanto que, em português, a duração é o principal correlato acústico do acento lexical (MASSINI, 1991) e frasal (BARBOSA, 2006). Isso talvez explique porque alguns falantes nativos do inglês têm a percepção de que a fala do aprendiz brasileiro de inglês apresenta pouca inflexão tonal (LIEFF e NUNES, 1993).

Por fim, conforme alerta Madureira (2004), engana-se quem pensa que a fala que apresenta poucas variações de altura (*pitch*) não é expressiva: ela pode ser interpretada pelo ouvinte como falta de entusiasmo, apatia, desinteresse, entre outros sentidos. Segue-se que essas questões de expressividade revestem-se de grande relevância em contextos multiculturais e, em particular, no ensino de inglês para brasileiros.

CAPÍTULO 2 QUESTÕES SOBRE O RITMO

DA FALA

Why do we study rhythm? People have a basic need for rhythm, or for the effect produced by it. Rhythm gives pleasure and a more emotional response to the listener or reader because it establishes a pattern of expectations, and rewards the listener or reader with the pleasure that comes from having those expectations fulfilled.

(Brian Black)

2.1 CONCEITUAÇÃO DO RITMO

Platão (*As Leis*, livro II, p. 93) deu uma definição singular de ritmo: “o ritmo é ordem em movimento”.

Ao ritmo são associados os aspectos de regularidade e estruturação (FRAISSE, 1974). A regularidade é percebida na repetição de certas unidades proeminentes da fala, enquanto a estruturação lida com os elementos constitutivos daquilo que está sendo repetido periodicamente (BENGUEREL e D’ARCY, 1986).

O ritmo é, portanto, repetição periódica de um evento e, “quando dizemos que há ritmo em algo, ou que alguma coisa é rítmica, pressupomos a coexistência, ao longo do tempo, de repetição ou regularidade com estrutura ou padrão” (BARBOSA, 2006).

Nos próximos parágrafos, abordarei estudos sobre o ritmo das línguas e, no processo, focalizarei a proposta teórica de Bolinger (1986) para a análise do ritmo do inglês norte-americano.

2.2 ESTUDOS SOBRE O RITMO DAS LÍNGUAS

Em 1940, Lloyd James descreveu dois tipos de ritmo baseados na organização temporal dos eventos recorrentes na fala. Em um tipo de ritmo, esses eventos regulares corresponderiam à duração das sílabas e, em outro, aos intervalos entre os acentos (*stresses*).

Pike (1945) foi o primeiro a utilizar o termo *stress-timed* e *syllable-timed* para apontar tendências nas estruturas rítmicas das línguas, dependendo de qual evento, acento ou sílaba repetia-se em intervalos regulares. O autor já apontava maior *tendência* do inglês norte-americano ao ritmo acentual e do espanhol ao ritmo silábico: “Many non-English languages (Spanish, for instance) tend to use a rhythm which is more closely related to the syllable than to the regular stress-timed type of English” (PIKE, 1945, p. 35).

Pike, no entanto, jamais defendeu diferença categórica entre as línguas baseadas nos dois ritmos. Para o autor, o inglês apresenta aspectos de *syllable-timing* em alguns contextos discursivos, como em alguns trechos de elocução e em trechos da fala cantada¹¹. Neste trabalho, mantereí os termos *stress-timing* e *syllable-timing* para me referir a esses dois tipos de organização temporal.

Abercrombie (1967), em *Elements of General Phonetics*, defende a hipótese de que todas as línguas podem ser classificadas em duas classes rítmicas mutuamente exclusivas: línguas de ritmo acentual e línguas de ritmo silábico:

As far as is known, every language in the world is spoken with one kind of rhythm or with the other. In the one kind, known as syllable-timed rhythm, the periodic recurrence of movement is supplied by the syllable producing process: the chest pulses, and hence the syllables recur at equal intervals of time – they are isochronous. (...) in the other kind, known as stress-timed rhythm, the periodic recurrence is supplied by the stress-producing process: the stress pulses, and hence the stressed syllables are isochronous.

¹¹ Barbosa (2000), como exemplo, cita o discurso de Martin Luther King, que ora apresenta trechos de *syllable-timing*, ora de *stress-timing*.

A seguir, na Figura 9, apresento a representação gráfica desses dois tipos de organização temporal para o ritmo, o *stress-timing* e o *syllable-timing*:

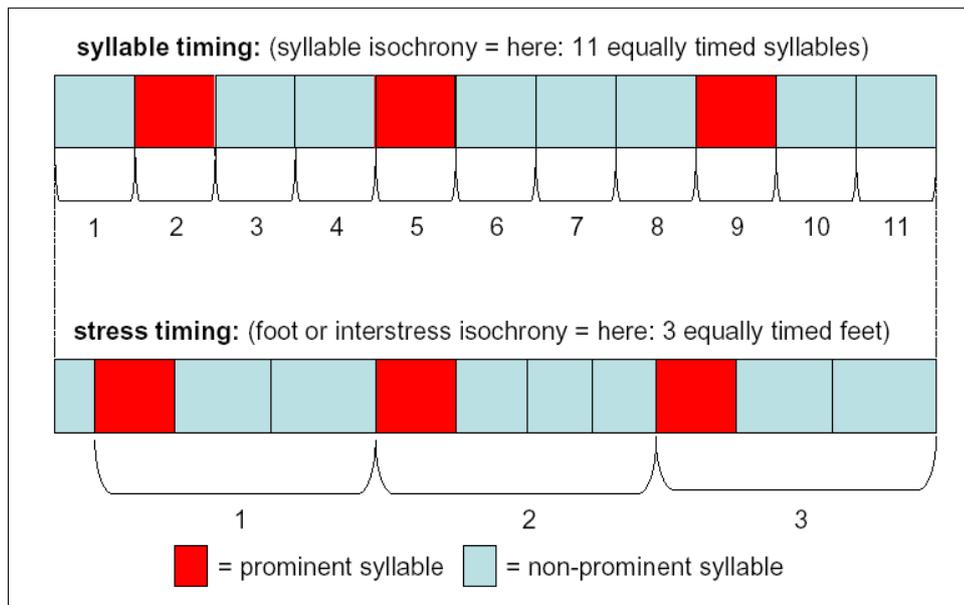


Figura 9 – Representação gráfica de stress-timing e syllable-timing.
Fonte: <http://www.phonetiklabor.de>

No ritmo silábico (topo da Figura 9), cada uma das sílabas, numeradas de 1 a 11, teriam a mesma duração. Haveria, portanto, isocronismo silábico (ocorrência das sílabas a intervalos iguais), independentemente de a sílaba ser acentuada/proeminente ou não¹². As sílabas proeminentes (2, 5 e 9) seriam intercaladas por sílabas não proeminentes de igual duração. Por conta dessa característica, esse ritmo recebe algumas denominações de cunho impressionístico (apenas de outiva) tais como *morse-code like*, *bouncing*, *heart beat* e *tum-te-tum* (CRYSTAL, 2010).

¹² Segundo Barbosa (2000, p. 370), “o termo *isossilabismo* é normalmente usado neste caso, para marcar uma distinção em relação ao uso exclusivo do termo de isocronismo – “isochrony” – para referir apenas às línguas com sílabas acentuadas isócronas”.

O ritmo acentual, por sua vez, seria caracterizado por apresentar os pés métricos acentuais, compostos de sílabas acentuadas que podem vir seguidas de outras sílabas não acentuadas (reduzidas). Para que houvesse isocronia acentual, cada pé acentual deveria se repetir no mesmo intervalo de tempo. Isso faria com que as sílabas entre os acentos fossem comprimidas de modo a manter a regularidade rítmica.

Os quadrados vermelhos da porção inferior da Figura 9 representam os acentos, delimitando três pés rítmicos acentuais de 3, 4 e 3 sílabas, respectivamente. Tomemos os dois primeiros pés: para que houvesse isocronismo acentual, seria necessário: (a) alongamento de uma ou mais sílabas do primeiro pé ou (b) encurtamento de uma ou mais sílabas do segundo pé. Há metáforas e analogias para caracterizar as impressões auditivas desse padrão rítmico, tais como: *machine-gun like*, *pattering*, *staccato* e *rat-a-tat* (CRYSTAL, 2010).

Segundo Patel (2010), os estudos sobre a duração de eventos prosódicos relacionados ao ritmo da fala pecavam justamente por essa imprecisão: a **falsa** asserção de que haveria periodicidade de acentos ou sílabas. Essa busca no domínio da produção pelo isocronismo absoluto foi, segundo Barbosa (2006), consequência da afirmação **não consubstanciada** de Abercrombie (1967) em favor da divisão mutuamente exclusiva para o ritmo das línguas do mundo.

Cauldwell (2002) acredita que a periodicidade na fala esteja apenas superficialmente relacionada à organização temporal, ou *timing*. O autor considera esses episódios como, na realidade, demonstrações da plasticidade da fala:

Coincidental rhythmicity is most likely to occur where there are equal numbers of syllables between stresses. In spontaneous speech, the speaker's attention is on planning and uttering selections of meaning in pursuit of their social-worldly purposes, and this results in an irrhythmic norm which aids comprehension (CAULDWELL, 2002, p. 1).

Para que houvesse isocronia acentual, ou seja, regularidade temporal entre as proeminências, certas propriedades fonológicas precisariam ser atendidas, sendo que as mais apontadas pelos defensores da ocorrência da isocronia acentual e silábica são:

- (a) contraste nítido nas propriedades fonéticas de sílabas proeminentes e não proeminentes;
- (b) acentuação tônica das palavras determinadas por regras lexicais (e, portanto, específicas de cada língua);
- (c) maior alongamento compensatório das sílabas proeminentes em contextos em que ocorre choque de proeminências, de modo a fornecer o espaço temporal necessário para o possível movimento tonal;
- (d) maior encurtamento compensatório de sílabas proeminentes, redução vocálica e compressão temporal de sílabas não proeminentes, para que as diferenças de tamanho de pés rítmicos não comprometam a regularidade.

Vários pesquisadores passaram a investigar se haveria realmente isocronismo na fala. Couper-Kuhlen (1993), com o intuito de analisar o grau de *stress-timing* no inglês, conduziu experimento perceptivo em que dois sujeitos ouviram uma gravação de 2 minutos extraída de um programa de rádio que recebia ligações de ouvintes. A

cada acento percebido, os sujeitos foram instruídos a bater o lápis na mesa ou menear a cabeça.

Os sujeitos identificaram 48 episódios de *stress-timing* na fala, mas houve momentos em que não perceberam nenhuma regularidade. Couper-Kuhlen acredita que os dados sugerem que o inglês não é uma língua isócrona do ponto de vista acentual quando vista de uma macro-perspectiva de toda a extensão temporal do texto falado. Por outro lado, sob a micro-perspectiva das características internas de cada um dos 48 trechos, ele demonstrou ser isócrono.

Roach (1982), por sua vez, desenvolveu um trabalho de investigação fonética para averiguar se as afirmações de Abercrombie poderiam ser validadas. Ele utilizou as mesmas línguas (francês, telugu, ioruba, inglês, russo e árabe) em *corpus* de fala semiespontânea. A seguir, apresento as duas afirmações de Abercrombie seguidas dos resultados obtidos por Roach:

(1) A duração silábica estaria sujeita a maior *variabilidade* nas línguas de ritmo acentual do que nas línguas de ritmo silábico, uma vez que, se a duração do pé rítmico¹³ é constante, deduz-se que a duração da sílaba tem de variar em função do número de sílabas nesse grupo.

Os dados de Roach (1982), o desvio padrão (em ms) das medidas de duração dos enunciados nas seis línguas, mostraram que, realmente, o inglês língua classificada como de ritmo acentual, obteve o maior valor (86 ms) e o ioruba, tradicionalmente classificado como de ritmo silábico, obteve o menor valor (66 ms). Entretanto, nas demais línguas, os valores encontrados eram muito próximos

¹³ Grupo acentual composto de sílaba acentuada + sílaba não acentuada.

(75.5, 81, 76 e 77 ms) e, portanto, a distinção feita por Abercrombie não pôde ser corroborada por dados acústicos.

(2) A duração do pé rítmico deveria variar em graus diferentes para os dois tipos de língua: ela deveria ser significativamente maior nas línguas de ritmo silábico do que nas línguas de ritmo acentual, pois apenas nas línguas de ritmo acentual ocorreria a compressão das sílabas para compensar o número maior de sílabas no grupo acentual.

Os dados de Roach (1982), desvio padrão de duração do grupo acentual (pé rítmico), também não corroboraram a segunda afirmação de Abercrombie: na verdade, houve maior variabilidade nos pés rítmicos nas línguas de ritmo acentual.

O estudo sobre a relação entre duração do grupo acentual e número de sílabas prossegue até hoje. Em trabalho recente, Barbosa (2000) propôs uma equação de reta com as variáveis “número de sílabas (n)” e “duração do pé rítmico (l)” (ou “grupo acentual”, segundo a terminologia adotada pelo autor):

$$l = a + b.n$$

sendo que:

a = ponto de intersecção da reta no eixo das ordenadas

b = inclinação da reta (tangente do ângulo entre a reta e o eixo das abscissas)

Deduz-se que só poderia haver isocronismo acentual absoluto quando a duração do pé l fosse constante, independentemente do número de sílabas. Isso só seria

possível se b fosse igual a zero, como representado por uma reta paralela no eixo das abscissas, no Gráfico 1, a seguir:

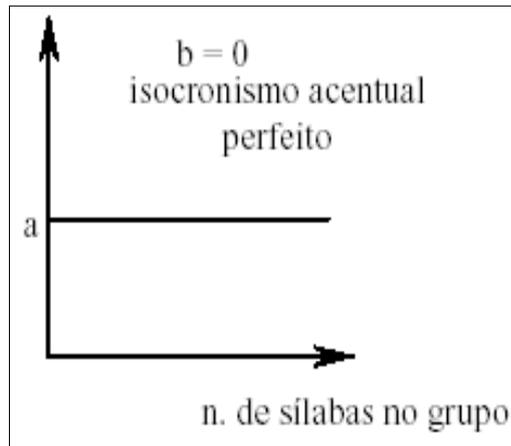


Gráfico 1 – Isocronismo acentual ideal. Fonte: Barbosa, 2006

Sendo que:

$$l = a + b.n$$

$$l = a + b.0$$

$$l = a$$

No caso do isocronismo silábico, a duração dos pés seria diretamente proporcional ao número de sílabas daqueles, ou seja, quanto mais sílabas houver no pé maior o tempo que se leva para produzi-lo. Isso só seria possível se a fosse zero, representado por uma reta passando pela origem, no Gráfico 2, a seguir:

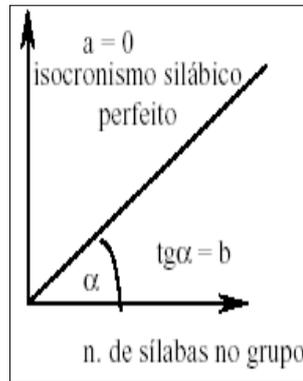


Gráfico 2 – Isocronismo silábico ideal. Fonte: Barbosa, 2006

Em que:

$$l = a + b.n$$

$$l = 0 + b.n$$

$$l = b.n$$

Finalmente, o terceiro padrão parece ser o que representa de maneira mais precisa a estrutura rítmica das mais diversas línguas: nem a nem b são zero, o que significa que a duração do pé rítmico é variável. Conclui-se não haver nenhuma língua com isocronismo acentual ou silábico, mas, sim, momentos de *stress-timing* e momentos de *syllable-timing*. Graficamente, isso é representado por uma reta iniciando-se em um ponto da ordenada, e não paralela à abscissa, conforme o Gráfico 3, a seguir:

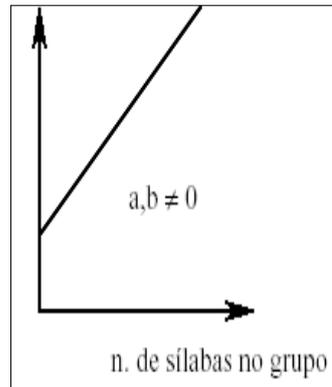


Gráfico 3 – Ausência de isocronismo: pé métrico com duração variável. Fonte: Barbosa, 2006

Onde:

$$l = a + b.n$$

Dauer (1983) estudou o ritmo das línguas a partir da análise do grau de *stress-timing* de quatro línguas (inglês, espanhol, grego e italiano). Conforme explica Barbosa (2000), Dauer propôs uma escala, caracterizada pela dimensão *stress-based* (que se poderia chamar *stress-timing*), em relação a qual “a language is more or less stress-based, depending on how large a role stress plays in that language, especially with respect to the three areas above” (DAUER, 1983, p. 59).

Em outras palavras, as línguas ditas de ritmo acentual e de ritmo silábico teriam um conjunto de propriedades fonéticas e fonológicas distintas que as caracterizariam em grupos distintos (TENANI, 2006). As três propriedades descritas por Dauer são: (1) a estrutura silábica (predominância na língua em questão de sílabas abertas ou fechadas), (2) o grau de extensão dos fenômenos de redução vocálica e (3) a realização fonética do acento e seu papel no sistema

fonológico da língua em questão. Nos parágrafos seguintes, detalharei cada uma dessas propriedades.

Ao comparar a estrutura silábica do inglês – língua tradicionalmente relacionada à ocorrência de ritmo acentual – com a estrutura silábica do espanhol – língua tradicionalmente relacionada à ocorrência de ritmo silábico – a autora percebeu haver uma diferença significativa entre os tipos de sílabas mais frequentes em cada uma das línguas, conforme mostra a Tabela 1:

Tipo de sílaba	Inglês (%)	Espanhol (%)
CV	34	58
V	8	6
CVC	30	22
VC	15	6
CVCC	6	-

Tabela 1 – Porcentagens de tipos silábicos em inglês e em espanhol. Fonte: Dauer (1983, p. 56)

Percebe-se pelos dados que as sílabas abertas (aquelas que terminam em vogal) constituem minoria em inglês. As sílabas fechadas (aquelas que terminam em consoante) constituem maioria do total em inglês, contra minoria em espanhol. Além disso, no inglês, as sílabas pesadas (aquelas compostas por muitos segmentos) têm tendência maior ao acento e as sílabas leves (aquelas contendo poucos segmentos) tendem a ser não proeminentes.

Além disso, Dauer (1983) pôde verificar aspectos relevantes sobre a redução vocálica: em inglês, 92% das vogais em sílabas CV não acentuadas foram

realizadas foneticamente como vogais centralizadas (as quais Bolinger afirma pertencerem ao subgrupo de vogais reduzidas [i, ɪ, e])¹⁴.

Em espanhol, por sua vez, 90% das vogais das sílabas não acentuadas apresentam-se como vogais periféricas mais longas, não se observando, assim, regular redução das vogais que ocorrem em posição átona. Essas observações sugerem que a variabilidade duracional percebida entre vogais tônicas e átonas é mais marcante em inglês do que no espanhol, o que indica que a qualidade vocálica exerce papel relevante no ritmo do inglês.

Dauer (1983) também observou que a proeminência acentual realiza-se foneticamente de forma diferente nas 4 línguas por ela estudadas. A autora verificou, por exemplo, que, no inglês, a complexidade dos parâmetros acústicos para a realização das proeminências acentuais e o número de regras relacionadas são maiores do que nas outras 3 línguas por ela estudadas.

Em trabalho posterior (DAUER, 1987), a pesquisadora, verificou que, em inglês, as sílabas proeminentes, e especialmente as vogais dessas sílabas, são pelo menos 1,5 vezes mais longas do que as sílabas não proeminentes. Por outro lado, no espanhol, não há diferença tão significativa entre a duração das sílabas proeminentes e a das não proeminentes. A autora verificou, também, uma diferença interessante entre o inglês e o francês: em inglês “pitch (usually high or changing) correlates with accent” enquanto que em francês “intonation and accent are independent” (DAUER, 1987, p. 448).

14 Sobre o grupo das vogais reduzidas, Laver (2002) afirma: “short central vowels are thus the most frequent of all English vowels in running speech, and can be thought of as a continually present background against which the less-frequently occurring longer stressed vowels can be perceived as standing out more prominently.”

Assim, Dauer posiciona o inglês no extremo *stress-based* da escala, por manifestar essas três propriedades (estrutura silábica, redução vocálica e realização fonética do acento) em alto grau.

Para a autora, “rhythm is a total effect involving phonetic and phonological as well as segmental and prosodic phenomena” (DAUER, 1987, p. 447) e as diferenças rítmicas entre as línguas está relacionada com “what goes on within rather than across interstress intervals” (DAUER, 1983, p. 55), ou seja, “any subjective differences in language rhythms are due to factors other than interstress intervals” (COUPER-KUHLEN, 1993, p. 13).

No começo da década de 1990, a discussão sobre as classes rítmicas foi interrompida devido à dificuldade de se medir os correlatos acústicos do ritmo. Conseqüentemente, abriu-se caminho para várias teorias ligadas à percepção e o ritmo passou a ser visto como um fenômeno perceptivo (BENQUEREL e D'ARCY, 1986; BECKMAN, 1992).

As pesquisas recentes sobre o ritmo da fala descartam a dicotomia categórica entre línguas de ritmo silábico e línguas de ritmo acentual, substituindo-a por uma orientação multidimensional que inclui diversas propriedades fonológicas das línguas (SCHLÜTER, 2005). O ritmo da fala passa, portanto, a refletir vários componentes linguísticos, tais como a diversidade da estrutura silábica, os processos de redução vocálica e a influência do acento na duração das vogais (PATEL, 2010).

Recentemente, alguns pesquisadores, em estudos experimentais sobre o ritmo, passaram a estudar também a relação entre o ritmo e a taxa de elocução (BARRY *et al*, 2003; DELLWO e WAGNER, 2003; BARBOSA, 2000; 2006; MEIRELES, 2009)

ao sexo do falante (BARBOSA, 2000) e o dialeto (MEIRELES *et al*, 2010).

Meireles *et al* (2010), em estudo com dois dialetos do português brasileiro (mineiro e baiano) observou (1) que características de ritmo silábico ou ritmo acentual são exacerbadas em uma taxa de elocução mais rápida, (2) que o ritmo do dialeto mineiro é mais acentual que o dialeto baiano e (3) que a fala masculina tende a ser mais rápida. Barbosa (2000), por sua vez, comparando a taxa de dois falantes do português brasileiro, com controle na taxa de elocução (média) e no tamanho do pé rítmico (menos que 5 unidades), concluiu que a produção de fala da mulher aproximava-se mais do ritmo silábico do que a do sujeito masculino.

Baseando-se no modelo de O'dell e Nieminem (1999) de osciladores acoplados, Barbosa (2006) propõe seu próprio modelo de osciladores acoplados para o modelamento do ritmo do português brasileiro. Nesse modelo, um oscilador¹⁵ do tamanho de uma sílaba (o oscilador silábico) interage com um oscilador acentual. O oscilador silábico e o acentual são comuns a todas as línguas posto que são osciladores acoplados no sistema cognitivo. Entretanto, a força de acoplamento entre os dois osciladores é específica para cada língua.

Nesse sentido, a proposta de Barbosa (2006) guarda algumas semelhanças com a teoria de ritmo de Bolinger (1986), na medida em que Bolinger também concebe duas camadas, ou níveis, para o ritmo: silábico e acentual, os quais também interagem.

O desafio dos pesquisadores atuais continua sendo, portanto, o mesmo daqueles do final dos anos 1960: tentar achar correlatos acústicos do ritmo da fala no sinal

¹⁵ Oscilador é definido como um sistema que gera comportamento periódico.

acústico, fazendo correlação entre padrões duracionais e os fenômenos fonológicos envolvidos.

A seguir, apresento a teoria para o ritmo do inglês norte-americano proposta por Bolinger (1986), que servirá de base para a análise rítmica das produções dos falantes norte-americanos e brasileiros.

2.3 RITMO DO INGLÊS NORTE AMERICANO

Segundo Bolinger (1986), assim como há dois tipos de vogais, há também dois tipos de ritmo no inglês norte-americano: o ritmo silábico e o ritmo acentual.

The rhythmic patterns that concern us come in two layers, one based on syllable structure, the other on accents. In the first, or syllabic rhythm, what determines the foot structure is the presence of a full syllable with its following reduced syllable or syllables if any. In the second, accentual rhythm, it is the accented syllable plus its following unaccented ones (BOLINGER, 1986, p. 63).

Bolinger acrescenta que esses dois níveis de ritmo, silábico e acentual, são independentes: “In speech, the two schemes, syllabic rhythm and accentual rhythm, are independent” (BOLINGER, 1986, p. 64).

Bolinger (1978) também aponta a relevância da duração da sílaba proeminente na estrutura rítmica. Ele afirma que, no inglês, pés rítmicos baseados na qualidade vocálica (sílaba plena seguida de sílaba(s) reduzida(s)) têm uma tendência maior a ocorrerem em intervalos de tempos *aproximados* do que aqueles baseados em acentos de *pitch*.

Dada a importância das vogais para o entendimento da estrutura rítmica do inglês, os dois próximos tópicos consideram o inventário vocálico do inglês e apresentam a tipologia de Bolinger para as vogais plenas e reduzidas.

2.3.1 Vogais do inglês

Na Figura 10, a seguir, é apresentado o quadro dos fonemas vocálicos do inglês, de acordo com o IPA (International Phonetic Alphabet):

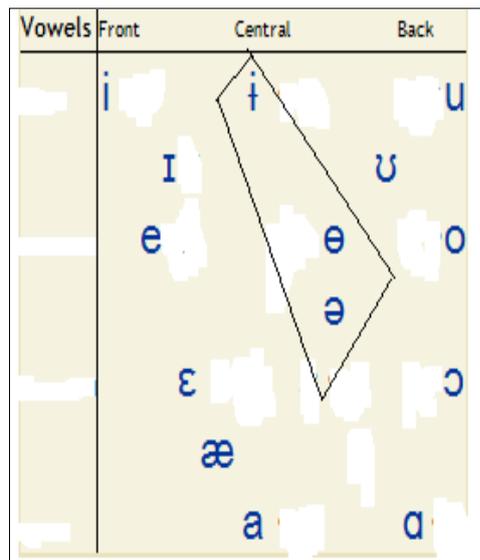


Figura 10 – Quadro das vogais do inglês

Maidment (1981) afirma que a descrição articulatória de um segmento permite que ele seja classificado, facilitando, assim, a especificação de sua função em sistemas e estruturas da língua. Os segmentos vocálicos são descritos considerando-se os seguintes aspectos articulatórios: (1) altura da língua; (2) anterioridade/posterioridade da língua e (3) arredondamento dos lábios.

A altura da língua representa a dimensão vertical ocupada pela língua dentro da cavidade bucal. São consideradas vogais altas /i/, /ɪ/, /i/, /u/ e /ʊ/ ; vogais médias /eɪ/, /ɛ/, /ə/, /ʌ/, /θ/, /o/, /ɔ /; e vogais baixas /ɑ/, /æ/.

A anterioridade/posterioridade da língua refere-se à posição dela na dimensão horizontal durante a articulação do segmento vocálico. A cavidade bucal é dividida em três partes simétricas, e as vogais podem ser classificadas em anteriores (frente da cavidade bucal): /i/, /ɪ/, /eɪ/, /ɛ/, /æ/; centrais (parte central da cavidade bucal): /ə/, /ʌ/, /θ/ e /i/, e posteriores (final da cavidade bucal): /u/, /ʊ/, /o/, /ɔ /, /ɑ/. Durante a articulação de algumas vogais, os lábios ficam arredondados. As vogais arredondadas são as vogais /u/, /ʊ/, /o/, /ɔ /.

Os segmentos fônicos (vogais e consonantes) assim como os prosódicos podem ser analisados quanto aos três parâmetros acústicos: duração, frequência fundamental e intensidade. No caso dos segmentos, referimo-nos à duração, frequência fundamental e intensidade intrínsecas.

As vogais possuem duração intrínseca. Entretanto, a duração vocálica está correlacionada com sua altura: quanto mais baixa for a vogal, e, portanto, maior o caminho a percorrer para sua produção, maior será sua duração intrínseca¹⁶.

Em relação à frequência fundamental, em um contexto fonético, as vogais altas como [i] e [u] apresentam valores mais altos do que as vogais baixas como [a] (PETERSON e BARNEY, 1952; LADD e SILVERMAN, 1984; STEELE, 1986) e,

¹⁶ Lehiste (1970) usa a expressão *the-further-the-longer* para se referir a esse fato.

nas vogais que seguem consoantes não vozeadas, a frequência fundamental apresenta valores mais altos do que no início das vogais que seguem as consoantes vozeadas (SILVERMAN, 1986).

Em relação à intensidade dos segmentos, as vogais são os sons que apresentam maior intensidade enquanto as fricativas /θ/ e /f/, presentes em palavras como *thin* e *fin*, são os sons mais fracos¹⁷.

2.3.1.1 VOGAIS PLENAS E VOGAIS REDUZIDAS

Bolinger (1986) considera a qualidade vocálica uma pista relevante relacionada ao ritmo. Ele defende a existência de um *sistema* composto de duas classes de vogais: vogais plenas (*full vowels*) e vogais reduzidas (*short vowels*).

As vogais reduzidas situam-se dentro do trapezóide no *vowel chart* na Figura 10. As demais vogais serão consideradas plenas. Neste trabalho, utilizarei os termos *plena* e *reduzida* para as vogais por achá-los mais apropriados do que os termos *long* e *short* que podem, a meu ver, suscitar divergências conceituais em relação à duração das vogais.

A proposta teórica de Bolinger (1986) mostra-se bastante relevante na construção de uma tipologia de ritmo¹⁸. Como vimos, a fundamentação de sua teoria é a noção

¹⁷ A banda dinâmica da fala – seu alcance de energia – é de 60 dB, ou seja, os sons mais fortes são cerca de 60 dBs mais intensos que os sons mais fracos. Via de regra, a banda de um falante em particular pode ser facilmente aferida quando ele pronuncia a palavra *thaw* que contem os dois sons limite (KENT, 1992).

¹⁸ Sua teoria tem sido, ao longo dos anos, de grande valia para os estudos prosódicos relacionados ao ritmo e à entoação, conforme afirmam Waugh e Schuuneveld (1980, p. 298): “The study of stress and intonation has been one of Bolinger’s primary interests and his characteristically insightful

de que existem dois conjuntos distintos de vogais no inglês: vogais plenas e vogais reduzidas. Entretanto, o autor considera como *reduzidas* apenas as três vogais /ɪ/, /ə/, /ə/. Para Bolinger, essas três vogais reduzidas fazem parte de um subgrupo de vogais com características próprias.

Assim como Dauer (1983), Bolinger (1986) considera a qualidade vocálica um fator relevante na estrutura rítmica do inglês. A qualidade vocálica é, segundo o autor, uma pista para as proeminências na medida em que sílabas com vogais reduzidas são impossibilitadas de receber proeminência e as sílabas com vogais plenas são aptas a receber proeminência.

A Tabela 2 apresenta uma lista das vogais plenas do inglês com algumas palavras-chaves, a partir de exemplos de Bolinger (1986):

palavra	vogal	palavra-chave	vogal
keyed	[i]	cod	[a]
kid	[ɪ]	cawed	[ɔ]
cade	[e]	code	[o]
ked	[ɛ]	could	[ʊ]
cad	[æ]	cooed	[u]
cud	[ʌ]		

Tabela 2 – Vogais plenas do inglês. Fonte: Bolinger (1986)

contributions have greatly enlivened and advanced the field. Probably everyone who works in this area has had the experience of arriving at a hard-thought-out conclusion about the operation of some intonational mechanism, only to find later that the same perception appeared in a paper of Bolinger's some twenty years earlier (in a footnote). One of his most frequently assumed roles is that of debunker: he is always ready with the counter example which forms an immovable obstacle to the sweeping generalization."

Para Bolinger, são três as vogais reduzidas do inglês: [ɪ], [ə] e [e], as quais são exemplificadas na Tabela 3 a seguir:

palavra	vogal
Willie	[ɪ]
Willa	[ə]
willow	[e]

Tabela 3 – Vogais reduzidas do inglês. Fonte: Bolinger (1986)

Para Bolinger, cada uma dessas vogais reduzidas apresenta características de fonemas. Um som é definido como fonema em função de sua capacidade de distinguir palavras. Por exemplo, /i/ e /ɪ/ constituem-se fonemas distintos no inglês posto que, em um mesmo contexto, eles conseguem distinguir duas palavras (*leave* [liv] vs *live* [lɪv]).

O autor acredita que essas 3 vogais pertençam a uma *classe* (fonológica) de vogais reduzidas do inglês que, portanto, comportam-se diferentemente de outras vogais. Segundo o autor, assim como as vogais plenas, essa subclasse de vogais também distingue palavras, ocorrendo, entretanto, principalmente, em sílabas em posição final de palavra. Como exemplo, Bolinger (1986, p. 38) cita as palavras: *sallied salad swallowed*, respectivamente [ˈsæliɪd], [ˈsæləd] e [ˈsæləd].

Enquanto as vogais plenas distinguem-se uma da outra em três dimensões: (1) altura da língua; (2) anterioridade e posterioridade da língua e (3) arredondamento dos lábios, as vogais reduzidas são indiferentes à primeira e à terceira dessas

dimensões. Em termos práticos, distinguem-se apenas pela anterioridade e posterioridade da língua na cavidade bucal. Em *sallied*, a vogal /ɪ/ é anterior; em *sallowed*, a vogal /ə/ é posterior; e em *salad*, a vogal /ə/ é central.

Uma frase que mostra bem a diferença entre duas vogais reduzidas, /ə/ e /ɪ/, é *Rosa's roses*, utilizada nos trabalhos experimentais de Flemming e Johnson (2007). *Rosa* é uma palavra que já termina com o /ə/. Ao adicionar o 's para mostrar posse – *Rosa's* – o *schwa* fica intacto, ou seja, basta acrescentar o [z] do genitivo: [rozɚz]. Já, *roses*, plural de substantivo terminando em “e”, requer [ɪ]: [rozɪz]¹⁹.

Um aspecto relacionado às vogais reduzidas comumente encontrado em livros de pronúncia é a distinção entre o som da vogal no artigo definido *the*: [ðɪ] quando seguido de vogal (*the apple*) e [ðə] quando seguido de consoante [*the grape*]. Bolinger (1886, p. 357) afirma que essa distinção ocorre de modo a preservar o hiato, ou seja, o contraste entre as duas vogais sucessivas, daí *the apple* ser pronunciado [ðɪæpl]. O hiato pode ocorrer também dentro da mesma palavra, como em *geography* [dʒ'ɪɑgr əfi].

A vogal reduzida [ə] é menos comum do que [ə] e [ɪ]. Embora a posição com maior incidência desse som seja em posição pós-tônica (como em *sorrow*), ela pode também ocorrer em posição pré-tônica, como nas palavras *mortality*, *obese*, *donation*, *procede* (vs *precede* com [ə] ou [ɪ]) (BOLINGER, 1986, p. 357).

¹⁹ Bolinger (1986, p. 348) aponta haver variação dialetal na produção do [ɪ]. A região nordeste dos Estados Unidos, com exceção de Nova Iorque e New England, tende a uma vogal mais alta, próxima ao [i] enquanto o sul dos Estados Unidos privilegia uma posição mais baixa.

Diferentemente de Bolinger, Teschner e Whitler (2004), entre outros autores, agrupam todas as vogais reduzidas como [ə] e, portanto, centralizadas. Bolinger credita essa classificação aos inúmeros casos de [ə] alternando-se com [i] e [e] e à maior incidência de [ə]. A forma da palavra *police*, pode ilustrar tal afirmação:

[pə'lis] ----- > [pə'lis] ----- > [plis]

O fato de, em taxas de elocução mais rápidas, vogais reduzidas evoluírem para o [ə] explica porque muitos analistas veem a redução como um processo e não como um estado e, portanto, para eles, essas vogais não fazem parte de uma classe de vogais diferenciadas. Essa não é, entretanto, a opinião de Bolinger, posto que, conforme detalhei anteriormente, as três vogais reduzidas têm a capacidade de desambiguar palavras muito próximas como, por exemplo, as vogais [ə] e [i] nos pares *allegator/alligator* e *affect/effect*.

2.3.2 Ritmo silábico (*full vowel timing*)

Diferentemente de outros autores que caracterizam o ritmo silábico por intervalos regulares entre sílabas, para Bolinger (1986) o ritmo silábico manifesta-se na alternância entre sílabas plenas e reduzidas.

Para o autor, portanto, o ritmo silábico do inglês norte americano é determinado estruturalmente e independe das proeminências por acento de *pitch*.

The main difference between the two (syllabic and accentual rhythm) is that syllabic rhythm is structurally determined (by the arrangement of full and reduced syllables), whereas accentual rhythm is partly determined by, and must always adjust itself to, the accentual highlighting that the speaker imposes for the sake of his meaning (BOLINGER, 1986, p. 65).

Em outras palavras, no ritmo silábico, “there is a basic level of temporal patterning that is independent of tonal patterning” (BOLINGER, 1981, p. 24). Segue-se que o ritmo silábico é observável mesmo na fala produzida sem grandes variações de f_0 .

Nos parágrafos seguintes, apresentarei algumas características do ritmo silábico no inglês norte-americano, segundo a teoria rítmica proposta por Bolinger (1986).

2.3.2.1 ALONGAMENTO SILÁBICO

O efeito mais consistente relacionado à duração silábica no ritmo silábico ocorre entre sílabas adjacentes: quando uma sílaba plena é seguida de outra sílaba plena, a primeira recebe uma duração extra. Bolinger (1986) exemplifica esse alongamento intersilábico por meio de uma sentença extraída de um anúncio do sabonete Boraxo, publicado na revista *Life*, na edição de 13 de agosto de 1951, o qual reproduzo na Figura 11, a seguir:



Figura 11 – Anúncio do sabonete Boraxo

A sentença que Bolinger utilizou para exemplificar o ritmo silábico e seus efeitos prosódicos no inglês encontra-se na posição médio-baixa do enunciado (ver seta) onde se lê:

Gets out dirt plain soap can't reach

Ocorre, assim, alongamento nas sílabas em adjacência ao longo de toda a extensão do enunciado, uma vez que todos os monossílabos são compostos de sílabas plenas. O autor afirma que a razão para o alongamento silábico é provavelmente acentual: se o falante tem motivos para dar proeminência acentual a duas palavras seguidas – a eficácia do anúncio consiste em se fazer justamente isso – a duração extra fornece o espaço temporal para a inflexão de *pitch* (f_0). Em

outras palavras, há alongamento compensatório da vogal plena que antecede outra vogal plena para preservar o espaço temporal para a mudança de *pitch* (se houver). Bolinger acrescenta: “As with other parts of the sound system, we do not wait to find out whether we need it; instead, we provide it consistently in case we might need it” (BOLINGER, 1986, p. 39).

Dasher e Bolinger (1982) resumem esse fenômeno da seguinte forma: “a syllable containing a full vowel will be longer when followed (in the same phrase or breath group) by another syllable containing a full vowel than when followed by a syllable containing a reduced vowel” (DASHER e BOLINGER, 1982, p. 58-9).

Faber (1986) acredita que essa proposta de Dasher e Bolinger de alongamento da vogal de uma sílaba plena que precede outra sílaba plena é superior a que utiliza a noção de “*stress*” (como composto de sílabas fortes (s) e fracas (w)) como marcador das proeminências. Ele acrescenta que a proposta de Bolinger pode ser usada para explicar padrões rítmicos que as teorias de *stress-timing* não conseguem dar conta, como, por exemplo, porque a sílaba *car* é mais curta em:

Ask	Mr	Carter
tipo de sílaba		plena reduzida

do que em:

Ask	Mr	Cartwright
tipo de sílaba:		plena plena

Ou porque a sílaba *man* é mais curta em:

Have you	seen the	manor ?
tipo de sílaba:		plena reduzida

do que em:

Have you	seen the	manhole?
tipo de sílaba:		plena plena

Esses exemplos parecem evidenciar o fato de que, uma vez a duração silábica estando relacionada à característica vocálica da sílaba que a sucede, o ritmo do inglês apresenta forte componente estrutural. Bolinger (1986) dá outros dois exemplos:

The effect of full syllables remains the same: a full followed by another full has the predicted extra length. This happens with the *pon-* [pan] of *pontificate*, where *-tif-* is full, but not with the *pon-* of *pontiff*, where *-tiff* is reduced. It happens to *one* in *Ten times one makes ten* since it is followed by full *makes*, but not to *one* in *Ten times one is ten*, since *is* is reduced (BOLINGER, 1986, p. 41).

Bolinger não é o único a defender essa proposta para o ritmo do inglês motivado por duas sílabas plenas em adjacência. Gilbert (2001), por exemplo, afirma que há, no inglês, regras rítmicas ligadas à sílaba que fazem com que as sílabas tenham durações variáveis. A autora alerta para o efeito do alongamento mútuo de vogais plenas sucessivas no ritmo do inglês, citando como exemplo a sentença a seguir:

Give me your permission to **go right now**.

Os três últimos monossílabos da sentença acima, **go right now**, são compostos de vogais plenas [gou], [raɪt] e [naʊ] e, assim, há maior precisão na articulação das vogais e distensão na duração de cada palavra.

Embora indiretamente, Dauer (1993) também menciona essa característica rítmica do inglês, quando, em livro de pronúncia dirigido a aprendizes estrangeiros de inglês, na instrução de uma atividade para a prática do ritmo, pede ao aluno: “try to keep the same spacing between the “beats” (the stresses) *by lengthening the stressed syllables when there are not unstressed syllables between them*” (DAUER, 1993, p. 87).

Em suma, Bolinger acredita que “vowel quality is a given and not a product of stressing rules or rhythm rules, and can be used to predict the possibility of accentuation” (DASHER e BOLINGER, 1982, p. 60) e que “the lengthening keyed to the fullness of the vowel is part of the underlying system and occurs automatically” (BOLINGER, 1986, p. 45).

Faz-se importante analisar como a duração da sílaba plena seria afetada quando seguida de uma sílaba reduzida. É sobre isso que falaremos no próximo item.

2.3.2.2 THE BORROWING RULE

Conforme discutido na seção anterior, quando uma sílaba plena fosse seguida de outra sílaba plena, a primeira receberia *extra length*. Por outro lado, quando uma sílaba plena fosse seguida por uma sílaba reduzida, o alongamento da vogal não mais se faria necessário, uma vez que a sílaba reduzida já forneceria o espaço

temporal para a inflexão tonal. Assim, as sílabas reduzidas estariam “emprestando tempo” das sílabas plenas. Como resultado, a combinação formada por uma sílaba plena e uma sílaba reduzida (nesta ordem) seria produzida aproximadamente no mesmo espaço de tempo da sílaba plena, como exemplificado a seguir:

Kids want sweets

[kɪdz wʌnt swits]

Kids wanted sweets

[kɪdz wʌntɪd swits]

No exemplo, a vogal [ʌ] da palavra *want* no primeiro enunciado seria mais longa do que no segundo enunciado em *wanted*, já que ambas palavras tenderiam a ser produzidas em tempos aproximados, uma vez que a duração dos pés rítmicos troqueus²⁰ (*wanted*) e *monobeats* (*want*) são muito próximas.

Dados experimentais mostram que, a medida que aumentamos o número de sílabas reduzidas dentro de um pé rítmico, ele tende a aumentar em duração. Kim e Cole (2007) desenvolveram um experimento para o inglês norte-americano no qual foram feitas as medições de duração de pés rítmicos de tamanhos variados em três posições: dentro de frases intermediárias, em fronteiras intermediárias e em fronteiras finais, os quais foram plotados no Gráfico 4 abaixo:

²⁰ Pé rítmico formado por uma sílaba plena e uma reduzida.

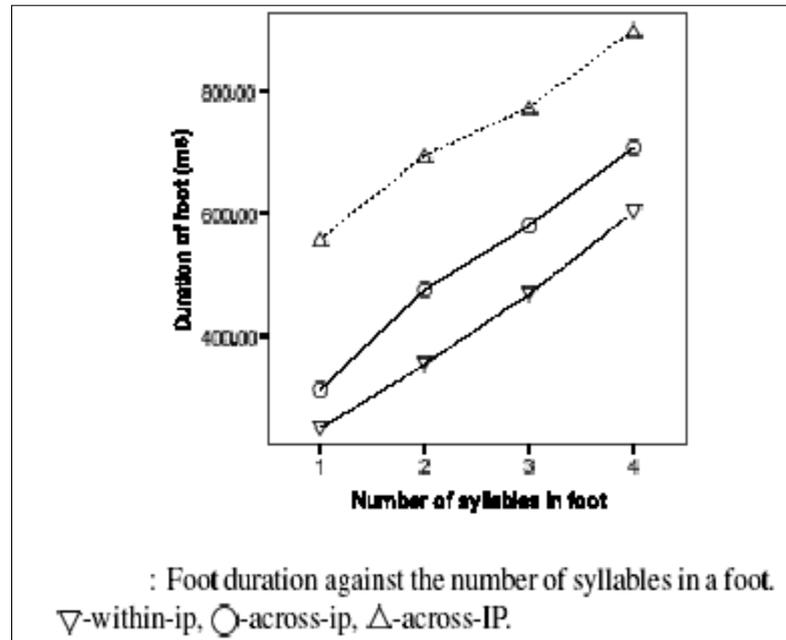


Gráfico 4 – Duração do pé rítmico vs número de sílabas no pé. Fonte: Kim e Cole (2007)

Kim e Cole (2007) concluíram que, quando o número de sílabas em um pé rítmico aumenta, a duração do pé aumenta, mas a duração da rima silábica da sílaba acentuada é reduzida. Os autores acrescentam que “this finding suggests that rhythmic regularization in American English is manifested not in terms of isochronous foot duration, but through an adjustment of the duration of stressed syllables”. Uma vez que esse encurtamento da vogal ocorre apenas dentro da frase intermediária, os resultados do experimento sugerem que a estrutura prosódica desempenha um papel importante no inglês norte-americano.

Tal implementação fonológica pode ser explicada de forma mais simplificada através da chamada *Borrowing Rule*, ou regra do empréstimo: a sílaba com a vogal reduzida [tɪd] estaria “tomando tempo emprestado” da sílaba longa que a precede [wʌn]. Em outras palavras, quando a sílaba plena é seguida de sílabas reduzidas,

ela teria sua duração reduzida, de modo a manter a regularidade, conforme, portanto, mostraram os dados de Kim e Cole (2007).

Bolinger (1986) também acredita na existência da *Borrowing Rule* para o inglês norte-americano. Segundo o autor, as características rítmicas relacionadas à sílaba são observadas na alternância entre sílabas plenas e reduzidas e ao modo como as vogais plenas têm sua duração alterada quando sílabas reduzidas são adicionadas ao mesmo pé rítmico. Não se trata, portanto, de *isossilabicionismo*, tendência de ser dar a cada sílaba, independentemente de ser proeminente ou não, a mesma duração²¹.

Isso já seria em si suficiente para que se traçasse base para distinção tipológica entre o ritmo de duas línguas. Por exemplo, seria possível testar a hipótese de que as línguas classificadas como de ritmo acentual apresentam maior contraste na duração de vogais em sílabas adjacentes do que as sílabas classificadas como de ritmo silábico.

Há vários trabalhos acústicos contrastivos entre o ritmo das línguas envolvendo a inspeção da qualidade vocálica e a duração silábica que parecem corroborar a *Borrowing Rule*.

Udofot (2003), por exemplo, desenvolveu experimento acústico de análise do ritmo da fala de 60 sujeitos nigerianos oriundos de três categorias distintas: *substandard*, *standard* e *sophisticated English*, contrastando as produções desses falantes com a de um falante nativo de inglês britânico. Ele percebeu que nigerianos tendem a dar a mesma proeminência a todas as sílabas, sem diferenciar vogais plenas e vogais

²¹ O termo isossilabicionismo é recorrente em Barbosa (2006), em estudo sobre as características rítmicas do português brasileiro sob um enfoque dinâmico-funcional.

reduzidas. Como resultado, a duração, tanto das sílabas quanto dos enunciados, mostra-se maior na fala dos falantes nigerianos.

O pesquisador percebeu haver também diferenças em relação ao maior número de sílabas proeminentes na fala dos nigerianos, sendo que o padrão rítmico desses falantes não alterna sílabas proeminentes e não proeminentes de forma consistente. A análise acústica demonstrou que o grupo dos nigerianos usou mais tempo para produzir as sílabas reduzidas, palavras e enunciados e menos tempo para produzir *accented syllables* do que o falante britânico, resultando em uma tendência de valores de duração de sílabas mais aproximados.

O autor afirma que talvez uma análise de ritmo baseada na descrição do intervalo de ocorrência das vogais plenas (*full vowel timing*) seja mais compatível com o inglês nigeriano, já que essa análise não depende nem do número de acentos de *pitch* nem do número de sílabas, mas do padrão formado pela mistura de sílabas plenas e reduzidas. Em outras palavras, o veio estrutural do enfoque proposto por Bolinger para o ritmo do inglês, ou seja, que as sílabas reduzidas “borrow time from the full syllable” mostrou-se mais produtora na análise das diferenças entre os padrões rítmicos do inglês nigeriano e do inglês britânico do que aquele baseado apenas nos “*stresses*”. O autor concluiu que o inglês nigeriano apresenta características rítmicas mais próximas do ritmo silábico do que do ritmo acentual.

Rietveld e Koopmans-van Beinum (1987) investigaram a relação entre o grau de redução vocálica e a percepção de proeminência acentual quando outros parâmetros, tais como *f0*, intensidade e duração, foram mantidos constantes; eles descobriram que, nesse contexto, a ausência de vogais reduzidas era uma pista relevante para a percepção das proeminências.

Householder (1971), por sua vez, em estudo de *corpus* de inglês coloquial, descobriu que a diferença na posição do acento lexical na grande maioria de pares de palavras com mesma grafia (ex: 'record (n) vs re'cord (v)) era sinalizada pela redução vocálica (portanto, manipulação de qualidade de voz), bem como por diferenças acústicas nos parâmetros de f_0 , intensidade e duração. De modo similar, Fear, Cutler e Butterfield (1995) afirmam que as sílabas “fortes” e “fracas” no inglês são distinguidas pelo “*stress*” (ex: f_0 , duração e intensidade), pela qualidade vocálica ou por ambos.

2.3.2.3 PÉS RÍTMICOS SILÁBICOS

Um pé rítmico silábico é, segundo Bolinger, qualquer combinação de sílabas plenas que podem vir ou não seguidas de sílabas reduzidas. Assim, todas as palavras em inglês são passíveis de serem analisadas em termos de seus pés rítmicos. Dentre os pés rítmicos silábicos, podemos destacar três tipos: (1) espondeus (duas sílabas plenas, como em *Utah* [juta]), (2) troqueus (sílabas plena seguida de sílabas reduzida, como em *morning* ['mɔ:rnɪŋ]) e (3) dátilos (sílabas plena seguida de duas sílabas reduzidas, como em *mountainous* ['maʊntənəs]).

A fala encadeada apresenta características rítmicas similares às palavras polissilábicas. Bolinger (1986) afirma que a combinação e intercalação de sílabas plenas e reduzidas exerce papel preponderante no ritmo, não apenas no nível lexical como também no nível frasal. Existe uma estrutura de pés rítmicos silábicos, ou seja,

uma estrutura rítmica composta de sílabas plenas que podem ou não vir seguidas de sílabas reduzidas.

Segundo Bolinger (1981), “syllabic rhythm is based on duration, not on intensity, nor on recurring accents.” A regularidade existe em função da alternância entre sílabas plenas (*long*) e sílabas reduzidas (*short*).

De modo a exemplificar esse típico de ritmo no inglês norte-americano, utilizarei o enunciado analisado por Mauad (em preparação), extraído da série norte-americana *Desperate Housewives*, na Figura 12, no qual o personagem Bob (Tuc Watkins) convida Bree (Marcia Cross) para uma festa de Halloween.

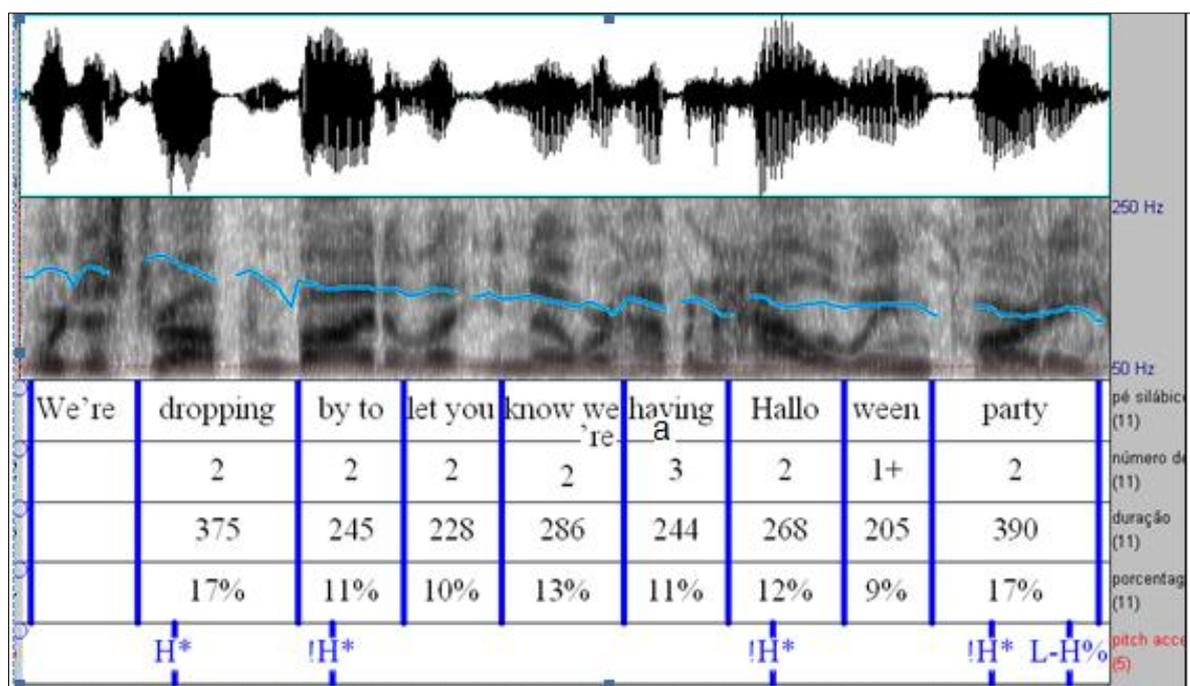


Figura 12 – Forma da onda, contorno entoacional, pés silábicos, número de sílabas por pé, duração do pé silábico, porcentagem da duração do pé em relação ao todo e *pitch accents* do enunciado *We're just dropping by to let you know we're having a Halloween party*

Na Figura 12, observa-se que apenas 4 vogais plenas recebem acentos de *pitch* (tons H*, !H*, !H* e !H*) e que os pés silábicos, excetuando-se o inicial e o final,

apresentam valores relativos (percentuais) de duração muito próximos (11%, 10%, 13%, 11% e 12%). De acordo com Bolinger (1986), o ritmo seria, assim, implementado pelos pés silábicos, definidos como a combinação de uma sílaba plena e demais sílabas reduzidas (o símbolo “+” abaixo da sílaba “*ween*” sinaliza que esta sílaba receberia um alongamento extra por preceder outra sílaba plena “*par*”).

Há dois pontos sobre o ritmo silábico do inglês que gostaria de destacar. O primeiro é que o local onde é feita a separação dos pés rítmicos é indiferente, pode vir no meio da palavra, como em *Halloween* (2 pés separados por 2 vogais plenas), bem como entre palavras. O segundo é que a contagem de sílaba dos pés rítmicos é raramente regular. Na sentença em questão, temos, por exemplo, 8 pés rítmicos silábicos, sendo que a contagem de sílabas nos pés rítmicos consecutivos é: 2-2-2-2-3-2-1-2, uma sequência irregular, portanto.

Entretanto, é possível construir-se pés rítmicos regulares, os quais podem veicular sentidos de ordem pragmática. A seguir apresento três exemplos, extraídos de Bolinger (1986), de sequências de pés rítmicos de igual tamanho e duração (isocrônicos) inseridos em enunciados e os efeitos de sentidos que esses padrões de ritmo silábico podem causar no ouvinte nativo de inglês.

(1) Sequência de monossílabos com sílabas com vogais plenas

A sequência de monossílabos com sílabas plenas é rara no inglês. Bolinger (1986) afirma que enunciados com ritmo silábico regular composto de monossílabos podem causar estranheza no ouvinte e dar a impressão de terem sido formados previamente ao momento de fala. Dauer (1993, p. 103), indiretamente, parece corroborar essa

percepção de Bolinger quando afirma que “good writers and speakers try to avoid a sequence of ‘stressed’ syllables like this, by changing the grammar and vocabulary.”

Exemplo: **Pat / made / John / tell / who / fired / those/ guns.**

(2) Sequência de troqueus silábicos (sílabas com vogal plena seguida de sílabas reduzidas)

A sequência de troqueus silábicos soa mais natural por conta da sílaba reduzida, intercalando cada proeminência, mas ainda assim é notada pelo ouvinte. Vale ressaltar que, embora haja ausência de *pitch accents* nas sílabas proeminentes, há regularidade, nitidamente marcada pela intercalação de sílabas plenas e reduzidas.

Exemplo: I **had** to / **make** an / **effort** / **just** to / **get** a / **breath** of / **air**.

(3) Sequência de dátilos silábicos (sílabas com vogal plena seguida de duas sílabas reduzidas)

Segundo Bolinger (1986), no caso de dátilos silábicos sucessivos, o ritmo silábico regular deixa de soar estranho aos ouvintes nativos do inglês. Dátilos silábicos também favorecem o “encaixe” entre os pés acentuais e silábicos, na medida em que as duas sílabas reduzidas fornecem o espaço temporal para a inflexão de *pitch* (se houver). Exemplo: He / **wanted** to / **give** them a / **piece** of his / **mind**.

Portanto, pelo exposto acima, conclui-se que o ouvinte nativo tem, mesmo que de forma inconsciente, ciência das combinações de *padrões* rítmicos que soam naturais em sua língua e percebe quando um enunciado foi artificialmente modificado.

Embora seu foco fosse o texto escrito, Lorant (1920), há quase um século, já discutia a percepção do ritmo e a impressão negativa que um texto deliberadamente ritmado pode suscitar no leitor:

Now, why do we become uneasy the moment we suspect a writer of aiming at musical effects? It is because we know instinctively that every thought creates its own rhythm, and that when a writer's attention is upon his rhythm, he is bent upon something else than his thought processes. The only way of giving the impression of thought that is not original or spontaneous is by imitating the rhythm of that thought. For real meanings cannot be borrowed. They are always new. Real thought is an action, an original adventure. It pulsates, and the body pulsates with it. No writer can produce this sense of original adventure in us unless he has it himself (LORANT, 1920, p. 2).

Sobre esse aspecto do ritmo da fala, Bolinger (1981, p. 33) acrescenta que:

There's no a great deal one can do intentionally in the heat of putting a sentence together to pick words that will make the metrical foot more or less even, but what happens **unintentionally** testifies to some kind of underlying rhythmic tendency.

Tendo exposto as características do ritmo silábico do inglês norte-americano, passo agora ao ritmo acentual.

2.3.3 Ritmo acentual (*pitch accent timing*)

Como a fala, dada sua expressividade, raramente é produzida sem variação de f_0 , Bolinger concebe outra camada (ou nível) de ritmo: o ritmo acentual. Ele envolve relações temporais entre *acentos* marcados por variações de f_0 , os *pitch accents*, e a acomodação dos mecanismos da fala.

De acordo com Bolinger, a principal diferença entre o ritmo silábico e o ritmo acentual é que, enquanto o ritmo silábico é determinado estruturalmente (pela

combinação de sílabas plenas e reduzidas), o ritmo acentual existe em função das proeminências impostas pelo falante em prol do significado que ele quer expressar (BOLINGER, 1986).

No ritmo acentual, as proeminências são carregadas pelo acento, marcado principalmente pela inflexão de *pitch* – daí o termo *pitch accent*.

Antes de continuarmos nossa discussão sobre o ritmo acentual, farei, nos itens 2.3.3.1 e 2.3.3.2, breve explicação sobre dois termos recorrentes nos estudos do ritmo de inglês, *stress* e (*pitch*) *accent*, explicando também a distinção feita por Bolinger para esses termos.

2.3.3.1 STRESS

Já em 1949, Trubetzkoy apontava a relevância da sílaba na prosódia da fala, ao afirmar que ela fornece base para as características prosódicas distintivas: “Les particularités prosodiques n’appartiennent pas aux voyelles en tant que telles mais aux syllabes”²² (TRUBETZKOY, 1949, p. 196). Considerada por muitos a unidade rítmica da fala, a sílaba desempenha papel importante na cognição, uma vez que parece facilitar o processamento do fluxo de fala.

Em função de sua proeminência nas palavras, as sílabas podem ser classificadas em tônica e átona. A sílaba tônica é composta de uma vogal plena que carrega o acento lexical e, portanto, recebe a maior proeminência: é mais

²² As particularidades prosódicas não pertencem às vogais como tal, mas às sílabas.

longa, tem maior intensidade e é produzida em um tom mais alto do que as demais. Na palavra *apartment*, por exemplo, *-part* é a sílaba tônica e *-a* e *-ment* são sílabas átonas.

De acordo com Laver (2002), a manifestação fonética da sílaba tônica varia de língua para língua. O inglês especificamente explora os quatro parâmetros acústicos: *pitch*, intensidade, duração e qualidade. Entretanto, segundo o autor, a maioria das outras línguas parece utilizar-se de apenas três desses parâmetros: *pitch*, intensidade e duração. Segundo Lehiste (1970), há, na fala encadeada, forte correlação entre a sílaba tônica e os pulsos elétricos gerados pelas contrações dos músculos respiratórios.

Há palavras nas quais mais de uma sílaba recebe proeminência. Nesse caso, a sílaba tônica recebe a maior proeminência: o acento primário (*primary stress*). Uma outra sílaba também composta de uma vogal plena (*long syllable*), a sílaba subtônica, recebe a segunda maior proeminência na palavra: o acento secundário (*secondary stress*). As demais sílabas, compostas de vogais reduzidas (*short syllables*), não são proeminentes (são *stressless*). Por exemplo, na palavra **generation** (gen - e- ra- tion), *ra* é sílaba tônica (acento primário); *gen-* é sílaba subtônica (acento secundário); e *-e* e *-tion* são sílabas átonas (sílabas reduzidas, i.e. isentas de acento).

Em Bolinger (1964), os termos *long syllable* e *short syllable*, são, por vezes, utilizados para referir-se à sílaba com vogal plena e à sílaba com vogal reduzida, respectivamente. Uma vez que os adjetivos *long* e *short*, quando aplicados à sílaba, podem remeter à sua duração e neste trabalho serão feitas medições de duração silábica, optei por evitá-los ou colocá-los entre

parênteses. Assim, os termos que utilizarei com mais frequência serão: vogal plena, vogal reduzida, sílaba com (ou composta de) vogal plena e sílaba com vogal reduzida.

Segundo Dauer (1993) há, no inglês, tendência, no nível lexical, de se alternar vogais plenas e vogais reduzidas em palavras polissilábicas:

pho to graph	pho to gra phy	pho to graph ic
['fou tə græf]	[fə 'tə grə fi]	[,fo tə 'græf ik]
plena reduzida plena	reduzida plena reduzida plena	plena reduzida plena reduzida

Nas palavras *photograph* e *photography*, o acento tônico lexical recai, respectivamente, sobre as sílabas *pho-* e *to-*. Essas palavras não têm acento secundário, uma vez que não há outra sílaba com vogal plena antes da sílaba tônica.

A palavra *photographic*, por sua vez, tem o acento primário em *-graph* e o secundário em *pho-*. É a única palavra em que o acento secundário pode, dependendo do contexto, receber mais proeminência do que o acento primário, pois há nessa palavra uma sílaba com vogal plena (*pho-*) anterior à sílaba tônica (*graph*).

Conforme afirma Bolinger (1986, p. 60) “It’s not a matter of shifting stress, since the stress as a potential for accent, is already there; it is rather the choice of a secondary over a primary in a situation where intonational advantage accrues”. Essa

possibilidade de acento secundário será abordada futuramente quando iniciarmos a discussão sobre o ritmo acentual.

O inglês exibe algumas tendências em relação ao *stress* (acento lexical), tais como:

(a) nos substantivos e adjetivos, a sílaba tônica (*the stressed syllable*) *tende* a ser inicial: *accident*, *happy*, *comfortable*, *television*, etc.

(b) nos verbos dissílabos, a sílaba tônica *tende* a vir no final das palavras, favorecendo, portanto, as oxítonas: *complete*, *invade*, *arrive*, *remain* etc.

No entanto, há várias exceções à essa regra: *frustrate*, *lengthen*, *combat* etc.

Há, contudo, um elemento de previsibilidade nas proeminências que pode auxiliar no estudo do ritmo do inglês: a sílaba tônica (*the stressed syllable*), ou seja, aquela com o potencial de receber o acento (de *pitch*), não será uma sílaba átona, ou seja, uma sílaba que contém uma vogal reduzida. Por exemplo, na palavra *fanfare*, as duas sílabas são compostas de vogais plenas, conseqüentemente, tanto *fan-* como *-fare* poderiam receber o acento tônico (também chamado de *primary stress*). A segunda sílaba receberia o *secondary stress*.

Por outro lado, na palavra *fancy*, a sílaba *-cy* contém uma vogal reduzida, caracterizada pela diminuição de duração e pela incerteza do timbre. Portanto, *-cy* não pode receber *stress*. Seguindo essa linha de raciocínio, conclui-se que a sílaba tônica (*the stressed syllable*) é facilmente identificável na maioria dos contextos. Por exemplo, na palavra, ***formidableness***, a primeira sílaba, *for-*, é *stressed*, já que todas as outras sílabas são compostas de vogais reduzidas.

2.3.3.2 *PITCH ACCENT*

A proeminência na fala encadeada difere da proeminência em palavras isoladas. Na fala encadeada, o *accent* é determinado em função da intenção comunicativa do falante. Por exemplo, no diálogo:

MARY: Did you move back to your parent's house?

PETER: **No**. I've **bought** an **apartment**.

As sílabas em negrito recebem proeminência frasal. Na palavra *apartment*, o *pitch accent* recai justamente sobre a sílaba tônica (a *stressed syllable*), posto que a sílaba tônica tem o potencial para o *acento* (BOLINGER, 1986, p. 15). Se fizermos uma interpretação gestáltica do *pitch accent*, as sílabas acentuadas representam a Figura (*figure*) e as demais o fundo (*ground*).

Pitch accents podem ser definidos como “configurações tonais que descrevem o movimento de *f0* específico associado com sílabas acentuadas dentro da frase entoacional” (ROCCA, 2003, p. 92). Os padrões de proeminência (por *pitch accent*) na fala encadeada são distribuídos mais livremente do que os padrões de proeminência na palavra (o acento lexical ou *stress*).

Segundo Ladd (1996), um *pitch accent* é “a local feature of pitch contour – usually but not invariably a *pitch change*, and often involving a local maximum or minimum – which signals that the syllable with which it is associated is prominent” (LADD, 1996, p. 46).

Barbosa (2006, p. 166) também vê no *pitch*, ou seja, na frequência fundamental (*f0*) um fator essencial na marcação dos *accents* em inglês. O autor afirma que “em

virtude do controle sistemático e frequência da f_0 para assinalar proeminência através da produção de *pitch accents*, é mais fácil a detecção das sílabas proeminentes em línguas como inglês e sueco do que no português, pelo menos para a frase lida”.

É importante ressaltar que próprio termo *pitch accent* traz em si uma relação entre entoação (*pitch*) e ritmo (*accent*), relação esta evidenciada na declaração de Fletcher e Harrington (1999, p. 173): “Words in a languages like English can have an intonational event (*pitch accent*) associated with the rhythmically strongest syllables that makes the word accentually prominent relative to other words in a phrase.”

Os termos *stress* e *accent* correspondem, portanto, respectivamente, às proeminências nas sílabas tônicas de palavras (acento lexical) e nas sílabas tônicas das palavras relevantes na frase (acento de *pitch*). Citando Bolinger (1998, p. 47): “Accents are abrupt changes in pitch usually combined with increases in length and loudness, imposed, as a rule, on the ‘stressed’ syllable of a word that is to be highlighted.”

Embora muitos linguistas usem os dois termos, *stress* e *accent*, indistintamente, neste trabalho prevalece a visão de Bolinger para *stress* como sendo “an abstract lexical property of individual syllables”, e (*pitch*) *accent* como “actual prominence in an utterance” (LADD, 1996, p. 46).

Esse sistema é resumido por Bolinger (1964, p. 22) na Tabela 4:

sílaba com <i>pitch accent</i> (<i>accented syllable</i>)	sílaba sem <i>pitch accent</i> (<i>unaccented syllable</i>)	Qualquer sílaba tônica pode receber o <i>pitch accent</i> . Quais delas receberão essa proeminência depende da intenção comunicativa do falante
sílaba tônica (<i>stressed syllable</i>)	sílaba átona (<i>unstressed syllable</i>)	Sílabas que contêm vogais plenas podem receber <i>stress</i> (acento lexical); apenas uma sílaba da palavra recebe o <i>primary stress</i> : a sílaba tônica
sílaba composta de vogal plena (<i>long syllable</i>)	sílaba composta de vogal reduzida (<i>short syllable</i>)	<i>Long syllables</i> contêm vogais plenas; <i>short syllables</i> contêm vogais reduzidas

Tabela 4 – A relação *stress* e *accent* na marcação de proeminências em inglês.
Fonte: Bolinger (1964, p. 22)

Quando um falante destaca uma sílaba por meio de *pitch*, significa que ela é importante na expressão do significado de seu enunciado. Como imaginamos um significado para tudo o que ouvimos, mesmo a fala sem inflexões tonais pode ser *percebida* como marcada por acentos de *pitch*.

Conforme afirma Classe (1939, p. 18), “It is quite likely that I shall feel a stress [=accent] in the very place at which I would have put one myself, even if my interlocutor has spoken in a perfectly even manner”.

A afirmação de Classe (1939) vai ao encontro dos estudos de Allen (1978), que demonstrou que, quando ouvintes foram expostos a uma série de ruídos absolutamente idênticos e de igual intervalo, eles percebiam um desses *clicks* como sendo mais proeminente do que outros, embora não houvesse nenhum dado acústico/físico presente que corroborasse essa percepção.

Entretanto, embora o *pitch accent* tenha realmente um componente perceptual na medida em que o ouvinte percebe e identifica as sílabas acentuadas das palavras proeminentes na frase entoacional, creio que a análise acústica na identificação desses alvos tonais é extremamente valiosa, uma vez que, por meio dos dados acústicos sobre o contorno entoacional, ou mesmo sobre outros elementos prosódicos, tais como a duração e a qualidade vocal, o transcritor pode corroborar essa percepção, integrando, assim, o produzido ao percebido.

2.3.3.3 CONTENT WORDS VS FUNCTION WORDS

Nem todas as palavras possuem uma sílaba proeminente capaz de receber o *pitch accent*. No nível frasal, a sílaba proeminente geralmente recai sobre a sílaba tônica da palavra de maior relevância semântica, ou seja, a palavra com maior conteúdo informativo. Essas palavras são denominadas palavras de conteúdo (*content words*), dentre as quais estão: (a) substantivos, (b) adjetivos, (c) verbos ((principais (*work, live, etc*), auxiliares negativos (*don't, can't, aren't, etc*), auxiliares não seguidos de verbo principal (*Yes, I do, I think he can, etc*)), (d) advérbios, (e) numerais, (f) *question words* (*what, how, why, etc*) e pronomes demonstrativos (*that, these etc*).

Por outro lado, as palavras que desempenham função gramatical e cujo sentido depende da relação que elas estabelecem com outras palavras geralmente não recebem proeminência. Essas palavras são denominadas palavras de função gramatical (*function words*), dentre as quais encontramos: (1) artigos, (2) pronomes, (3) preposições, (4) conjunções e (5) verbos auxiliares afirmativos quando seguidos

de verbo principal (ex: *has* em *Joe has bought the car*) e *be* como verbo de ligação seguido de um adjetivo ou de um sintagma nominal (ex: *is* em *The cat is white*) ou de um sintagma nominal (ex: *was* em *It was a group of children*).

2.3.3.4 OUTRAS PISTAS PARA O ACCENT

Além da inflexão de *pitch*, há outras formas de se destacar uma sílaba com proeminência acentual, dentre as quais: aumento de duração, maior intensidade, *delayed release* (liberação tardia) da consoante e voz soprosa (*breathy voice*). Listarei brevemente duas dessas pistas acústicas para o *accent*: *breathiness* e intensidade.

2.3.3.4.1 *Breathiness*

Segundo Bolinger (1986, p. 19), *breathiness* é um mecanismo curioso utilizado no discurso oral para marcar proeminência, pois, ao invés de dar mais intensidade à palavra saliente, o falante a produz com menos intensidade. Ele parece querer transmitir um grau de surpresa tão grande que o faz calar, reduzindo a intensidade da sílaba acentuada para quase um sussurro.

Como lembra Camargo (2002), a voz modal é caracterizada pelo ajuste neutro em relação à condição das pregas vogais nos aspectos de adução, tensão e características de vibração. *Breathiness* é uma variação do ajuste modal,

resultando em modo de vibração insuficiente e acompanhado por leve e discreta fricção audível.

Tal vibração insuficiente pode ser observada no espectrograma da Figura 13, a seguir:

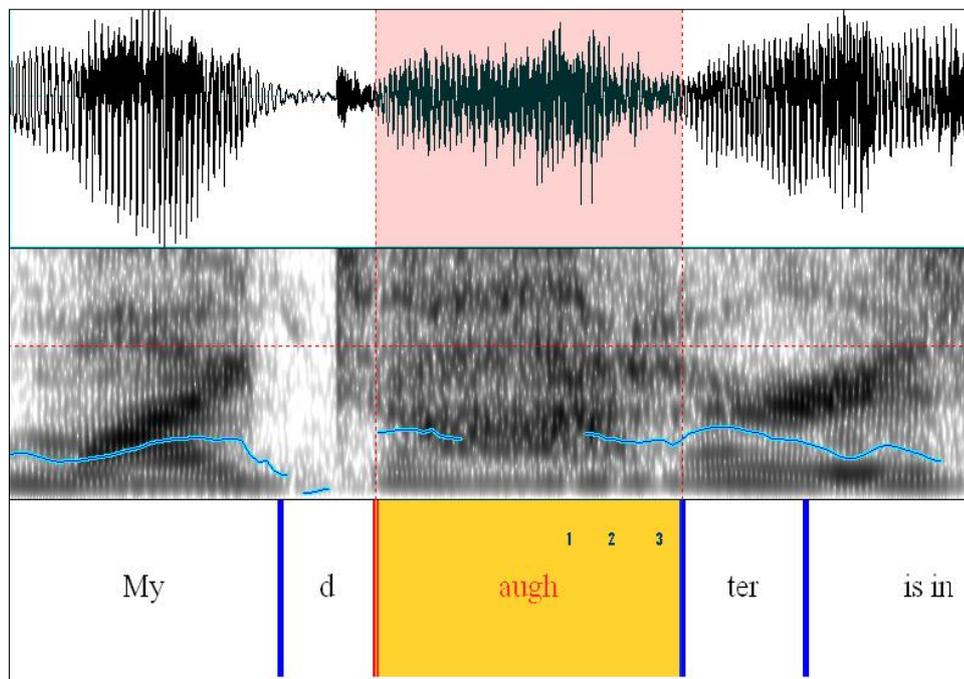


Figura 13 – Forma da onda, curva entoacional, espectrograma de banda larga e segmentação com transcrição ortográfica da palavra *daughter* no enunciado *My daughter is in Switzerland*, ilustrando episódios de *breathiness* como marcador de proeminência acentual

Pode-se observar que a consoante vozeada plosiva /d/ foi quase que totalmente desvozeada, como nota-se pela ausência de barra de vozeamento.

A vogal /ɔ/ correspondente ao grafema *augh* também apresenta episódios de *breathiness*. Segundo Crystal (1985), “a breathy vowel is one produced by allowing a great deal of air to pass through a slightly open glottis, the aperture between the vocal cords”.

Se atentarmos ao espectrograma na Figura 13, veremos que, acima dos números 1, 2 e 3, encontram-se três espaços mais claros, o que mostra vibração insuficiente das cordas vocais. Além disso, as estrias escuras verticais irregulares na parte superior do espectrograma demonstram haver fricção.

Em consonância com sua visão icônica da entoação e, por extensão, aos outros elementos prosódicos como o ritmo, Bolinger (1986) afirma que, quando uma sílaba recebe proeminência por meio de *breathiness* em situações em que o falante está surpreso, a fala é acompanhada de um gesto facial específico: “eyes wide open and jaw slack and sagging”.

As duas fotos abaixo, na Figura 14, analisadas em Mauad (em preparação) foram tiradas no instante de produção da sílaba *-daugh* em intervalo aproximado de 1 segundo. Os olhos arregalados e a posição rebaixada da mandíbula corroboram as afirmações de Bolinger.



Figura 14 – Duas fotos com intervalo de 1 segundo, ilustrando a qualidade de voz de *breathiness* que acompanha a atitude de surpresa. Fonte: série *Desperate Housewives*

2.3.3.4.2 Intensidade

Outra forma de se marcar proeminência ocorre por meio de variações de intensidade. A intensidade das vogais está correlacionada com o grau de abertura da mandíbula. Em estudo sobre o papel da intensidade na marcação das proeminências, Ladefoged (2003) apresentou três enunciados: (1) *There were two cars*, (2) *There were three cars*, e (3) *There were four cars*. Todas as palavras foram produzidas com o mesmo grau de proeminência. O autor observou que as palavras *three* e *two*, compostas de vogais altas, têm valores cerca de 6 dB menor de intensidade do que *four*, com vogal baixa, na qual a mandíbula está abaixada.

Esse é um indício de que a intensidade não é um bom parâmetro acústico para a marcação das proeminências. Ladefoged observou que o mesmo ocorre em contextos onde há contraste, como em *I see three bees (not a swarm of them)*, *I see three bees (not three wasps)* e *I see three bees (but I can't hear them)*. Ele verificou que a intensidade das palavras proeminentes não foi maior do que as não proeminentes e alerta o leitor para o fato ao afirmar que “despite what you may read elsewhere, intensity as shown in dB is usually not a very useful acoustic property to measure. It is seldom one of the phonetic characteristics of a language” (LADEFOGED, 2003, p. 91)²³.

²³ Madureira (comunicação pessoal) aponta que até o advento da espectrografia da fala, a intensidade era considerada o parâmetro mais relevante na marcação das proeminências acentuais.

2.3.3.5 REGRA RÍTMICA E COLISÃO ACENTUAL (*STRESS CLASH*)

Segundo a regra de alternância rítmica (*The Rhythm Rule*), haveria na fala tendência de se alternar sílabas mais proeminentes e menos proeminentes motivada por forças rítmicas que procurariam manter essas proeminências em distância confortável entre si. Assim, a característica fundamental do ritmo acentual seria o espaçamento dos *pitch accents*, de modo a evitar o encontro de *pitch accents* sucessivos.

Dessa forma, proeminências que se encontram muito próximas estariam sujeitas a mecanismos linguísticos, atuando no sentido de evitar esse choque ou colisão acentual (*stress clash*). Por exemplo, na sequência *thirteen men*, segundo Liberman e Prince (1977), pode ocorrer colisão acentual entre as sílabas *teen* e *men*, e, para evitá-la, a proeminência seria deslocada para a esquerda da primeira sílaba acentuada (*teen*), portanto *thirteen men*.

Madureira (2002) investigou, do ponto de vista fonético-acústico, a ocorrência de colisão acentual para o português brasileiro e não encontrou evidências a favor desse fenômeno. A sentença analisada foi: “Não poderia tomar **café frio** de novo” (a colisão acentual encontra-se negritada). O fato mais relevante de sua investigação foi a presença de um movimento de *f0* ascendente na sílaba “fé”, mas que se estende até a palavra *frio*, onde ocorre o pico de *f0* – na verdade, para onde é deslocado o acento de *pitch* principal.

É importante ressaltar a ocorrência de um deslocamento para a direita do pico de *f0*, característico do padrão entoacional do português. Ao analisar resultados

de trabalhos em língua inglesa sobre o *stress clash*, Madureira aponta o predomínio, nessa língua, de movimentos à esquerda, tais como os adjetivos que precedem os substantivos (nível sintático) e os processos de assimilação progressiva no inglês (nível fonológico)²⁴.

No inglês, há alguns estudos perceptivos em favor de um princípio de alternância rítmica em experimentos que mostram que falantes do inglês ajustam os padrões de proeminência de modo a torná-los mais regulares. Por exemplo, Kelly e Bock (1988) realizaram experimento com 16 sujeitos norte-americanos, membros da comunidade acadêmica da Cornell University.

As autoras criaram 48 pseudo-palavras (*pseudowords*), ou seja, termos em conformidade com as regras fonológicas do inglês, mas que não existem no vocabulário do idioma. Para garantir que todos esses logotomas fossem pronunciados como palavras dissílabas, todos tinham uma consoante medial flanqueada por vogais.

Trinta e dois pares de frases-veículos compostas de 6 sílabas foram construídas, e um logotoma com função de substantivo foi inserido aleatoriamente a cada par de frases. Em uma das sentenças, o pseudo-substantivo ocupava um contexto rítmico que favoreceria a ocorrência de troqueus acentuais (sílabas acentuadas seguidas de sílabas não acentuadas) e, na segunda, de iâmbicos (sílabas não acentuadas seguidas de sílabas acentuadas).

²⁴ Por assimilação, entende-se a modificação de um som por influência do som vizinho que com ele passa a partilhar traços articulatórios. No inglês, a assimilação progressiva de um som pode ser verificada por influência do som anterior como a pronúncia da terminação -s para o plural: /z/ antes de sons vozeados, como em *eggs*, e /s/ antes dos sons desvozeados, como em *asks*. No português, por sua vez, a assimilação é regressiva. A assimilação regressiva no português pode ser verificada no vozeamento do -s, quando seguido de um som vozeado, como na frase *olhos azuis*.

As sentenças condicionadas ao contexto substantivo/troqueu apresentavam a estrutura *verbo – artigo – substantivo – advérbio*, e as sentenças condicionadas ao contexto substantivo/iâmbico tinham a estrutura *artigo – adjetivo – substantivo – verbo*, conforme exemplificado a seguir:

- (1) Save the pernew quickly (contexto sustantivo/troqueu)
- (2) The big pernew escaped (contexto substantivo/iâmbico)

Em seguida, os 48 logatomas foram extraídos das frases-veículo e gravados, um a um, em arquivo sonoro. Cada logatoma assim gerado foi submetido à análise perceptiva de dois juízes norte-americanos, estudantes de graduação da mesma universidade sem conhecimento das condições e objetivos do experimento, para que escolhessem a sílaba mais proeminente em cada logatoma dissílabo.

O dado perceptivo mais importante do estudo foi que “pseudowords preceeded by stressed syllables were more likely than pseudowords preceded by stressed syllables to be pronounced with a strong-weak pattern” (KELLY e BOCK, 1988, p. 397). O experimento sugere que “speakers tend to adjust lexical stress in a way that arrays strong beats over time in a regular way, perhaps creating an alternation between strong and weak beats across full sentences” (KELLY e BOCK, 1988, p. 397).

Bolinger (1986, p. 60) apresenta uma análise diferenciada desse fenômeno quando afirma que

it is not a matter of shifting the stress, since the stress, as a potencial for accent, is already there; it is rather the choice of the secondary stress over the primary in a situation where some intonational advantage accrues.

Em *thirteen man*, por exemplo, as sílabas *thir* e *teen*, por serem ambas compostas de vogais plenas, poderiam receber o acento (o *primary stress*).

Vogel, Bunnell e Hoskins (1995) acreditam que a *percepção* de *stress shift* pode estar relacionada à natureza *relacional* das proeminências: uma proeminência menor em *teen* causaria a *impressão* de uma proeminência maior em *thir*.

Os autores desenvolveram um estudo experimental de análise acústica e tarefa perceptiva com 144 estímulos, compostos de frases em contextos variáveis, em relação ao fenômeno de *stress clash*.

Eles concluíram que realmente não há evidência acústica para a *Reversal Analysis* (RA), ou seja, “the shifting of the primary stress of the first word from the final syllable to an earlier syllable in a word”, mas há evidências acústicas (medidas de duração, *pitch* e intensidade) para a *Deletion Analysis* (DA), “stress reduction on the final syllable of the first word with no claim made about a concomitant change in any other syllable”. Os dois tipos de análise da Regra Rítmica são exemplificados na Figura 15 a seguir, em que a grade métrica de nível 3 representa o acento primário:

Reversal Analysis				Deletion Analysis			
x	x			x	x		
x	x	x		x	x	x	
x	x	x	x	x	x	x	x
<i>thirteen women</i> → <i>thirteen women</i>				<i>thirteen women</i> → <i>thirteen women</i>			

Figura 15 – Frase composta de duas palavras exemplificando duas análises para a regra rítmica: a *reversal analysis* (RA) e a *deletion analysis* (RR). Fonte: Vogel, Bunnell e Hoskins (1995)

Patel (2010, p.143) fornece uma explicação, pautada no aspecto articulatório, para esses mecanismos de “deslocamento de acentos”, a qual acredito ser bastante oportuna, pois ela relaciona esse fenômeno à taxa de elocução e à fala espontânea:

The reason such mechanisms exist may be rooted in the mechanics of articulation. Stressed syllables tend to be made with larger jaw movements than unstressed syllables, and it may be biomechanically advantageous to avoid crowding larger jaw movements together when speaking at the fast rates typical of normal conversation.

Bolinger (1986, p. 61) acrescenta que “rhythmic regularity is something that speakers unconsciously strive for” e “they get a tendency not only to space out the accents, but to do it more or less rhythmically”, uma vez que “phased automatic activities are most efficient when they recur at regular intervals” (BOLINGER, 1986, p. 63).

Embora não haja consenso sobre a regra rítmica e a colisão acentual nos estudos linguísticos, optei por comentá-las neste trabalho por considerá-las, no nível perceptivo, relevantes para o estudo das diferenças rítmicas entre o português brasileiro e o inglês norte-americano.

2.3.3.6 PÉS RITMOS ACENTUAIS

Os pés rítmicos acentuais ocorrem na intercalação entre proeminências marcadas por *pitch accents*. São formados pela sílaba proeminente e pelas demais sílabas não proeminentes que a sucedem. Os tamanhos podem ser variáveis, e eles estão vinculados à intenção comunicativa do falante.

A Figura 16, segmentação de um trecho da fala espontânea da cantora Norah Jones analisado em Mauad (em preparação), exemplifica bem as proeminências acentuais carregadas por acentos de *pitch* (*pitch accents*). O enunciado é composto de três pés rítmicos acentuais de 3, 4 e 1 sílabas cada. As sílabas proeminentes são, respectivamente, *watch*, *vi* e *fun*, todas sílabas plenas, portanto. As sílabas não proeminentes são *ing*, *the*, *de*, *os* e *is*. A proximidade da duração dos pés pode dar a impressão de isocronia acentual.

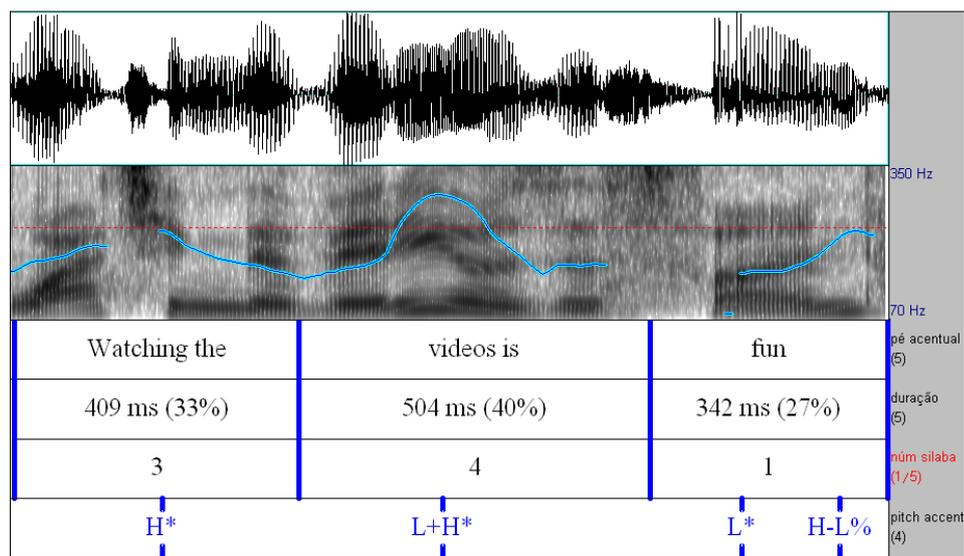


Figura 16 – Forma da onda, espectrograma, curva entoacional, pés rítmicos acentuais, duração do pé, número de sílabas por pé e *pitch accents* do enunciado *Watching the videos is fun*, produzido por Norah Jones. Fonte: revista *Speak-up*

É possível analisar os efeitos que uma sucessão de pés rítmicos acentuais de tamanhos similares causaria. Na sequência, serão considerados os seguintes padrões: (1) sequência de *monobeats*, (2) sequência de troqueus e (3) sequência de dátilos.

Monobeats sucessivos ocorrem em enunciados em que o falante enfatiza cada sílaba de modo a expressar uma atitude, como, por exemplo, a situação em que a filha recusa-se a chamar a irmã e o pai lhe diz furioso: **Call Pat right now!** (um acento de *pitch* em cada palavra).

Outro exemplo desse padrão é o refrão da canção *The Wall*, da banda inglesa *Pink Floyd*, entoado por um coro de jovens, a seguir: **We don't need no education.** Embora se possa aventar a hipótese de que, mesmo agramatical, a opção pela palavra *no* atenda a necessidade de um registro mais informal por se tratar de uma canção popular, acredito, que ela esteja mais relacionada à estrutura rítmica. A opção por *no* contribui para o tom contestatório da canção, pois, dessa forma, todas as sílabas do refrão recebem proeminência (a vogal reduzida em *-tion* /ə/ é “upgraded” para /Λ/):

we! don't! need! no! e! du! ca! tion!

[i ʊ i ʊ ε u eɪ Λ]

Caso o autor da canção tivesse optado por *any* ao invés de *no*, a sequência regular de *monobeats* (1-1-1-1-1-1-1-1) seria substituída pela irregular (1-1-1-2-1-1-1-1), uma vez que *any* constituiria um pé dissilábico. Além disso, como o alongamento entre as sílabas sucessivas não ocorreria por inteiro, a ênfase na fala seria afetada, atenuando, assim, o tom contestatório do refrão.

A sucessão de troqueus, por sua vez, implica na excursão de *pitch* em cada sílaba

proeminente, o que pode levar à incômoda percepção de um jingle²⁵. Bolinger (1986) exemplifica esse padrão com a sentença: I **ha'd** to/ **ma'ke** an / e'ffort /ju'st to/ **ge't** a/ **brea'th** of /ai'r. O autor afirma que, de modo a manter a naturalidade em sua fala, o falante pode, nesse exemplo, interromper o enunciado, pausando após a palavra *effort*, ou até mesmo substituir *a breath of* por *some*, diminuindo, assim, o número de pés rítmicos sucessivos do enunciado.

Em relação a dátilos sucessivos, esse padrão rítmico, composto de uma sílaba proeminente seguida de duas sílabas não proeminentes, soa natural, sendo, portanto, bem aceito pelos interlocutores: He / **wa'nted** to/ **gi've** them a/ **pie'ce** of his / **mi'nd** (BOLINGER, 1986). Nessa sequência de dátilos, especificamente, as sílabas que não recebem proeminência acentual são todas compostas de vogais reduzidas.

2.3.4 Interação entre os ritmos silábico e acentual no inglês

Bolinger (1986) acredita que o ritmo do inglês apresente aspectos rítmicos silábicos e aspectos ritmos acentuais, e que esses dois tipos (ou níveis) de ritmo, silábico e acentual, atuam de modo independente.

At this stage of our knowledge of rhythm in ordinary spoken language there is little that one can say except in general terms – to observe, for instance, that rhythm does exist in the two broad layers of the syllable and the accentual phrase, and to relate it to efficiency in communication (BOLINGER, 1986, p. 73).

²⁵ Bolinger (1986, p. 378) acrescenta que o falante evita produzir um jingle da mesma forma como procura evitar a métrica poética na fala, uma vez que o foco está na mensagem em si e não no veículo que transmite essa mensagem.

A seguir, serão apresentadas as características rítmicas do enunciado *Sampan hardtack tastes like hell*, extraído de Bolinger (1986), tanto em relação às sílabas plenas, quanto aos *pitch accents*.

A Tabela 5 apresenta as sete proeminências silábicas carregadas pelas sílabas (com vogais) plenas. Há apenas uma sílaba em cada um dos sete pés rítmicos, todos de mesmo tamanho (1 sílaba).

pé silábico	Sam	pam	hard	tack	tastes	like	hell
vogal	/æ/	/æ/	/ɑ/	/æ/	/eɪ/	/aɪ/	/ɛ/
número de sílabas no pé silábico	1	1	1	1	1	1	1

Tabela 5 – Enunciado composto de 7 pés silábicos, exemplificando o ritmo silábico do inglês.
Fonte: Bolinger (1986)

O mesmo enunciado, na Tabela 6, apresenta 5 proeminências acentuais carregadas pelos *pitch accents*. Em cada um dos dois primeiros pés, há duas sílabas e, nos três últimos, há uma sílaba. No ritmo acentual, as sílabas não proeminentes podem ser tanto aquelas com vogais plenas quanto reduzidas (neste caso, são plenas).

pé acentual	Sampam	hardtacks	tastes	like	hell
número de sílabas no pé acentual	2	2	1	1	1

Tabela 6 – Enunciado composto de 5 pés acentuais, exemplificando o ritmo acentual do inglês.
Fonte: Bolinger (1986)

Há também enunciados em que toda a sílaba com vogal plena recebe um acento de *pitch*. Haveria assim um “encaixe” entre os dois ritmos, acentual e silábico. É o caso do enunciado, analisado em Mauad (em preparação), da Figura 17, extraído do filme *The Purple Rose from Cairo*, em que a personagem Ema, interpretada por Dianne Wiest, encontra Tom Baxter, interpretado por Jeff Daniels, a seguir:

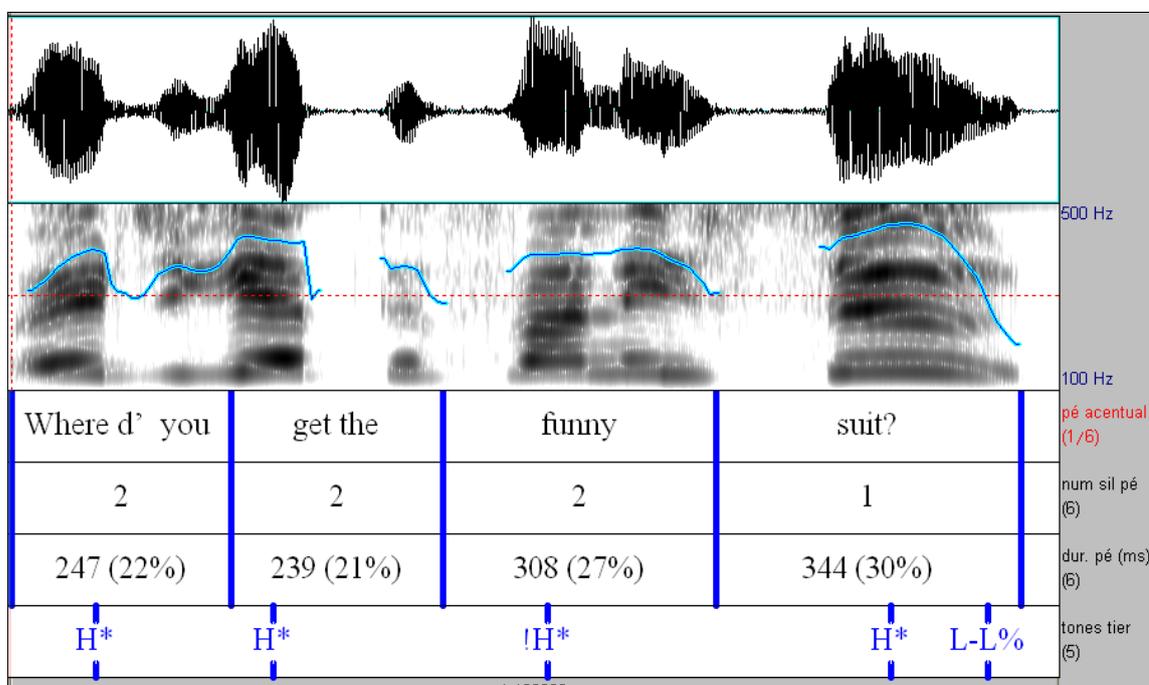


Figura 17 – Forma da onda, curva entoacional, pés rítmicos acentuais, número de sílabas por pé, duração do pé e *pitch accents* do enunciado *Where did you get the funny suit?*, ilustrando o “encaixe” entre o ritmo silábico e o ritmo acentual. Fonte de áudio: filme *The Purple Rose of Cairo*

Na produção do enunciado na Figura 17, verifica-se um encaixe perfeito entre o que Bolinger (1986) denomina ritmo silábico (*full vowel timing*) e o ritmo acentual (*pitch accent timing*), uma vez que os acentos de *pitch* incidem justamente sobre as sílabas com vogais plenas. É importante ressaltar que, nesta emissão, as durações relativas dos pés rítmicos estão relativamente próximas (22%, 21%, 27%

e 30% da duração total do enunciado), o que pode levar o ouvinte à *percepção* de isocronia acentual.

Conforme já discutido, o ritmo silábico é determinado estruturalmente, e, portanto, independe do ritmo acentual. Assim, nem sempre cada sílaba plena recebe um acento de *pitch*. Em outras palavras, é muito comum não haver o “encaixe” perfeito entre os dois ritmos, acentual e silábico, por conta da “rough and ready nature of the accentual rhythm that is grafted onto the much more predictable syllabic rhythm” (BOLINGER, 1986, p. 72).

A Figura 18, referente ao enunciado *I'm up for new experiences* (continuação do diálogo anterior), ilustra bem essa diferença entre os dois ritmos:

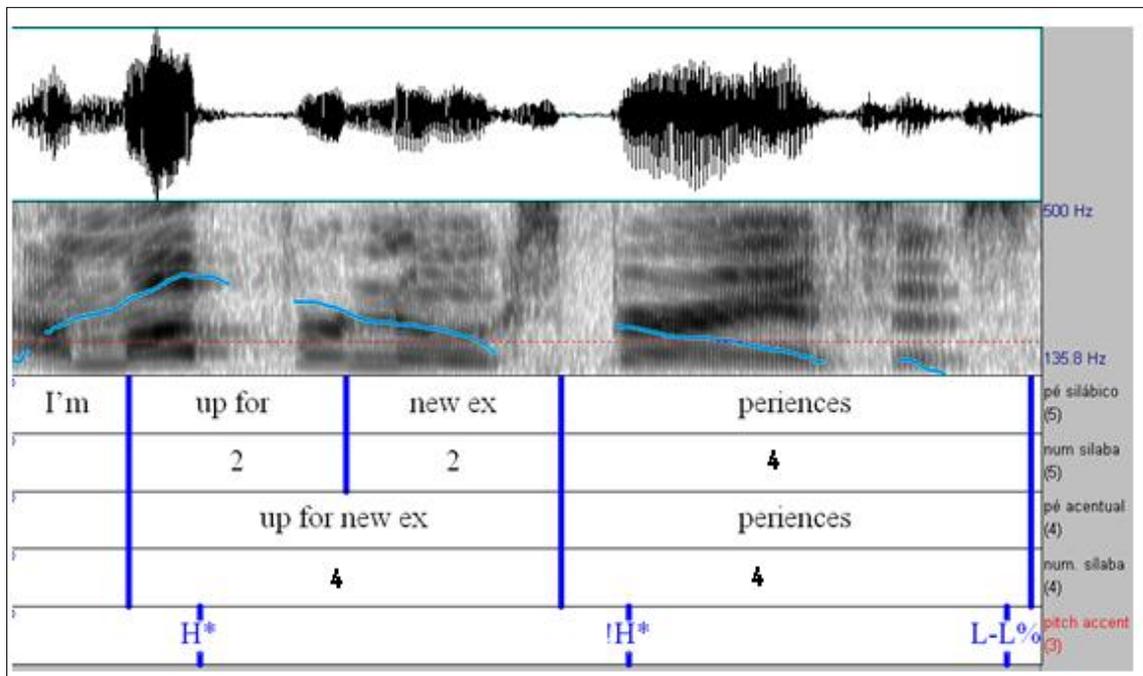


Figura 18 – Forma da onda, curva entoacional, pés rítmicos silábicos, número de sílabas por pé rítmico silábico, pé rítmico acentual, número de sílabas por pé rítmico acentual e *pitch accent* do enunciado *I'm up for new experiences*, ilustrando a independência entre o ritmo silábico e o ritmo acentual. Fonte de áudio: filme *The Purple Rose of Cairo*

As sílabas *up* e *per*, por receberem o *pitch accent*, destacam-se das demais. A sílaba *new*, dada a sua qualidade vocálica (vogal plena), distingue-se das sílabas *for*, *ex*, *rien* e *ces* (vogais reduzidas). Entretanto, embora tenha uma vogal plena, sem um acento de *pitch*, a sílaba *new* não se destaca em relação às outras.

O enunciado pode, assim, ser dividido em três pés rítmicos acentuais, com 2, 2, 2 e 4 sílabas cada e 2 pés rítmicos silábicos, com 4 sílabas cada.

Nota-se que, no enunciado da Figura 18, os pés rítmicos acentuais têm duração mais próxima do que os pés rítmicos silábicos. Assim, nesse caso, estaria o *pitch accent*, e não a sílaba plena, mais propenso à regularidade temporal.

Entretanto, para Bolinger (1986), esse não seria o padrão rítmico mais observável no inglês norte-americano. Para o autor, haveria maior probabilidade de se verificar intervalos de tempos aproximados entre pés silábicos do que entre pés acentuais, uma vez que os pés silábicos são determinados estruturalmente, e o *pitch accent*, evento que marca o pé rítmico acentual, depende da intenção comunicativa do falante, estando, assim, mais sujeito à variação. Conforme afirma Bolinger (1978),

There is supposed to be a basic distinction between “stress-timing and syllable timing” and stress-timing is always exemplified by a supposed tendency to isochrony between accentual peaks. But since accents respond to the meanings that the speaker wishes to highlight, unless meaning is somehow adjusted to rhythm the result can hardly be a succession of equal intervals. English is claimed to be a stress-timed language, but in a phrase such as “*the quite unnecessary incomprehensibility of his words*” the central interval (*ry incomprehensi*) is substantially longer than the flanking ones (*unneces e of his*). What English does have is a striking near-isochrony based on vowel quality. Syllables with reduced vowels “borrow time” from preceding syllables with full vowels, so that two such sentences as *Jon drank beer* and *Johnny drank a beer* are of nearly equal length. Much more needs to be known about rhythm before it can be discussed in the context of linguistic universals” (BOLINGER, 1978).

Tendo exposto a fundamentação teórica sobre o ritmo, no próximo capítulo apresento a fundamentação teórica sobre a entoação.

CAPÍTULO 3 QUESTÕES SOBRE A ENTOAÇÃO

DA FALA

The surface of the ocean responds to the forces that act upon its movements resembling the ups and downs of the human voice. If our vision could take it all in at once, we would discern several types of motion, involving a greater and greater expanse of sea and volume of water: ripples, waves, swells and tides (...)The ripples are the accidental changes in pitch, the irrelevant quavers. The waves are the peaks and valleys that we call accent. The swells are the separations of our discourse into its larger segments. The tides are the tides of emotion.

(Bolinger)

3.1 CONCEITUAÇÃO DE ENTOAÇÃO

A entoação corresponde a variações na melodia da fala. Em termos acústicos, ela corresponde à variação de f_0 ao longo dos enunciados orais. Dada a relevância que essa variação de f_0 desempenha na prosódia como um todo, os termos *entoação* e *prosódia* são, para alguns autores, considerados equivalentes. Neste trabalho, entretanto, entoação deve ser entendida como apenas mais um dentre os elementos prosódicos, assim como o ritmo e a tonicidade.

A entoação pode ser analisada em seu aspecto perceptual, acústico e articulatório. Percebemos, por exemplo, que a segunda sílaba da palavra *Hello?*, ao atendermos uma ligação telefônica, é mais “alta” (aguda) do que a primeira. Em outras palavras, *-lo* está em um *pitch* mais alto. *Pitch* é, portanto, uma propriedade auditiva – algo que percebemos.

O f_0 , ou frequência fundamental, medida em Hertz, correlaciona-se com o número de vezes que as vibrações das cordas vocais (abertura/abdução e fechamento/adução) repetem-se no tempo. Por exemplo, uma frequência de 100 Hz significa que, no intervalo de um segundo, as cordas vocais abriram-se e fecharam-se 100 vezes. Por ser muito rápido, esse movimento é invisível a olho nu.

A frequência fundamental (f_0) é o parâmetro acústico mais importante para o estudo da entoação. A estrutura laríngea, que inclui as pregas vocais, é responsável não só pela alternância entre os sons vozeados e não vozeados, mas também pela modulação de *pitch*. O movimento vibratório das pregas vocais resulta acusticamente em um sinal que tem seu tom fundamental, ou frequência fundamental (f_0), como o primeiro harmônico, sendo que os harmônicos

subsequentes possuem frequência múltipla do fundamental, conforme ilustra a Figura 19 a seguir:

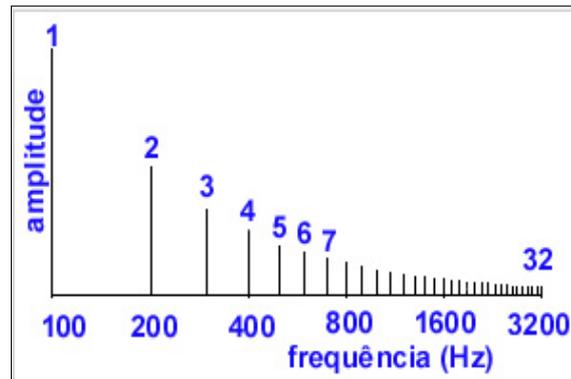


Figura 19 – Frequência fundamental de 100 Hz e subsequentes harmônicos múltiplos

Segundo Laver (2002), o *pitch* é relativo: (1) no que se refere ao julgamento perceptivo feito pelo ouvinte sobre as variações melódicas da voz do falante; (2) no que diz respeito à sensação de *pitch* de dada sílaba na cadeia da fala em relação às sílabas vizinhas, que pode ser igual, mais alto ou mais baixo.

Diferentes tipos de extensão de *pitch* podem ser categorizados para uma determinada voz:

1. A extensão orgânica de *pitch* refere-se à capacidade máxima que um falante é fisicamente capaz de produzir, segundo fatores biologicamente determinados pela anatomia de sua laringe e de fatores fisiológicos;
2. A extensão paralinguística de *pitch* refere-se ao ajuste, no âmbito orgânico, da extensão de *pitch*, com o propósito de sinalizar informação atitudinal (como surpresa, raiva, tristeza, impaciência etc.);

3. A extensão linguística de *pitch* refere-se à variação de *pitch* que é fonologicamente relevante na voz habitual do falante em conversação paralinguisticamente não marcada, ou seja, de atitude neutra.

Todas essas variações de *pitch* são acompanhadas por mudanças de ajustes de qualidade de voz. Os níveis mais altos na extensão de *pitch* estão relacionados com o falsete, os mais baixos com a voz rangida (*creaky-voice*) e, entre os dois, com a voz modal.

Prator (1985) afirma que a queda de *pitch* no final de uma sentença pode ser vista como um tipo de pontuação, um ponto final vocal/verbal, indicando que o pensamento está completo.

Quando a queda de *pitch* no término de sentenças declarativas, imperativas e *wh-questions* é pequena²⁶, o ouvinte acredita que há mais a ser dito. Ele pode sentir uma sensação desagradável e confusa, já que permanece a impressão de um pensamento não concluído, uma vez que “it implies nonfinality instead of finality” (PIKE, 1945, p. 49).

3.2 FUNÇÕES DA ENTOAÇÃO

Ao utilizar um determinado padrão entoacional, o falante o faz de acordo com sua meta comunicativa. A entoação geralmente não expressa significados por si só, mas em relação ao conteúdo semântico do enunciado. A entoação, portanto, adquire diversas funções em consonância com a pluralidade de aspectos de natureza

²⁶ Correspondente ao *pitch level 3* de Pike (1945).

gramatical, atitudinal, discursiva, pragmática, sociolinguística etc, que entram em ação no ato de comunicação.

3.2.1 Função gramatical

Em sua função gramatical, a entoação fornece informação sobre a estrutura gramatical ao, por exemplo, estabelecer fronteiras entre orações subordinadas e distinguir diferentes modalidades (declarativas, interrogativas, exclamativas etc). Segundo Chun (2002), no inglês norte-americano, sentenças declarativas, *wh-question*, ordens e exclamações são tradicionalmente descritas como tendo declinação final descendente. Já as perguntas totais (*yes/no questions*) geralmente terminam em tons ascendentes²⁷.

3.2.2 Função discursiva

Para Chun (2002), a função discursiva da entoação pode ser estudada sob quatro aspectos: (1) a estrutura informacional, (2) os atos de fala, (3) a coesão textual e (4) o controle de interação entre os participantes do discurso (o *turn-taking*).

A estrutura informacional está relacionada à perspectiva baseada em pressuposições que o falante faz em relação ao ouvinte e a seu conhecimento. Ela

27 Bolinger, diferentemente de Halliday, acredita que, por razões pragmáticas, tais “regras” não se sustentam, visto que “the encounters between intonation and grammar are casual, not causal” (BOLINGER, 1958, p. 37).

marca conhecimento compartilhado no discurso (informação dada), informação relevante no discurso (informação nova), além de estabelecer as fronteiras que determinam o discurso. A estrutura informacional permite que os interlocutores realcem certas palavras com relevância semântica, como no diálogo a seguir:

JOE: Do you you like that *shirt*?

BEA: *Which* one?

JOE: The *brown* one.

As palavras *shirt*, *which* e *brown* são pronunciadas em um *tom* mais alto do que as demais por apresentarem informação nova.

Segundo a teoria dos atos de fala (SEARLE, 1962), os enunciados são ações verbais. Por exemplo, na produção de um enunciado como *It's hot in here*, a entoação do falante pode expressar se ele está fazendo uma simples afirmação, uma queixa, ou se está pedindo indiretamente a alguém que abra a janela (CHUN, 2002). A entoação nos atos de fala expressa, portanto, a intenção do falante (exemplo: pedido vs sugestão; afirmação vs queixa).

Conforme afirma Leite (2009), “em uma produção linguística, o falante tem como objetivo transmitir informações. Assim, ele estrutura seu discurso de forma que os enunciados sejam considerados textos, e não somente frases isoladas.” Enquanto no texto escrito esse mecanismo é feito por conectivos (*então, além disso* etc), nos enunciados orais essa função coesiva pode ser exercida pela entoação.

Também inserido na função discursiva da entoação, está o controle das estruturas de interação entre os participantes do discurso, o *turn-taking*. A estrutura interacional está relacionada às intenções, atitudes e reações do falante em relação à

conversação em si. Por exemplo, ao mudar a entoação, o falante pode sinalizar a seu interlocutor estar dando continuidade ou finalizando um tópico, controlando, assim, o processo de *turn-taking*.

3.2.3 Função atitudinal

Durante a produção de um enunciado, mudanças na frequência fundamental podem expressar a atitude e o estado emocional do falante. Embora tanto as emoções quanto as atitudes constituam estados afetivos, somente as atitudes podem ser analisadas em função da *configuração* do contorno entoacional.

A partir de resultado de experimento cujo *corpus* foi constituído de logatomas (frases-veículo sem sentido) interpretados por atores empenhados em vocalizar várias emoções, Bänzigher e Scherer (2004) constataram que há contornos entoacionais específicos para diferentes emoções.

Foram extraídos os valores médios e desvios padrões das elevações e quedas de f_0 para cada emoção expressa. As inclinações tenderam a ser mais marcadas nas emoções mais fortes – especialmente euforia –, e menos marcadas nas emoções com menor grau de excitação – especialmente tristeza, alegria e ansiedade. A entoação acrescenta, portanto, um aspecto atitudinal ou emocional ao enunciado²⁸.

²⁸ Segundo Fonagy (1993), a emoção está relacionada à descarga espontânea de uma tensão psíquica e, portanto, foge do controle do indivíduo. São consideradas emoções a raiva, a tristeza, a vergonha, o medo e o desespero. A atitude, por sua vez, designa comportamento determinado, consciente, controlado e que tem componente moral e intelectual. São consideradas atitudes o desprezo, a frieza, o entusiasmo, a reprovação, a circunspeção e o sarcasmo.

3.2.4 Função sociolinguística

A entoação pode assumir a função sociolinguística: determinar contrastes que possibilitem ao ouvinte identificar falantes como sendo membros de diferentes grupos sociais, determinados por ordem de sexo, idade, ocupação, assim como identificar indivíduos com base em características idiossincráticas. Brend (1975) afirma que, no inglês, a maioria dos homens tem três níveis contrastivos de entoação, enquanto que as mulheres apresentam quatro níveis. McConell-Ginet (1983) acrescenta que a fala feminina apresenta variedade maior de padrões e inflexões tonais. A voz masculina, por sua vez, tem tendência maior à queda para tons mais baixos de *f0* (voz rangida ou *creaky voice*) do que a voz feminina.

Geoffrey Sampson (*apud* Mauad, 2007), linguista britânico, entretanto, quando indagado se também acredita que *creaky voice* no inglês norte-americano está associada à voz masculina, afirma que essa qualidade de voz apresenta-se também na fala das norte-americanas e que seu uso pode ocasionar conflitos em contextos interculturais:

From a British point of view, what you say about American speakers using a narrow pitch range and a lot of creaky voice is interesting: it matches my impressions (though whereas you write specifically about males, I notice it *_specially_* with American women -- it may be more noticeable to me with women because in a British context it seems specially inappropriate from women). And it certainly has strong potential for cross-cultural misunderstanding (MAUAD, 2007, p. 27).

Segue-se que o uso de determinados padrões entoacionais pode causar no ouvinte efeitos impressionantes e reforçar estereótipos associados à sua cultura.

3.2.5 Função acentual

Afirmar que a entoação tem função acentual implica dizer que a entoação desempenha um papel nas proeminências, e, conseqüentemente, no ritmo, conforme afirma Tench (1992, p. 30):

I will argue that the two most fundamental functions of intonation in English, often seen as realized as “tone of voice”, are in fact realized by rhythm: accent, which attracts the listener’s attention to the salient syllable in a rhythmic foot, and juncture, which segments speech into rhythmic groups, groups of words which form, in Halliday’s terms, distinct moves in the speech acts.

A função acentual da entoação está, portanto, relacionada ao modo como ocorre a inflexão tonal nas proeminências acentuais (ROACH, 1998). O falante pode dar proeminência a certas sílabas por meio do acento de *pitch*. *Pitch* e proeminência acentual não são, portanto, independentes, posto que a inflexão de f_0 é a pista principal para o acento.

Línguas como o inglês e o alemão fazem uso extensivo do *pitch accent* para marcar proeminência. Os parâmetros acústicos de f_0 , juntamente com intensidade e duração, são os três correlatos acústicos principais da proeminência, tanto no nível da palavra (*stress*) quanto no da frase (*accent*).

O primeiro a descrever a relação de f_0 com proeminências foi Fry (1958). O autor constatou que a percepção de *accent* envolve a interação de três parâmetros: duração, intensidade e f_0 , sendo que f_0 exerce o papel mais importante.

Nem todas as línguas utilizam *f0* da mesma forma na marcação das proeminências. (BARBOSA, 2006, p. 167) afirma ser mais fácil a detecção das sílabas proeminentes em inglês do que em português, “dado o controle sistemático e frequente de *f0* no inglês”²⁹.

3.2.5.1 FOCO

Uma das formas de modelar nossa percepção das unidades dos contornos entoacionais de um enunciado consiste em dividi-lo em duas partes: *tema* (aquilo que é dito, ou dado previamente ou inferível, subentendido pelo contexto) e *rema* (aquilo que é dito sobre o tema, que responde à uma pergunta subjacente e que pode trazer algo de novo ao enunciado). Essa configuração típica composta de dois acentos não adjacentes recebe, para o inglês, o nome de *the hat pattern*.

Ao analisarmos a segunda linha do diálogo abaixo:

A: Did you stay in the mall?

B: Nope. I walked to **school**.

Temos que *school* constitui a rema e “*I walked to*”, tema. Considerando-se a definição de Bolinger (1986) para o foco como “the chief accent in the rheme”,

²⁹ Beckman (1992, *apud* BARBOSA 2006) constatou, a partir de dados de Dauer (1983), que parágrafos lidos em línguas tradicionalmente referidas como de ritmo acentual produzem pés rítmicos curtos (até 5 sílabas) e parágrafos lidos em línguas tidas como de ritmo silábico, produzem pés rítmicos mais longos (até 10 sílabas). Barbosa (2006) acrescenta que foi justamente esse espaçamento maior entre as proeminências que ele encontrou ao analisar frases isoladas e parágrafos lidos em português brasileiro, língua que, segundo o autor, tenderia ao ritmo silábico.

conclui-se que o foco (também chamado núcleo, acento frasal, ou *sentence stress*³⁰) no enunciado em questão recai sobre *school*.

O foco realiza-se por um movimento mais amplo de *f0* – sendo, portanto, um *pitch accent* – e alinha-se sobre a sílaba tônica da palavra que, na frase, veicula a informação mais importante. Duração e frequência fundamental são os parâmetros mais atuantes na construção desse tipo de proeminência.

Dentro de um grupo prosódico (ou frase entoacional), qualquer sílaba acentuada lexicalmente pode receber o foco, desde que tenha relevância discursiva. Por exemplo, no diálogo a seguir, extraído de Grant (2004), o foco desloca-se para cada uma das palavras destacadas, uma vez que elas apresentam informação nova:

A: What's the matter?

B: Oh, I lost my **book**.

A: **Which** book?

B: My **Math** book.

Tal realce é acompanhado de uma excursão mais acentuada de *f0*, pelo prolongamento da vogal e pela maior precisão na articulação.

³⁰ Roach (1998) não considera *sentence stress* um termo apropriado, pois, enquanto *sentence* corresponde a uma unidade gramatical, *stress* corresponde a uma unidade fonológica.

3.3 MODELOS DE DESCRIÇÃO DA ENTOAÇÃO

A seguir, apresento modelos fonológicos para a análise da curva entonacional: os modelos *tune-based* e *tone-based* da escola Britânica, a teoria de *pitch accent*, o modelo *Discourse Intonation* e os modelos gerativos.

3.3.1 Escola Britânica

Na fala encadeada, o *f0* varia continuamente. Para que essa variação contínua de *pitch* possa ser analisada em termos de padrões/configurações de *pitch*, é necessário desmembrar o fluxo de fala em unidades. Essa análise pode ser feita de modo holístico, isso é, no contorno como um todo (*tune-based*) ou na soma das configurações que compõem o contorno (*tone-based*). A Escola Britânica privilegia o enfoque de caráter mais impressionístico (de outiva) para o estudo dos padrões de *pitch*. Os contornos são analisados a partir das unidades fonológicas, os grupos tonais, *tone units/groups*, os quais são subdivididos em *prehead*, *head núcleo* e *tail*, conforme a Figura 20.

	head			
anacrusis (pre head)	onset		núcleo	tail
I'll	PHONE	you to	MOR	row

Figura 20 – A estrutura de um grupo tonal em inglês. Fonte: Underhill (1994)

3.3.2 Teoria de *Pitch Accent*, de Bolinger

Assim como os teóricos da Escola Britânica, Bolinger acredita que as configurações de *pitch* são mais significativas do que os níveis de *pitch*. O *pitch accent* é um evento discreto, perceptível, enquanto a entoação engloba os aspectos não acentuais dos padrões de *f0*, ou seja, “the kinds of patterns that are said to be gradient rather than discreet” (CHUN, 2002, p. 22). Segundo Ladd (1994, p. 49), “Bolinger succeeds in making some sense of the relationship between stress and pitch, in languages like English and Dutch.”

O *pitch accent* é o elemento prosódico que é ao mesmo tempo marcador de proeminência e parte estruturante dos contornos de *f0*. Dessa forma, o autor estabelece uma relação estreita entre a entoação e o ritmo. No livro *Intonation and its Parts*, Bolinger (1986) distingue a prosódica melódica (*melodic prosody*) e a prosódia acentual (*accentual prosody*). A entoação insere-se na prosódia melódica e o ritmo, na prosódia acentual. Isso explica porque o autor afirma que “Rhythm (and the accents that punctuate it) is one kind of prosody in English; Intonation (and the tunes that it composes) is another. Both involve pitch” (BOLINGER, 1986, p. 37).

Bolinger refere-se a *profiles* como sendo as configurações de *pitch* fonologicamente contrastivas para o inglês norte-americano, sendo que as principais são *profile A*, *profile B* e *profile C*. A Figura 21 mostra três tipos de *profiles*:

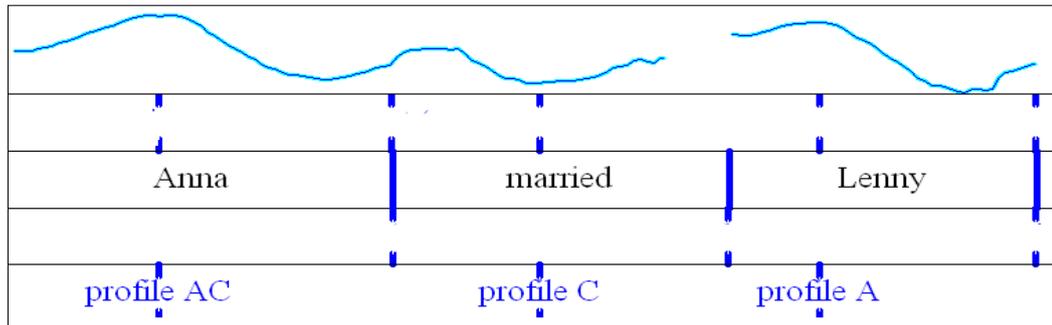


Figura 21 – Contorno de f_0 , palavras e notação entoacional de três configurações de *pitch* descritas por Bolinger (1986) (*profile AC*, *profile C* e *profile A*) do enunciado *Anna married Lenny*.
Fonte de áudio: Silverman *et al* (1992)

Embora Bolinger reconheça a importância de outras pistas acústicas como duração, intensidade, ritmo e qualidade vocálica para a proeminência acentual, o autor acredita ser o *pitch accent* a pista mais relevante.

3.3.3 *Discourse Intonation*

Este é um modelo de forte viés interacionista, uma vez que ele analisa o contorno entoacional a partir da análise da interação entre falante e ouvinte, ambos participantes de um ato comunicativo. Nesse modelo, não é feita nenhuma referência explícita à sintaxe, mas, sim, às intenções comunicativas dos falantes. Há, assim, um conjunto finito de configurações de *pitch*, cuja implementação visa a atingir coerência e coesão no discurso. Os cinco tons propostos por Brazil (1975) são: (1) descendente-ascendente – “r” (*referring tone*), (2) descendente – “p” (*proclaiming tone*), (3) ascendente “r+”, (4) ascendente-descendente – p+ e (5) *low-rising*. A Figura 22, a seguir, mostra dois grupos entoacionais, cada um com um tom:

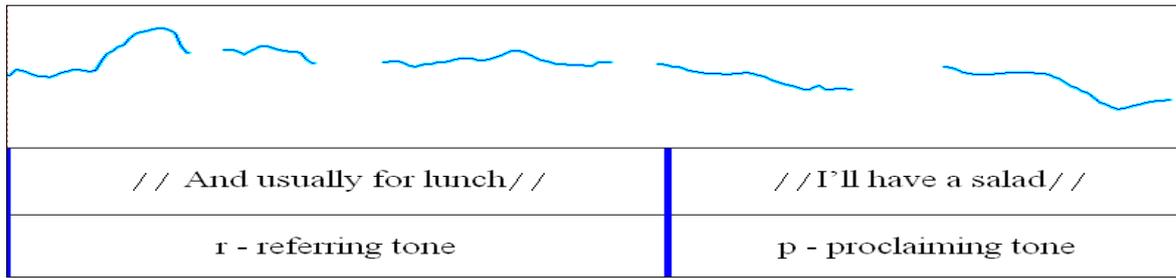


Figura 22 – Contorno de f_0 , segmentação em 2 grupos tonais e notação entoacional, segundo Brazil (1995), do enunciado *And usually for lunch, I'll have a salad*, extraído da série *Seinfeld*

3.3.4 Modelos gerativos

Como originalmente proposto por Chomsky e Halle, nesses modelos o *accent* não precisa estar marcado no léxico, uma vez que essa proeminência pode ser implementada por regras fonológicas. Essas regras são aplicadas à estrutura de superfície, a qual é rotulada a partir de seu componente sintático, por meio de siglas, tais como S (substantivo), V (verbo), SN (sintagma nominal), etc. Tal estrutura é, então, dividida em frases fonológicas, grades métricas e árvores métricas.

Segundo Abaurre (1993) a fonologia métrica, defendida por Liberman e Prince, questiona a análise linear do acento em inglês. A proeminência não deve ser atribuída a segmentos, “mas entendida como a proeminência de algumas sílabas que são colocadas em relação umas com as outras, nos domínios das unidades prosódicas de extensão variada”.

O modelo gerativo proposto por Pierrehumbert (1980) é baseado em um sistema binário de tons, H (high) e L (low), que compõem uma frase entoacional, sendo que esses dois níveis são relativos, e não absolutos, como nos quatro tons

de Pike (1945). Pierrehumbert (1980) resgata o termo *pitch accent*, originariamente proposto por Bolinger em 1958, considerando-o “building blocks of pitch contours” (LADD, 1994, p. 46).

Os *pitch accents*, que se alinham à sílaba tônica, são representados por um asterisco (H* e L*), os tons de frase intermediária que seguem os *pitch accents* são representados por um hífen (L- ou H-) e os tons na fronteira da frase entoacional são representados pelo símbolo % (H% ou L%), conforme a Figura 23, a seguir:

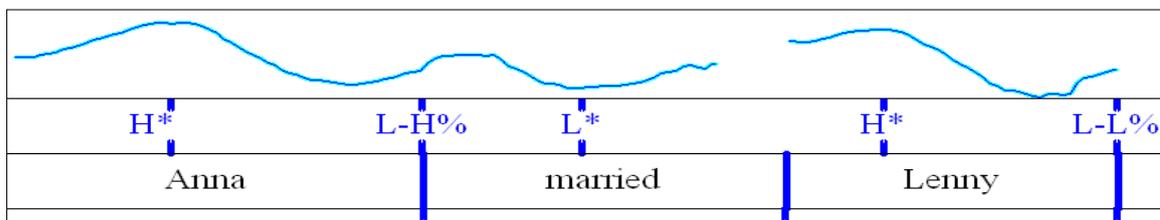


Figura 23 – Contorno de f_0 , grupos tonais, segmentação em 2 grupos tonais e notação entoacional segundo Pierrehumbert do enunciado Anna married Lenny. Fonte: Silverman *et al* (1992)

3.3.4.1 SISTEMA DE NOTAÇÃO ENTOACIONAL ToBI

O sistema de notação entoacional ToBI (*Tones and Break Indices*) (SILVERMAN *et al*, 1992) está inserido no programa da fonologia entoacional. Ele pressupõe que significados presentes na curva entoacional são passíveis de representação fonológica e podem ser codificados no sinal acústico por meio de regras de implementação fonética.

Para Ladd (1996), esse sistema é uma versão mais recente do modelo de Pierrehumbert (1980). Com o objetivo de tornar a análise mais simples e foneticamente orientada, há simplificação na notação de alguns tons (como, por

exemplo, o tom H* equivalendo ao tom H*+L de Pierrehumbert). Além disso, há um número maior de *tiers* (camadas de notação entoacional).

A transcrição de um enunciado por meio do sistema de notação ToBI requer o registro do contorno de *pitch* e rótulos simbólicos para dois eventos tonais: os *pitch accents* e os tons de fronteira (*boundary tones*).

Os *pitch accents* incidem sobre cada sílaba proeminente. A Tabela 7 apresenta os 5 tons para o *pitch accent*:

H*	Tom alto (situado na parte superior da <i>faixa de pitch</i> do falante na produção do enunciado). Inclui tons na faixa média de <i>pitch</i> , mas exclui tons muito baixos
L*	Tom baixo (situado na parte inferior da extensão/faixa de <i>pitch</i> do falante)
L*+H	Tom baixo alinhado à sílaba tônica imediatamente seguido de elevação acentuada de <i>f0</i> até a parte superior da extensão de <i>pitch</i> do falante
L+H*	Alvo de acento de <i>pitch</i> em tom alto imediatamente precedido de elevação acentuada a partir da parte inferior da faixa de <i>pitch</i> do falante
H+ !H*	Descida de um tom alto até a sílaba com <i>pitch accent</i> . Só é utilizado quando o material que antecede a sílaba com <i>pitch accent</i> está em um tom mais alto e não é <i>accented</i> . O símbolo ! significa que houve <i>downstep</i> ³¹

Tabela 7 – Rótulos dos 5 tons para o acento de *pitch* para o inglês, segundo o sistema de notação entoacional ToBI

³¹ O termo *downstep*, “equivalente ao termo ‘abaixamento’ no português”, (ROCCA, 2003, p. 92) é definido como uma compressão da extensão de *pitch*, acionada fonologicamente, que abaixa os alvos de *f0* de qualquer tom H subsequente a um acionamento (de *downstep*) (SILVERMAN *et al*, 1992).

Os acentos frasais são marcados nas *fronteiras intermediárias* do sintagma entoacional. A Tabela 8 apresenta os dois acentos frasais:

L -	fronteira intermediária baixa
H-	fronteira intermediária alta

Tabela 8 – Rótulos para os acentos frasais para o inglês, segundo o sistema de notação entoacional ToBI

Os tons de fronteira são marcados nas *fronteiras finais* do sintagma entoacional. Os três tons de fronteira são apresentados na Tabela 9:

L%	tom de fronteira final baixa
H%	tom de fronteira final alta
%H	tom inicial de fronteira alta

Tabela 9 – Rótulos para os três tons de fronteira, segundo o sistema de notação entoacional ToBI

Como os sintagmas entoacionais são compostos de uma ou mais frases intermediárias e de um tom de fronteira, as fronteiras finais do sintagma entoacional terão dois tons. Há quatro combinações de tons para as fronteiras finais dos sintagmas entoacionais (Tabela 10):

L- L%	Fronteira final baixa precedida por acento frasal baixo. Contorno característico das sentenças declarativas no inglês
L- H%	Fronteira final alta precedida por acento frasal baixo. Contorno característico da <i>continuation rise</i> no inglês
H- H%	Fronteira final alta precedida por acento frasal alto. Contorno característico das questões totais (<i>yes/no questions</i>) no inglês
H- L%	Fronteira final baixa precedida por acento frasal alto. Contorno caracterizado por um <i>plateau</i> na parte mediana da faixa de <i>pitch</i> do falante

Tabela 10 – Rótulos para os quatro tons de fronteira final, segundo o sistema de notação entoacional ToBI

Utilizando-se o programa PRAAT (Boersma e Weenick, 2005), são estabelecidas quatro camadas, ou *tiers*, na transcrição dos eventos tonais. Essas camadas são apresentadas na Tabela 11:

Camada ortográfica (orthographic tier)	Transcrição das palavras que compõem o enunciado
Camada dos tons (tone tier)	Inventário dos tons que compõem o enunciado
Camada de grau de junção (break-index tier)	Escala de 0 a 4, representando o grau de junção entre palavras ou entre as palavras finais e o silêncio no fim dos enunciados. (0: mais coesa, (ex: <i>did you</i>); 1: entre palavras; 2: pausa sem marcas tonais ou mais fraca que a da fronteira intermediária; 3: fronteira intermediária; 4: final do sintagma entoacional)
Camada sobre ocorrências verificadas na fala (miscellaneous tier)	Observações gerais sobre perturbações no enunciado, tais como: respiração, risada, hesitações, gaguejo etc.

Tabela 11 – Quatro camadas, ou *tiers*, na transcrição dos eventos tonais, segundo o sistema de notação entoacional ToBI

No exemplo da Figura 24, extraído do arquivo de áudio *Anna*, disponibilizado por Silverman *et al* (1992), podem ser visualizadas as quatro camadas (*tiers*) e seus eventos tonais para três enunciados ortograficamente idênticos, mas prosodicamente distintos.

Primeiro enunciado:

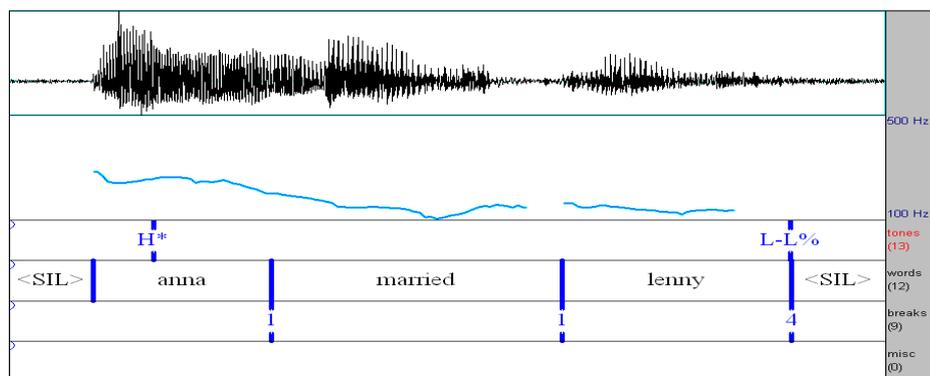


Figura 24 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI: (1) ortográfica, (2) de tons, (3) grau de junção e (4) ocorrências para o primeiro enunciado, do arquivo *Anna*.
Fonte: Silverman *et al* (1992)

Há, no primeiro enunciado, um único acento de *pitch* produzido na parte superior da extensão de *pitch* do falante, recaindo sobre *Anna* (H*). O grau de junção (*break-index*) 1 entre *Anna*, *married*, e *Lenny* demonstra haver bastante coesão e ausência de pausa entre essas palavras. O tom L-L% de fronteira final indica que se trata de uma sentença declarativa.

Segundo enunciado:

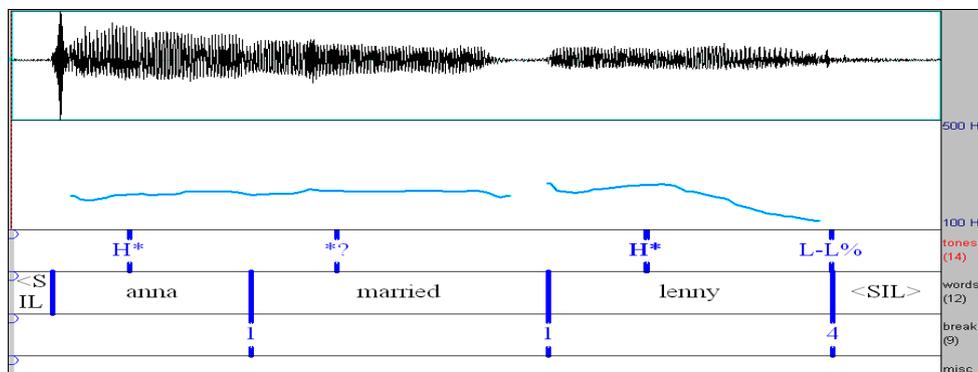


Figura 25 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI: (1) ortográfica, (2) de tons, (3) de grau de junção e (4) das ocorrências, para o segundo enunciado do arquivo *Anna*. Fonte: Silverman *et al* (1992)

Há, no segundo enunciado, na Figura 25, dois acentos de *pitch* na parte superior da extensão de *pitch* do falante, o primeiro incidindo sobre Anna (H*) e o segundo sobre Lenny (H*). O símbolo *? significa que o transcritor não tem certeza se há realmente um acento de *pitch* recaindo sobre *married*.³²

Terceiro enunciado:

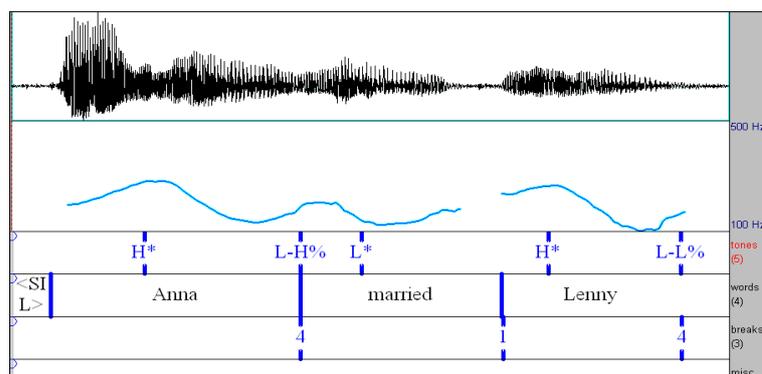


Figura 26 – Forma da onda, curva entoacional e as quatro camadas do sistema ToBI (1) ortográfica, (2) de tons, (3) grau de junção e (4) ocorrências para o terceiro enunciado do arquivo *Anna*. Fonte: Silverman *et al* (1992)

³² Minha interpretação é que *married* seja apenas um elemento da porção não acentuada do pé acentual *Anna married*.

O contorno da frase entoacional *Anna*, no terceiro enunciado (Figura 26), o tom H* L-H% (*fall-rise*), tem implicações discursivas: *Anna* já faz parte do conhecimento mútuo (*common ground*) entre falante e ouvinte (Brazil, 1975), ou seja, refere-se a algo já mencionado antes. Segundo Bolinger (1986), esse contorno pode ser usado em saudações (*Hel/oo!*), contraste seletivo (*I like **waffles**. Does that count as an answer?*), ênfase irônica (*Well **good** for you!*) e admoção (*Don't do **that!***)³³.

O tom de fronteira L-H% e o grau de juntura (*break-index*) 4 após *Anna* significa que o enunciado é composto de duas frases entonacionais.

3.3.5 Modelo dinâmico-entoacional

Segundo Lucente (2008) enquanto os modelos gerativos consideram a entoação como resultante da atuação de características fonológicas subjacentes ligadas a características fonéticas de superfície, a teoria dos sistemas dinâmicos considera a melodia da fala (entoação no sentido mais amplo) como “portadora de informações comunicativas que são produzidas unicamente pelo sistema articulatório e codificadas e transmitidas por meio de *f0*” (LUCENTE, 2008).

O modelo dinâmico-entoacional não propõe, portanto, dois níveis de representação distintos. Na realidade, ele *integra* os níveis fonético e fonológico e privilegia os contornos entoacionais, a extensão de *pitch* e o alinhamento específico de *f0* em relação à sílaba tônica.

³³ Falantes brasileiros de inglês têm dificuldade de produzir esse contorno, particularmente em um único núcleo silábico (ROCCA, 2003), como ocorre justamente na palavra *that* no enunciado *Don't do that!*.

O alinhamento do pico de f_0 em relação à sílaba tônica merece papel de destaque nos trabalhos experimentais desenvolvidos por Kohler (2006). O autor afirma que as três sincronizações do contorno do pico de f_0 , no início, meio ou final da vogal acentuada, constituem categorias distintas da fonologia do inglês e podem codificar diferentes sentidos de ordem pragmática:

a) pico de f_0 no *onset* da vogal (*early peak*): saber, terminar uma exposição, resumir, resignação, **finalidade** (*finality*);

b) pico de f_0 na porção medial da vogal (*medial peak*): observar, perceber, começar uma nova exposição, **receptividade** (*openness*);

c) pico de f_0 na porção final (*late peak*): observar, perceber que algo não correspondeu às expectativas, surpresa, **imprevisto** (*unexpectedness*).

Apresentada a fundamentação teórica sobre entoação passamos, no próximo capítulo, à consideração dos aspectos metodológicos.

CAPÍTULO 4 METODOLOGIA DE PESQUISA

Neste capítulo, descrevo a metodologia de pesquisa empregada, apresentando a natureza da pesquisa, o tipo de *corpus* utilizado, a caracterização dos sujeitos, os instrumentos e procedimentos utilizados para a coleta, o sistema de notação dos eventos tonais e os procedimentos de segmentação, etiquetagem e análise dos dados. Também teço algumas considerações para guiar a leitura de espectrogramas e traçados de frequência fundamental.

4.1 ESCOLHA DA METODOLOGIA DE PESQUISA

Nesta tese, desenvolverei um estudo experimental apoiado em instrumental de análise fonético-acústica. As deduções lógicas ao longo do trabalho serão feitas por meio de proposições verbais. Tais proposições serão baseadas na análise e no tratamento estatístico dos dados acústicos, na elaboração de gráficos comparativos, e em observações do próprio pesquisador.

Todo o estudo é intrinsecamente motivado pelo meu interesse em pesquisar o modo como aprendizes brasileiros de inglês implementam a estrutura rítmica e melódica nos enunciados orais de um diálogo, o que pode afetar a apreensão de sentidos. A partir da comparação com a produção do mesmo diálogo por falantes nativos, pretendo analisar os contrastes, que se verificam tanto nos aspectos rítmicos quanto entoacionais.

4.2 SUJEITOS DE PESQUISA

Para realização da pesquisa, foram selecionados quatro sujeitos do sexo feminino, sendo dois oriundos dos Estados Unidos e dois do Brasil.

Os sujeitos brasileiros (doravante S1 e S2) eram paulistanas e pós-graduandas na PUC-SP. S1 é professora de português para estrangeiros e S2 é professora de inglês em institutos de idiomas.

Os sujeitos norte-americanos, doravante chamados de S3 e S4, faziam, na época da gravação, intercâmbio cultural na PUC-SP. Eles estavam no Brasil há seis meses e um ano, respectivamente. S3 é natural de Mineápolis e S4, do Havaí, portanto falantes de dialetos diferenciados.

De modo a reduzir a margem de diferenças de natureza sócio-linguística, optou-se por falantes do mesmo sexo.

4.3 CORPUS DE PESQUISA

Conforme alerta Xu (2010, p. 333), “when an experimental design does not include the right condition to make the encoding of a particular function obligatory, the associated prosodic pattern is not guaranteed to occur”.

Assim, para a montagem do *corpus* do experimento – um diálogo composto de seis sentenças –, fiz um levantamento de trabalhos que descrevem os sistemas prosódicos das línguas inglesa e portuguesa, buscando contemplar contextos de interesse para a investigação.

Optei pela leitura e interpretação do diálogo à fala espontânea por acreditar que a leitura fornece-me maior controle sobre os padrões que pretendo analisar e consegue preservar as funções interativas da fala³⁴.

Segue, em negrito, o *corpus* de análise:

(SITUAÇÃO: Duas amigas estão organizando uma festa de aniversário)

A: A party without games is not a good party.

Let's get Chris to organize the activities.

B: The children don't like her so much.

A: What about Sandra, then?

B: No, not Sandra. She couldn't do it.

4.4 GRAVAÇÃO DO *CORPUS*

Para a gravação dos dados, foi utilizado o microfone *Sennheiser e835*, mesa *Shune M267*, placa de áudio *Delta 44*, e *software SOUND Forge 6.0* em uma taxa de amostragem de 22,05 KHz.

Cada sujeito recebeu um *script*, lido e interpretado em ambiente acusticamente isolado nas dependências do Estúdio de Rádio e TV da PUC-SP, sob a supervisão de um técnico de som.

³⁴ Xu (2010), ao rebater críticas de que a fala ensaiada (*lab speech*) não tem funções interativas, afirma, por exemplo, que no diálogo "Who did she play in the movie?" "She played the PRINCESS in the movie", o acento frasal (foco) em *princess* é implementado automaticamente pelo próprio contexto "whether the context clue is a replay of a pre-recorded audio (XU e XU, 2005), spoken live by another interlocutor (XU e WANG, 2009) or spoken by the speaker him/herself (XU, 1999) e que a natureza da interação para o foco é "informational rather than interpersonal" (XU, 2010, p. 335).

Em um primeiro momento, os sujeitos foram instruídos a lerem o diálogo silenciosamente. Logo a seguir, cada dupla, composta de um sujeito brasileiro e de um sujeito norte-americano, leu o diálogo em voz alta, alternando-se os papéis. Foram gravadas as três repetições de cada um dos 6 enunciados, totalizando, assim, 72 gravações.

4.5 NOTAÇÃO DOS PADRÕES ENTOACIONAIS

Para a notação dos eventos tonais, foi utilizado o sistema ToBI (SILVERMAN *et al*, 1992) de transcrição, que tem sido amplamente usado na descrição de variedades da língua inglesa e que foi explicitado no capítulo 3 desta tese. Como apoio para anotação dos tons, recorreu-se ao dado acústico, utilizando-se como referência o contorno de frequência fundamental.

Na Figura 27, apresentamos a onda sonora, o contorno de f_0 e a transcrição de tons de uma das emissões do experimento:

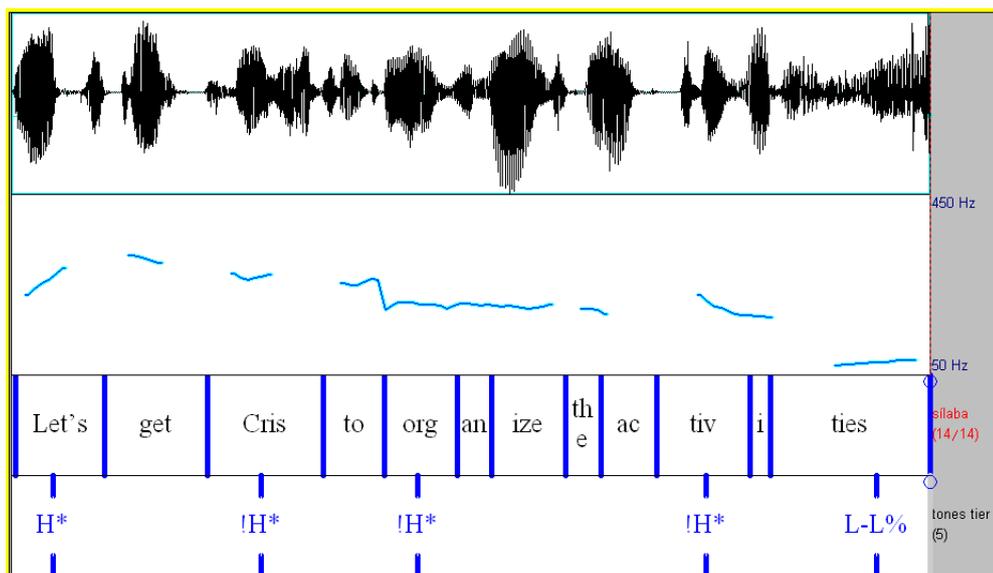


Figura 27 – Forma da onda, contorno de f_0 e transcrição de tons da primeira repetição de S4 do enunciado: *Let's get Cris to organize the activities*

4.6 ANÁLISE ACÚSTICA

4.6.1 Instrumentos, procedimentos e técnicas de análise

Para a análise acústica, empregou-se o *software* PRAAT (Boersma e Weenink, 2005), versão 4.1.3, de domínio público. A opção *pitch settings* do menu de ferramentas do programa PRAAT foi utilizada para ajustar a faixa de extensão relativa à frequência fundamental. Utilizou-se janelamento (*window length* em *Spectrum*, *Spectrogram Settings*) de 0.005 para se obter o espectrograma de banda larga. O computador que captou o registro acústico possuía as seguintes características: Pentium III, com 256Mb de memória RAM, 2Hd de 20Gb cada e processador 700MHZ. Da placa de som do computador que gerencia o estúdio, o *corpus* foi gravado diretamente em CD-ROM, de modo que cada grupo de repetições correspondesse a uma faixa.

As frases selecionadas para análise foram segmentadas a partir da forma da onda, tendo-se como referência o espectrograma de banda larga.

O espectrograma de banda larga é uma representação gráfica dos sons da fala. Nele, a ordenada corresponde à frequência; a abscissa, ao tempo; e a escala de cinzentos (produzindo contraste entre claro e escuro), à amplitude dos componentes espectrais. A frequência é medida em Hertz (Hz), a intensidade em decibéis (dB) e a duração em milissegundos (ms).

A Figura 28, a seguir, mostra a forma da onda, o espectrograma de banda larga e a segmentação da palavra *above*.

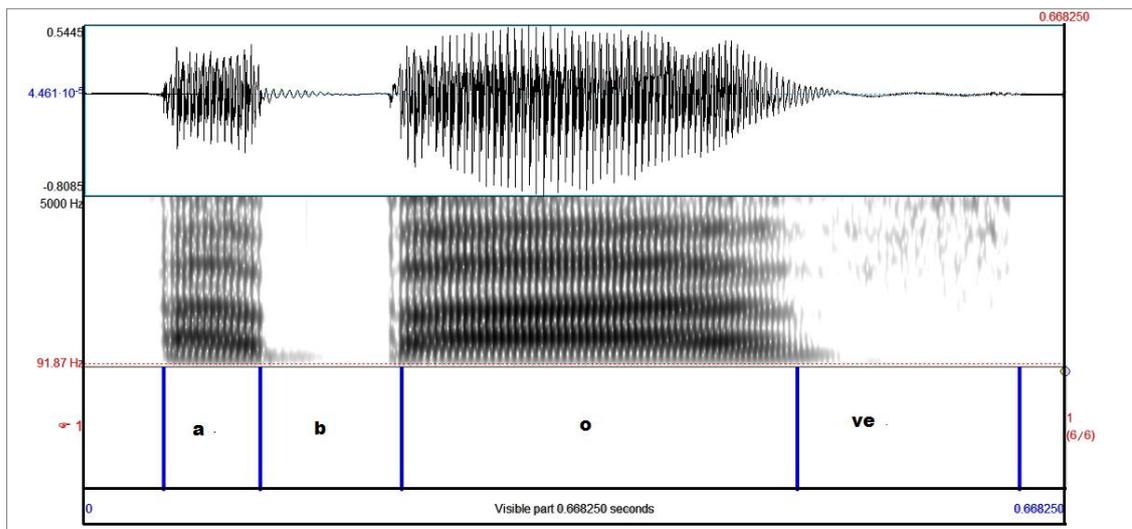


Figura 28 – Forma da onda, espectrograma de banda larga e segmentação da palavra *above*

Nem toda faixa de frequência faz-se importante na análise de um determinado dado acústico. O pesquisador dispõe de mecanismos para selecionar ou filtrar as faixas de frequência. O espectrograma de banda larga é gerado por larguras de banda de filtro relativamente largas, geralmente em torno de 300 Hz ou mais. Ele apresenta boa resolução no tempo, o que o faz uma boa opção para as medidas de duração.

Também possibilita boa visualização não apenas dos formantes (frequências de ressonância do trato vocal específicas dos sons, que, no espectrograma, apresentam-se como manchas escuras horizontais), mas também da transição entre os formantes. Por outro lado, o espectrograma de banda larga apresenta uma resolução grosseira em termos de frequência.

Ao observarmos o espectrograma de banda larga na Figura 28, percebemos que ele nos fornece boa visualização dos *formantes* da vogal (faixas de energia horizontais de tom escuro). Para tanto utilizei janelamento (recorte no tempo) pequeno, de 0,005 segundos. Dessa forma, obtive maior resolução no tempo, o que me permite trabalhar melhor o parâmetro de duração.

Para descrever as características fonéticas dos sons, o espectrograma de banda larga é muito útil, dada a variedade de informações sobre a produção dos sons que dele podemos inferir. Dentre essas informações, estão a barra de vozeamento (nas vogais e consoantes vozeadas); a plosão, correspondente ao ruído transiente nas plosivas como /p, b, t, d, k, g/, a fricção nas fricativas como /f, v, s, z, ʃ, ð/ correspondente ao ruído contínuo.

A identificação das características dos segmentos é crucial para a realização da segmentação, tarefa que realizamos para poder medir a duração dos segmentos fônicos, pois, como são coproduzidos na fala, não há limites precisos entre eles e é preciso delimitá-los com base em critérios que se baseiam no conhecimento de suas características acústicas e articatórias.

Os espectrogramas também oferecem evidência acústica para a tonicidade, pois permitem medir a duração dos segmentos e verificar a qualidade vocálica por meio da

medição dos formantes. Os correlatos acústicos do acento variam de língua para língua. Eles compreendem a duração, a frequência fundamental, a intensidade e a qualidade vocálica. Na Figura 29, a seguir, pode-se visualizar o espectrograma de banda larga com a sobreposição dos contornos de f_0 (em azul) e intensidade (em amarelo).

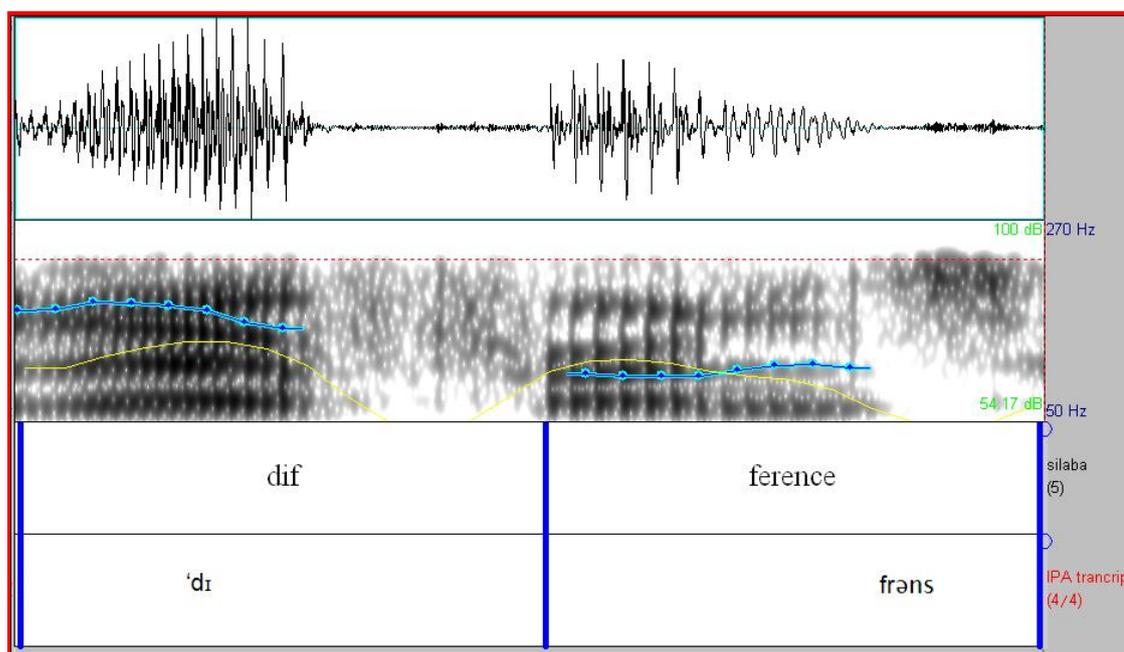


Figura 29 – Forma da onda, espectrograma de banda larga e camadas de segmentação das sílabas da palavra *difference* com transcrição ortográfica, na camada superior, e fonética, na camada inferior. Fonte: Silverman (1992)

As linhas verticais no espectrograma correspondem aos pulsos glotais que remetem às vibrações das pregas vocais. Na Figura 30, destacamos esses eventos:

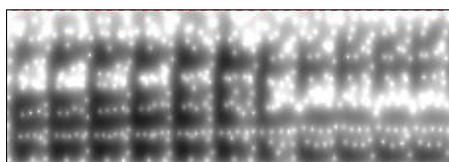


Figura 30 – Linhas verticais no espectrograma, representando os pulsos glotais

4.6.2 Procedimentos de segmentação e medição das unidades linguísticas

A segmentação das consoantes e vogais foi feita na forma da onda (oscilograma), porém considerando-se as informações do espectrograma de banda larga. Foram medidas, em milissegundos, a duração total dos enunciados e as durações das seguintes unidades fonéticas: (1) sílabas, (2) pés rítmicos acentuais, (3) pés rítmicos silábicos, (3) vogais plenas em adjacência e (4) unidades V-Vs ou GIPCs (*Group Inter-Perceptual Center*)³⁵.

A Figura 31, a seguir, correspondendo à primeira emissão do primeiro enunciado do sujeito brasileiro ilustra como foi feita a segmentação desses dados no PRAAT:

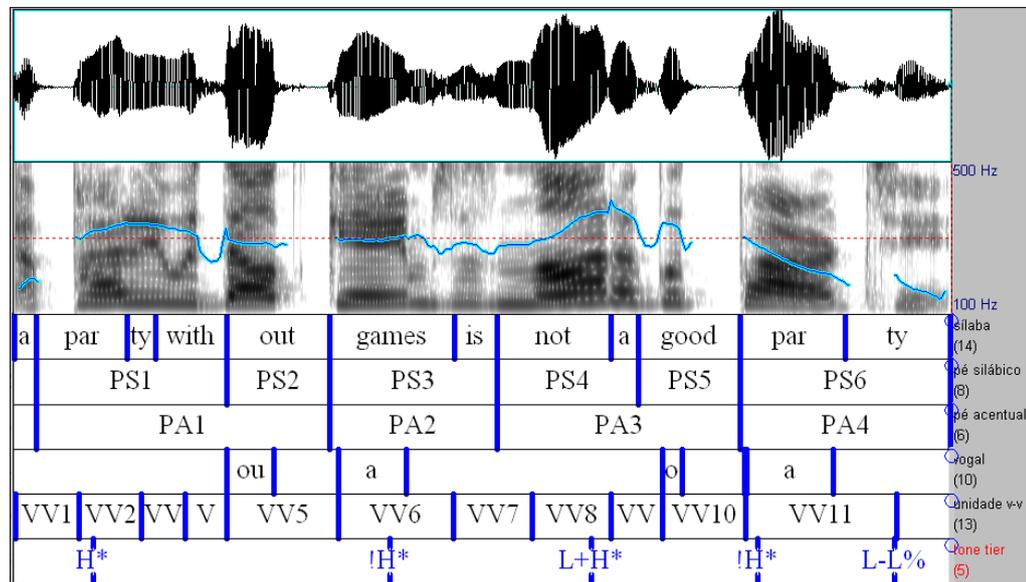


Figura 31 – Medidas de duração e marcação dos eventos tonais do primeiro enunciado produzido pelo sujeito brasileiro S1: *tier 1*: sílabas; *tier 2*: pés silábicos; *tier 3*: pés acentuais; *tier 4*: vogais plenas em adjacência; *tier 5*: unidade V-V; *tier 6*: eventos tonais

³⁵ O GIPC é uma unidade fonético-acústica que compreende o espaço delimitado pela vogal de uma sílaba (rima) e a vogal da próxima sílaba, incluindo as consoantes (entre as sílabas), mas excluindo a vogal da segunda sílaba. Sua importância nos estudos do ritmo das línguas decorre do fato de que o *onset* vocálico é considerado um ponto de saliência perceptual (teoria do *p-center*), pois a transição CV exerce um papel relevante no processamento periférico auditivo, uma vez que há predominância de neurônios primários respondendo a eventos do tipo *onset* (Barbosa, 2000).

As sílabas do enunciado foram divididas segundo alguns princípios de silabificação determinados pelo pesquisador e replicados para a segmentação de todos os enunciados produzidos pelos 4 sujeitos da pesquisa. No caso de plosivas em adjacência entre palavras (/t/ e /g/, como em *without games* e /d/ e /t/ como em *good party*), convencionou-se, para fins de padronização na segmentação das produções dos falantes, incluir na sílaba final da primeira palavra o intervalo de tempo até a liberação da oclusão da plosiva da sílaba inicial da segunda palavra.

Uma vez que os pés silábicos são determinados estruturalmente (combinação de sílaba com vogal plena seguida de sílabas com vogais reduzidas, quando houver), o número de pés rítmicos em cada enunciado é o mesmo, independentemente do sujeito.

Os pés rítmicos acentuais, por sua vez, variam de acordo com o sujeito, posto que eles são implementados em função da interpretação que cada falante fez na leitura do enunciado.

Uma vez que um dos fenômenos que se quer investigar é o alongamento silábico de sílabas com vogais plenas em adjacência, fez-se a medição da duração apenas dessas vogais, em relação à duração dos sintagmas onde ocorrem. Veja-se, por exemplo, na Figura 31, a delimitação de apenas 4 vogais: “ou” em *out* e “a” em *games*, na frase *without games*, e “oo” em *good* e “a” em *party*, na frase *good party*.

O número de unidades V-V em cada enunciado foi o mesmo para os 4 falantes nos enunciados de 1 a 5. Porém, no enunciado 3 (*The children don't like her so much*), S1 pronunciou a consoante /h/ na palavra de função gramatical *her*, não implementando, assim, a junção [ˈlaɪkə], da coda silábica em *like* /k/ e a rima de *her* /ə/, resultando nas

unidades V-Vs [aɪk] e [əʊs]. Na produção de S1, cada palavra foi pronunciada separadamente [ˈlaɪk] e [hər], resultando nas unidades V-V [aɪkh] e [ərs].

No enunciado 6 (*She couldn't do it*), entretanto, são observadas diferenças, entre os falantes, no número de unidades V-V. Essas diferenças ocorrem por conta da inserção, por parte de S1, da vogal /ə/ em *couldn't*, produzida como [ˈkudənt], e não [kudnt], como pelos demais sujeitos. Assim, enquanto para S1 observam-se duas unidades V-V [ˈʊd] e [əntd], para os demais sujeitos observa-se apenas uma [ʊdntd], resultando em uma unidade a menos na contagem do número de unidades V-V do enunciado.

4.7 ANÁLISE ESTATÍSTICA

As medições da duração das unidades linguísticas foram submetidas à análise de variância (ANOVA). Foram também verificados índices de variabilidade, dividindo-se o desvio padrão pela média de duração dos segmentos.

CAPÍTULO 5 RESULTADOS, ANÁLISE E INTERPRETAÇÃO DOS DADOS

Neste capítulo, abordaremos os resultados da análise dos padrões rítmicos e entoacionais dos enunciados e os discutiremos.

5.1 ANÁLISE DOS PADRÕES RÍTMICOS: RESULTADOS E DISCUSSÃO

De forma geral, os quatro falantes atribuíram proeminência às palavras relevantes no discurso. Houve variedade de pés ritmos acentuais, mostrando que cada falante não apenas leu a sua parte do diálogo, mas também a interpretou. A seguir, destacarei alguns contrastes entre as produções de cada unidade de análise que são relevantes para se entender os processos envolvidos.

5.1.1 Unidades V-V

Foram constatadas diferenças significativas entre as produções de unidades V-Vs, segregando o grupo das brasileiras e o das norte-americanas, em relação aos enunciados 2, 3, 4 e 5.

Em relação ao enunciado 2 (*Let's get Cris to organize the activities*), foram encontradas diferenças significativas entre as produções de fala das norte-americanas e das brasileiras em relação às seguintes unidades V-V: V-V1 “etsg” ($p=,010$), V-V3 – “ist” ($p=,010$), V-V4 – “oorg” ($p=,022$), e V-V9 – “it” ($p=,005$). As medidas de duração médias relativas podem ser observadas no Gráfico 5:

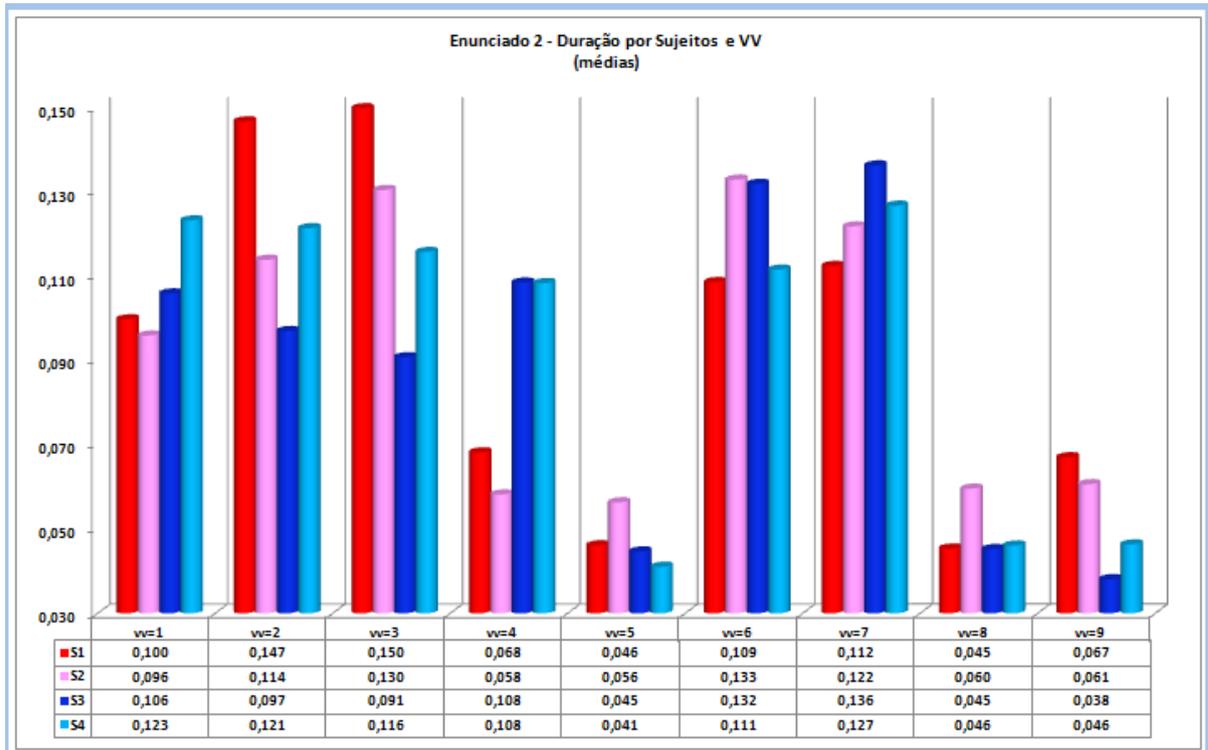


Gráfico 5 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 2

A unidade “etsg” apresentou-se mais longa nas produções de fala das norte-americanas. A unidade V-V “ist”, por sua vez, apresentou-se mais longa na produção das brasileiras. Esse alongamento pode ser explicado por conta de diferenças segmentais e acentuais: nas emissões das brasileiras, houve substituição da vogal relaxada /ɪ/ pela tensa /i/ e “Chris” recebeu o acento frasal. Em português, o correlato acústico de duração é o mais relevante nas proeminências frasais (BARBOSA, 2000).

A unidade “oorg”, por sua vez, foi mais longa nas produções das norte-americanas, que implementaram o acento lexical na primeira sílaba de “**organize**” e as brasileiras, na última (“**organize**”). Por fim, a unidade “it” foi mais longa nas produções das brasileiras, por nem sempre produzirem o flepe em “activities” ou substituírem a vogal relaxada /ɪ/ pela tensa /i/.

No enunciado 3 (*The children don't like her so much*), foram encontradas diferenças significativas entre as produções de fala das norte-americanas e das brasileiras em relação às seguintes unidades V-V: V-V1 “ech” ($p=,012$), V-V2 “ildr” ($p=,000$), V-V5 “ikh” ($p=,009$) e V-V 7 “om” ($p=,021$). As medidas de duração média relativas podem ser observadas no Gráfico 6:

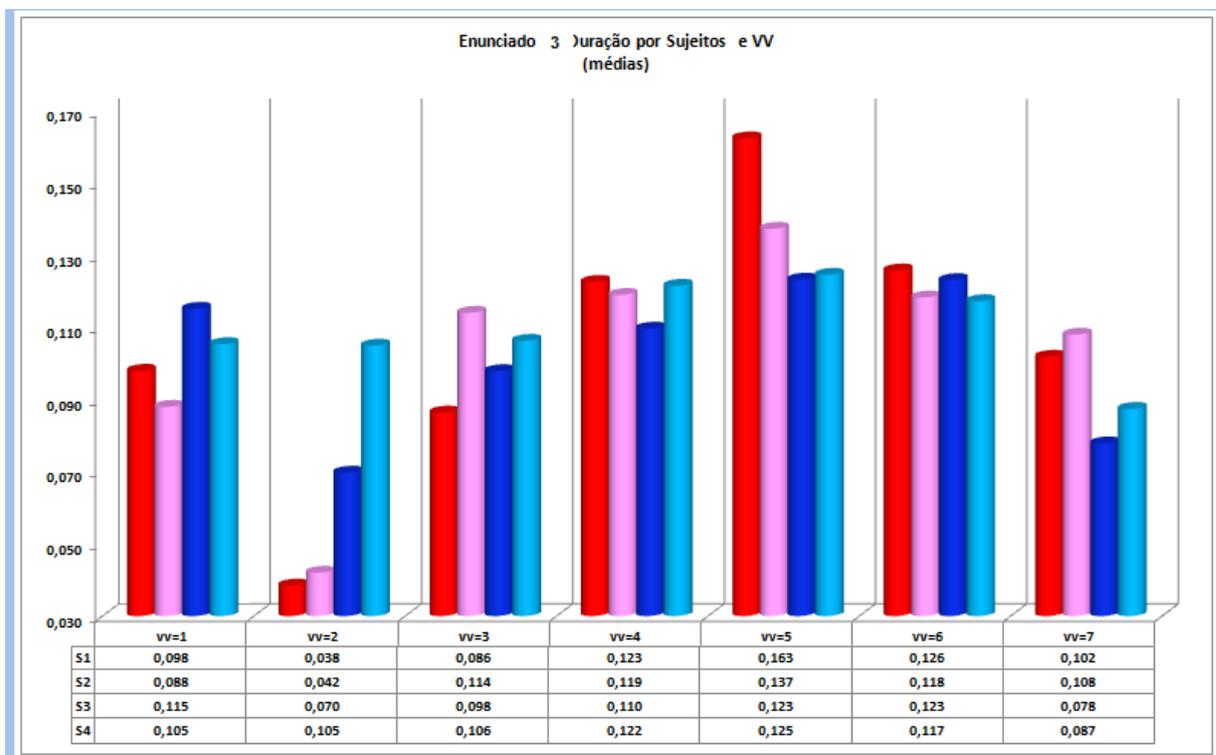


Gráfico 6 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 3

As unidades “ech” e “ildr” foram mais longas nas produções das norte-americanas. A unidade “ikh”, por sua vez, mostrou-se mais longa nas produções das brasileiras, por não elidirem sistematicamente a consoante fricativa surda /h/ de “her”.

No enunciado 4 (*What about Sandra, then?*), foram encontradas diferenças significativas entre as norte-americanas e brasileiras nas seguintes unidades V-V:

V-V2 “ab” ($p=,001$) e V-V3 “outs” ($p=,033$). As medidas de duração relativas podem ser observadas no Gráfico 7:

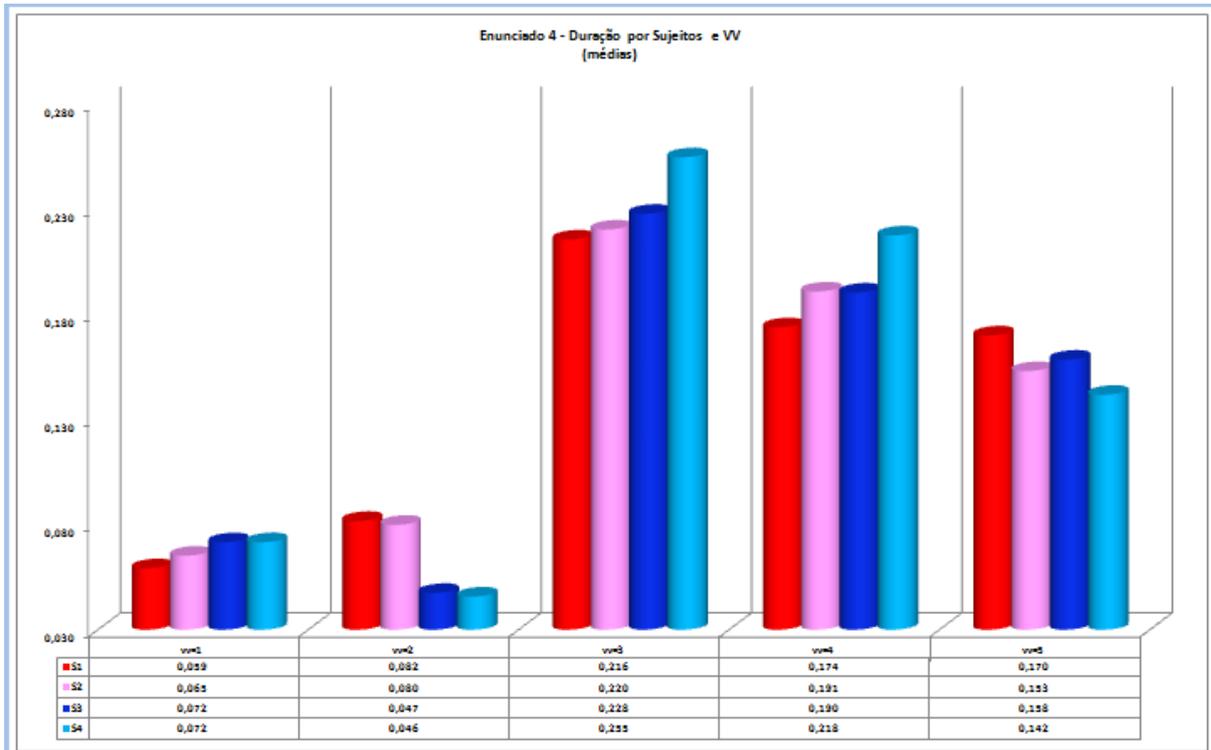


Gráfico 7 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 4

A unidade “ab” mostrou-se mais longa nas produções das brasileiras por conta da qualidade da vogal: as brasileiras não reduziram a pré-tônica “a” para o *schwa*.

A unidade “outs” foi mais longa nas produções das norte-americanas, devido ao contexto fonético (2 sílabas plenas em adjacência: a primeira sílaba recebe *extra length*), segundo a regra rítmica para o inglês norte-americano defendida por Bolinger (1986). As produções das brasileiras não apresentam esse alongamento.

No enunciado 5 (*No, not Sandra*), foram encontradas diferenças significativas entre as norte-americanas e as brasileiras nas seguintes unidades V-V: V-V1 “on” ($p=,050$)

e V-V2 “ots” ($p=,009$). As medidas de duração médias relativas podem ser observadas no Gráfico 8:

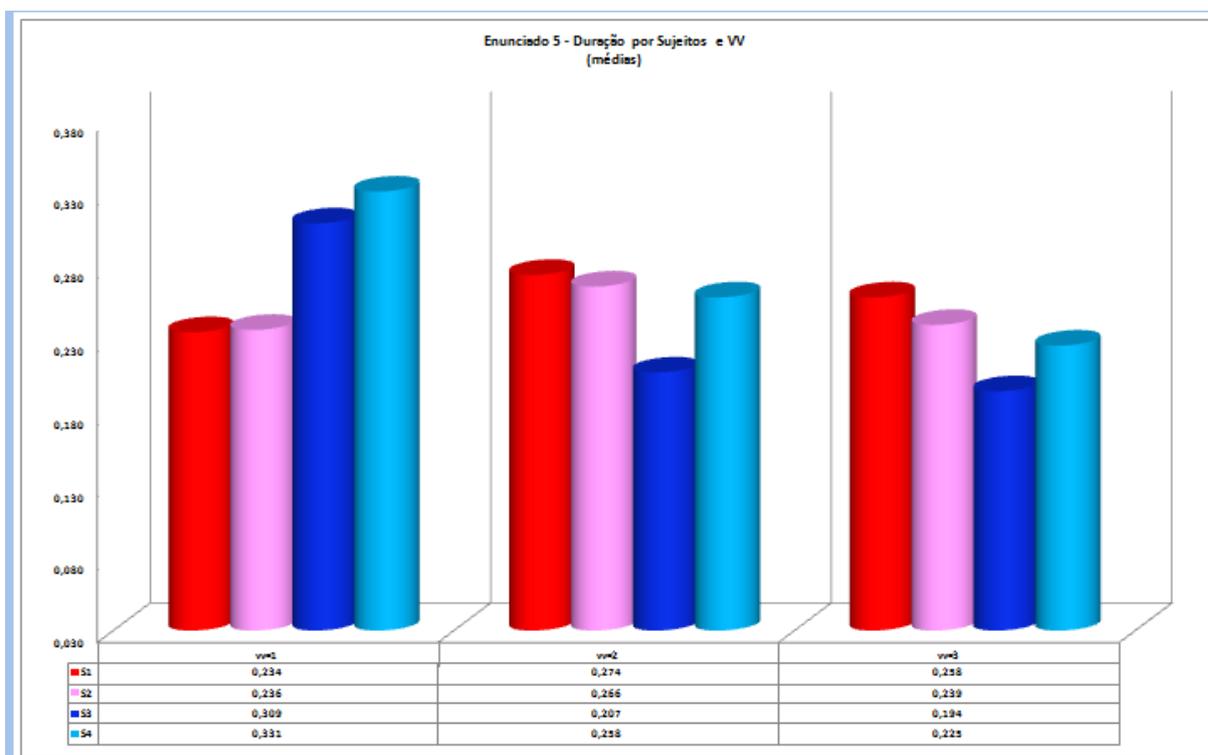


Gráfico 8 – Durações médias relativas nas unidades V-V por sujeitos para o enunciado 5

A primeira unidade “on” mostrou-se mais longa nas produções das norte-americanas por conta do alongamento do /o/ em posição de fronteira (L- e L-L%). A unidade “ots”, por sua vez, foi mais longa nas produções das brasileiras.

Vale ressaltar que o enunciado 5 foi o único no qual verificou-se o “encaixe” entre o ritmo silábico e acentual (na definição de ritmo silábico e ritmo acentual de Bolinger, 1986), pois todas as sílabas não proeminentes são compostas de vogais reduzidas e todas as sílabas plenas recebem o acento de *pitch*, conforme pode ser observado na Figura 32, a seguir:

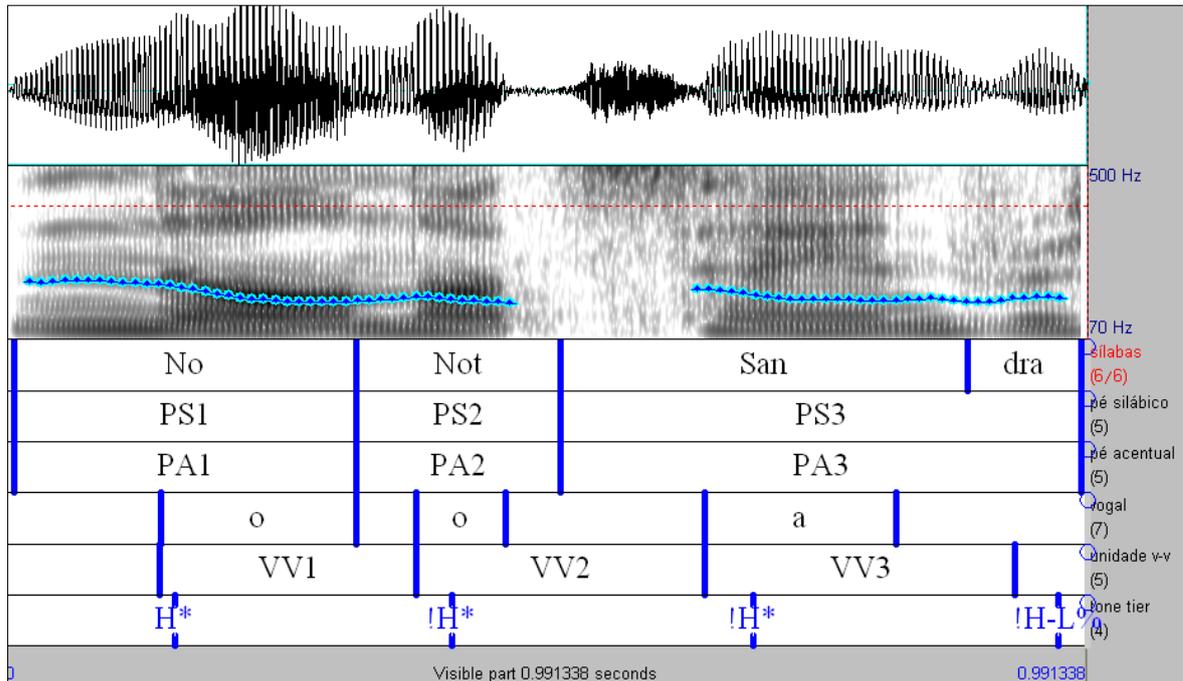


Figura 32 – “Encaixe” entre o ritmo silábico e o ritmo acentual na primeira repetição da emissão *No, not Sandra*, produzida pelo sujeito brasileiro S1

5.1.2 Sílabas

No enunciado 2, foram encontradas diferenças significativas entre as norte-americanas e as brasileiras nas seguintes sílabas: sil 2 “get” ($p=,001$), sil 5 “org” ($p=,048$), sil 6 “a” ($p=,035$), sil 8 “the” ($p=,015$), sil 10 “tiv” ($p=,017$) e sil 11 “i” ($p=,000$).

A sílaba “get” mostrou-se mais longa nas produções das brasileiras, que preservaram a qualidade vocálica de / ϵ /, enquanto as americanas produziram uma vogal reduzida. A sílaba “org”, por sua vez, mostrou-se mais longa nas americanas, pois é nela que se realiza o acento lexical, já na produção das brasileiras o acento lexical recaiu sobre “nize”.

A sílaba “a” foi mais longa nas produções das brasileiras, pois, enquanto as norte-americanas produziram o /ə/ ou simplesmente apagaram a vogal, as brasileiras produziram uma vogal nasalizada.

Em relação à sílaba “the”, ela foi mais longa nas produções das norte-americanas, que produziram o /i/ seguido do *glide* /j/ antes de “activities”. As brasileiras elidiram o /i/ e o *glide*.

A sílaba “tiv” foi mais longa nas brasileiras, que substituíram a vogal relaxada /ɪ/ pela tensa /i/. Por outro lado, a sílaba /i/ mostrou-se mais longa nas produções das norte-americanas, pois, enquanto essas produziram uma vogal reduzida, as brasileiras apagaram a vogal por ela estar em posição pós-tônica.

No enunciado 3 (*The children don't like her so much*), foram encontradas diferenças significativas entre as produções das norte-americanas e das brasileiras na sílaba 2 “chil” ($p=,002$), mais longa nas produções das norte-americanas.

5.1.3 Pés silábicos

Os pés silábicos são determinados estruturalmente (sílabas plenas seguidas de sílabas reduzidas, quando houver) e, portanto, o número de pés silábicos de todos os quatro sujeitos foi o mesmo para cada enunciado.

Pés silábicos com durações aproximadas podem indicar maior tendência à isocronia. A Tabela 12, a seguir, apresenta o índice de variabilidade de pés silábicos das

emissões dos quatro sujeitos. Quanto menor esse valor, mais próximas foram as medidas de duração de cada pé e maior a tendência à isocronia entre os pés.

	PS			
	Sujeito	Média	Desvio Padrão	Índice de Variabilidade
Enunciado 1	1	0,17	0,04	0,26
	2	0,17	0,05	0,31
	3	0,17	0,06	0,33
	4	0,17	0,06	0,36
	Total	0,17	0,05	0,31
Enunciado 2	1	0,14	0,08	0,58
	2	0,15	0,08	0,57
	3	0,14	0,08	0,59
	4	0,14	0,07	0,50
	Total	0,14	0,08	0,55
Enunciado 3	1	0,21	0,10	0,48
	2	0,20	0,07	0,34
	3	0,20	0,08	0,42
	4	0,20	0,08	0,40
	Total	0,20	0,08	0,41
Enunciado 4	1	0,25	0,13	0,51
	2	0,24	0,12	0,52
	3	0,26	0,09	0,34
	4	0,25	0,11	0,45
	Total	0,25	0,11	0,45
Enunciado 5	1	0,33	0,13	0,39
	2	0,33	0,17	0,52
	3	0,33	0,14	0,43
	4	0,32	0,11	0,35
	Total	0,33	0,14	0,41
Enunciado 6	1	0,33	0,25	0,76
	2	0,33	0,25	0,76
	3	0,34	0,28	0,83
	4	0,33	0,04	0,12
	Total	0,33	0,22	0,65

Tabela 12 – Índice de variabilidade dos pés silábicos

Assim, para cada enunciado, podemos considerar S4 como o falante que apresentou maior tendência à isocronia, isto é, valores mais aproximados, como pode ser sintetizado no quadro a seguir:

Enunciado	Sujeito cujas emissões apresentam maior tendência à isocronia dos pés rítmicos
1	S1
2	S4
3	S2
4	S3
5	S4
6	S4

Foram também observados dois pés silábicos, no enunciado 2 (*Let's get Chris to organize the activities*), que diferenciaram o grupo das brasileiras e o das norte-americanas: o pé silábico 2, que corresponde a palavra “get” ($p=,001$), e o pé silábico 5, “ize the” ($p=,030$).

Em relação ao pé silábico “ize the”, houve, principalmente nas produções da brasileira S2, alongamento desse pé por ela não ter atribuído à sílaba “ize” o acento lexical de “organize” e não ter coarticulado os sons /z/ (coda de “ize”) e /ð/ (ataque de “the”).

5.1.4 Pés acentuais

A Tabela 13, a seguir, mostra o número de pés acentuais nas emissões dos 4 sujeitos nas três repetições:

Enunciado	S1	S2	S3	S4
E1	10	9	9	12
E2	3	9	6	10
E3	4	8	9	9
E4	3	6	3	6
E5	5	9	9	9
E6	6	6	6	6
Total	31	47	42	52

Tabela 13 – Número de pés acentuais nas emissões dos 4 falantes

Observa-se, na Tabela 13, que o sujeito norte-americano S4 produziu o maior número de pés acentuais, e o sujeito brasileiro S1, o menor.

5.1.5 Vogais plenas em adjacência

A Tabela 14, a seguir, mostra as sílabas (e as vogais) em adjacência, que, segundo a regra rítmica fonológica do inglês defendida por Bolinger (1986), estariam sujeitas a receber alongamento extra por precederem outra sílaba plena:

Enunciado	Sílaba	Vogal
1	A party without games is not a good party	/ou/ e /u/
2	Let's get Chris to organize the activities	/ε/ e /ε/
3	The children don't like her so much	/ou/ e /ou/
4	What about Sandra, then?	/aʊ/
5	No, not Sandra	/ɑ/
6	She couldn't do it	/ʊ/

Tabela 14 – Vogais plenas em adjacência nos seis enunciados, sujeitas à regra de *extra lengthening* (BOLINGER, 1986)

Foram constatadas diferenças significativas nas medidas de duração, que segregaram as emissões das brasileiras e as das norte-americanas, em relação aos enunciados 4 e 5.

No enunciado 4, a vogal plena “ou” (em “about”, adjacente à vogal “a” em “Sandra”) segregou o grupo das norte-americanas do das brasileiras ($p=,000$). No enunciado 5, duas vogais plenas em adjacência, “o” (em “no” adjacente à vogal “o” em “not”) e “o” (em “not” adjacente à vogal “a” em “Sandra”), segregaram o grupo das americanas e o das brasileiras, ($p=,024$) e ($p=,000$), respectivamente.

Uma vez que, em todas as produções das norte-americanas e da brasileira S2 o enunciado foi dividido em duas frases intermediárias (No/not Sandra), não consideraremos a duração do primeiro /ou/ em *no*.

Assim, de um total de 8 vogais passíveis de receber esse alongamento, foram encontrados dois casos que diferenciaram as produções das norte-americanas das

produções das brasileiras, o que sugere que essa característica do inglês, conforme aponta Bolinger (1986), foi incorporada pelas aprendizes brasileiras na maioria dos contextos.

5.2 ANÁLISE DOS PADRÕES ENTOACIONAIS: RESULTADOS E DISCUSSÃO

Os resultados da análise perceptiva dos padrões entoacionais segundo o sistema de notação ToBI referentes aos contornos entoacionais das três repetições (R1, R2, R3) dos 6 enunciados (E1, E2, E3, E4, E5, E6) pelos 4 sujeitos de pesquisa brasileiros (S1 e S2) e norte-americanos (S3 e S4) serão apresentados nas tabelas a seguir:

Enunciado 1 – Sujeitos brasileiros (S1 e S2):

	A	par	ty	with	out	games	is	not	a	good	par	ty	
S1R1		H*				!H*		L+H*			!H*		L-L%
S1R2		H*				!H*				H*			L-L%
S1R3		L+H*				!H*		L+H*					L-L%
S2R1		L+H*				!H*		L+H*					L-L%
S2R2		L+H*				!H*		L*+H					L-L%
S2R3		H*				!H*		H*					L-L%

Tabela 15 – Eventos tonais nas 3 repetições do enunciado 1, produzidas pelos sujeitos brasileiros (S1 e S2)

Observa-se, na Tabela 15, que, na primeira repetição (R1) de S1, há 4 *pitch accents*, sendo que o último recai sobre a sílaba tônica da segunda ocorrência da palavra

“party”. Em R2, o foco realiza-se em “good” e, em R3, em “not”. Na produção de S2, por sua vez, observam-se três *pitch accents*, com foco estreito em “not”. O contorno global de *pitch* (f_0) do enunciado é descendente.

Enunciado 1 – Sujeitos norte-americanos (S3 e S4):

	A	par	ty	with	out	games	is	not	a	good	par	ty	
S3R1		H*				!H*		!H*					L-L%
S3R2		H*				!H*		!H*					L-L%
S3R3		H*				L*+H		L*+H					L-L%
S4R1		H*				!H*		!H*			!H*		L-L%
S4R2		H*				!H*		!H*			!H*		L-L%
S4R3		H*				!H*		!H*			!H*		L-L%

Tabela 16 – Eventos tonais nas 3 repetições do enunciado 1, produzidas pelos sujeitos norte-americanos (S3 e S4)

Observa-se, na Tabela 16, que, nas emissões de S3, ocorrem 3 *pitch accents* com foco em “not”. Em S4, por sua vez, há 4 *pitch accents*, cada um em um nível de *pitch* mais baixo que o anterior, caracterizando o *downstep*. Ocorre foco estreito em “not” na produção do S3, conferindo conotação incisiva, enquanto o foco largo e o *downstep* em S4 atribuem ao enunciado um caráter peremptório.

Enunciado 2 – Sujeitos brasileiros (S1 e S2):

	Let's	get	Chris	to	org	na	ize	the	ac	tiv	i	ties	
S1R1			L+H*										L-L%
S1R2			L+H*										L-L%
S1R3			L+H*										L-L%
S2R1	H*		!H*				L+!H*						L-L%
S2R2	H*		!H*				!H*						L-L%
S2R3	H*		L+H*				L+!H						L-L%

Tabela 17 – Eventos tonais nas 3 repetições do enunciado 2, produzidas pelos sujeitos brasileiros (S1 e S2)

Nota-se, na Tabela 17, que o tom L+H* em “Chris” e a ausência de *pitch accents* posteriores nas três repetições de S1 sugerem que “Chris” é informação nova ou contraste, como refere Pierrehumbert e Hirschberg (1990). Em S2, o foco em “ize” nas duas primeiras repetições e o acento de *pitch* em “ize” demonstram ter havido interferência da L1, o que pode acarretar consequências na estrutura rítmica da L2.

Enunciado 2 – Sujeitos norte-americanos (S3 e S4):

	Let's	get	Chris	to	org	na	ize	the	ac	tiv	i	ties	
S3R1	H*		!H*										L-L%
S3R2	H*		H*										L-L%
S3R3	H*		!H*										L-L%
S4R1	H*		!H*		!H*					!H*			L-L%
S4R2	H*		!H*		!H*					!H*			L-L%
S4R3	H*		H*										L-L%

Tabela 18 – Eventos tonais nas 3 repetições do enunciado 2, produzidas pelos sujeitos norte-americanos (S3 e S4)

Veja-se, na Tabela 18, a ocorrência do foco em “Chris” nas três repetições de S3. Diferentemente das brasileiras, que utilizaram um *pitch accent* bitonal (L+H*), o foco realiza-se com um tom (H*). Em S4, por sua vez, há 4 *pitch accents*, em R1 e R2, cada um em um nível de *pitch* mais baixo que o anterior (*downstep*), indicando o foco largo. Em R3, entretanto, o último *pitch accent* recai sobre “Chris”, sugerindo que “organize the activities” é subentendido e “Chris” é a informação que se quer destacar.

Enunciado 3 – Sujeitos brasileiros (S1 e S2):

	The	child	ren	don't	like	her	so	much	
S1R1					L*+H				L-L%
S1R2		H*			L*+H				L-L%
S1R3					L+H*				L-L%
S2R1		H*		L*	H*				L-L%
S2R2		H*		L*	H*				L-L%
S2R3		H*		H*					L-L%

Tabela 19 – Eventos tonais nas 3 repetições do enunciado 3, produzidas pelos sujeitos brasileiros (S1 e S2)

A Tabela 19 mostra que, nas emissões de S1, há foco estreito em “like” nas três repetições. S2, por sua vez, procurou destacar a palavra “don't” de duas formas: pela inflexão de *pitch*, com o tom L* medial, em R1 e R2, e pelo tom H* final, em R3.

Enunciado 3 – Sujeitos norte-americanos (S3 e S4):

	The	child	ren	don't	like	her	so	much	
S3R1		H*		!H*	!H*				L-L%
S3R2		H*		!H*	!H*				L-L%
S3R3		H*		!H*	!H*				L-L%
S4R1		H*		!H*	!H*				L-L%
S4R2		H*		!H*	!H*				L-L%
S4R3		H*		!H*	!H*				L-L%

Tabela 20 – Eventos tonais nas 3 repetições do enunciado 3, produzidas pelos sujeitos norte-americanos (S3 e S4)

Os dados da Tabela 20 demonstram que, nas produções das norte-americanas, houve 3 acentos de *pitch* com *downstep*. Chama a atenção a atribuição de *pitch accents* sucessivos em “don't” e “like” em todas as emissões.

Enunciado 4 – Sujeitos brasileiros (S1 e S2):

	What	a	bout	San	dra	then	
S1R1				L*+H			L-L%
S1R2				L*+H			L-L%
S1R3				L*+H			L-L%
S2R1	H*			!H*			L-L%
S2R2	H*			L+!H*			L-H%
S2R3	H*			L+!H*			L-L%

Tabela 21 – Eventos tonais nas 3 repetições do enunciado 4, produzidas pelos sujeitos brasileiros (S1 e S2)

Na Tabela 21, observa-se que S1 não atribuiu proeminência a “What” e conferiu o foco estreito à sílaba tônica de “Sandra” por meio do acento bitonal L*+H. Em S2, por sua vez, “What” recebe o primeiro acento de *pitch* seguido de “San” em um tom mais baixo. Em R2, observa-se o tom de fronteira final ascendente (L-H%), o que faz com que a sugestão soe menos incisiva.

Enunciado 4 – Sujeitos norte-americanos (S3 e S4):

	What	A	bout	San	dra	then?	
S3R1				L+ H*			L-L%
S3R2				H*			L-L%
S3R3				L+H*			L-L%
S4R1	H*			!H*			L-L%
S4R2	H*			!H*			L-H%
S4R3	H*			!H*			L-L%

Tabela 22 – Eventos tonais nas 3 repetições do enunciado 4, produzidas pelos sujeitos norte-americanos (S3 e S4)

Analisando a Tabela 22, observa-se que S3 não atribuiu proeminência à “What” e conferiu o foco estreito à sílaba tônica de “Sandra” por meio dos tons L+H* e H*. Embora esse padrão seja parecido com o utilizado por S1, os tons não são os mesmos: em S1, a subida de *pitch* coincidiu com o início da sílaba tônica (L*+H) e, em S3, com o final (L+H*). Em S4, por sua vez, nota-se o acento em “What” e o *downstep*, e, em R2, o tom de fronteira final ascendente L-H%.

Enunciado 5 – Sujeitos brasileiros (S1 e S2):

	No		not	San	dra	
S1R1	H*		!H*	!H*		L-L%
S1R2	H*					L-L%
S1R3	H*					L-L%
S2R1	H*	L-	H*	!H*		L-L%
S2R2	H*	L-	L*+H	!H*		L-H%
S2R3	H*	L-	H*	!H*		L-L%

Tabela 23 – Eventos tonais nas 3 repetições do enunciado 5, produzidas pelos sujeitos brasileiros (S1 e S2)

Nas emissões de S1 (Tabela 23), observam-se 3 acentos de *pitch* decrescentes (*downstep*) na primeira repetição e o foco estreito em “no” em R2 e R3. Nas emissões de S2, por sua vez, há 3 acentos de *pitch* e duas frases intermediárias: *no/not Sandra*.

Enunciado 5 – Sujeitos norte-americanos (S3 e S4):

	No		not	San	dra	
S3R1	H*	L-	H*	!H*		L-L%
S3R2	H*	L-	H*	!H*		H-L%
S3R3	H*	L-	H*	!H*		H-L%
S4R1	H*	L-	H*	!H*		L-L%
S4R2	H*	L-L%	H*	!H*		L-L%
S4R3	H*	L-	H*	!H*		L-L%

Tabela 24 – Eventos tonais nas 3 repetições do enunciado 5, produzidas pelos sujeitos norte-americanos (S3 e S4)

Nas emissões de S3 (Tabela 24), observam-se duas frases intermediárias (*No/not Sandra*) e o tom de fronteira final H-L% em R2 e R3. Nas de S4, notam-se duas frases intermediárias em R1 e R3 (*No/not Sandra*) e dois sintagmas entoacionais (*No//not Sandra*) em R2.

Enunciado 6 – Sujeitos brasileiros (S1 e S2):

	She	coudn't	do	it	
S1R1		H*	L+H*		L-L%
S1R2		H*	!H*		L-L%
S1R3		H*	H*		L-L%
S2R1		H*	!H*		L-L%
S2R2		L+H*	!H*		L-L%
S2R3		L+H*	L+!H*		L-L%

Tabela 25 – Eventos tonais nas 3 repetições do enunciado 6, produzidas pelos sujeitos brasileiros (S1 e S2)

Nas emissões de S1 (Tabela 25), observam-se 2 acentos de *pitch*, com *downstep* apenas em R2. S2 também implementa dois acentos, mas com *downstep* nas 3 repetições.

Enunciado 6 – Sujeitos norte-americanos (S3 e S4):

	She	coudn't	do	it	
S3R1		H*	!H*		L-L%
S3R2		H*	!H*		L-L%
S3R3		H*	!H*		L-L%
S4R1	H*		!H*		H-L%
S4R2	H*		!H*		L-L%
S4R3	H*		!H*		L-L%

Tabela 26 – Eventos tonais nas 3 repetições do enunciado 6, produzidas pelos sujeitos norte-americanos (S3 e S4)

Nas 3 repetições de S3 (Tabela 26), observa-se *downstep* e atribuição de *pitch accents* em “couldn’t” e “do”. Por outro lado, em S4 chama a atenção o acento de *pitch* na palavra de função gramatical “she” e um intervalo maior na atribuição dos acentos, se comparado aos demais falantes.

Em suma, nas produções das norte-americanas, observou-se a ocorrência de forma mais sistemática do fenômeno de *downstep*, e nas produções das brasileiras observou-se a ocorrência produtiva dos tons L*+H e L+H*. O tom L*+H é referido por Pierrehumbert e Hirschberg (1990) como enfático e o L+H*, como contrastivo ou veiculando um dado novo. Nas produções da norte-americana S4, constatou-se maior regularidade nas atribuições das proeminências acentuais.

A atribuição de *pitch accents* em palavras adjacentes ocorreu nas produções dos 4 sujeitos nos seguintes contextos: enunciado 3: “don’t like”; enunciado 5: “not Sandra”. No enunciado 6, “couldn’t do”, ela foi observada nas produções de S1, S2 e S3, mas não foi observada na fala de S4. Os dados sugerem haver diferenças

dialetais entre S3 e S4: enquanto S3 tende a atribuir foco estreito aos enunciados, S4 atribui foco largo.

Além disso, nas produções de S4, o espaçamento entre os acentos de *pitch* apresenta maior regularidade, o que, segundo o princípio de alternância rítmica, revela tendência nas produções de falantes nativos do inglês de “not only space out the accents but to do it more rhythmically” (BOLINGER, 1986, p. 63).

Os dados também revelam diferenças entre as produções de fala das brasileiras e das norte-americanas na implementação do foco estreito. O foco estreito em palavras, nas produções de fala das brasileiras, recaiu, por vezes, em sílabas átonas, como fez S2 ao atribuir proeminência à sílaba “ize” de “organize” em vez de na sílaba “org”. A mesma falante, ao enfatizar apenas a contração “don’t” na frase “don’t like”, sem atribuir acento de *pitch* a “like”, altera o sentido do enunciado, pois expressa uma correção de uma afirmação anterior de seu interlocutor e não apenas o fato de as crianças não gostarem da Chris.

Nas produções de S1 e S2, o *dowstep* não ocorreu com frequência, uma vez que o último acento de *pitch* do enunciado tende a ser mais alto que o anterior. Segue que as brasileiras não produzem o foco estreito da mesma forma que as norte-americanas. O foco largo, por sua vez, foi usado com mais frequência pelas norte-americanas.

CONSIDERAÇÕES FINAIS

Com o desenvolvimento deste trabalho, espero contribuir para aprofundar conhecimentos sobre o ritmo e a entoação no contexto de ensino de inglês para brasileiros.

Neste estudo, investiguei como foram implementadas, por falantes norte-americanos e aprendizes brasileiros de inglês, as seguintes unidades rítmicas: unidades V-V, sílabas, pés silábicos, pés acentuais e vogais plenas em adjacência. Também investiguei como os falantes diferenciaram-se na marcação das proeminências acentuais por meio de acento de *pitch*, bem como na escolha dos tons e na implementação do foco estreito.

Em relação às sílabas e às unidades V-V, constatei que transferências relacionadas à tonicidade (como em “organize”), aos segmentos (substituição de /ɪ/ por /i/ em “Chris” e em “it”), à redução vocálica (não implementação do *schwa* em “about” e “children”), à elisão segmental e à juntura (como em [laɪkə] em “like her”) afetaram a estrutura rítmica dos enunciados.

No tocante aos pés rítmicos, observei haver menor índice de variabilidade dos pés silábicos e maior número de pés acentuais nas produções de fala do sujeito norte-americano S4. Os dados sugerem maior tendência à isocronia acentual nas produções desse sujeito e possíveis diferenças dialetais na fala dos sujeitos norte-americanos.

Quanto ao alongamento da sílaba plena motivado por outra sílaba plena imediatamente posterior, constatei que, de um grupo de oito combinações possíveis, em apenas dois casos a duração dessa sílaba mostrou-se diferenciada, segregando o grupo das brasileiras e o grupo das norte-americanas. Os dados sugerem que essa característica do inglês foi, na maioria dos contextos, incorporada pelas falantes brasileiras.

Em relação à entoação, nas emissões das norte-americanas, observou-se de forma mais sistemática a ocorrência do fenômeno de *downstep* e, nas produções das brasileiras, a ocorrência produtiva dos tons L^*+H e $L+H^*$, o que pode passar certas atitudes diferenciadas. Como apontam Pierrehumbert e Hirschhberb (1990), $L+H^*$ veicula um contraste ou uma informação nova, enquanto o L^*+H é entendido como enfático. Os autores também apontam que L^*+H é relativamente mais raro comparado com o H^* ou o $L+H^*$. Ao utilizar o L^*+H em vez de H^* , sentidos diferenciados são veiculados pelo falante. Além disso, ressaltam-se questões de alinhamento do pico de f_0 que codificam sentidos de ordem pragmática (KOHLENER, 2006). Os dados também apontam para diferenças na implementação do foco estreito, mais frequentes na fala das brasileiras.

Verificou-se, igualmente, maior regularidade na marcação das proeminências acentuais nas emissões da norte-americana S4, na medida em que ela apresentou o menor índice de variabilidade duracional dos pés rítmicos silábicos e maior número de pés rítmicos acentuais.

Em relação à atribuição de *pitch accents* em sílabas adjacentes, ela ocorreu nas produções de todos os sujeitos no enunciado 3 em “don’t like” e, no enunciado 5, em “not Sandra”. No enunciado 6, ela foi observada em “couldn’t do” nas produções de S1, S2 e S3, mas não foi verificada nas produções de S4.

A pesquisa realizada nesta tese demandou vasta leitura sobre modelos prosódicos e características prosódicas da língua inglesa. Ao refletir sobre certos fundamentos da literatura revisada e sobre os resultados da análise efetuada, alguns pontos merecem ser destacados quando considerada a prática docente no contexto de ensino de língua inglesa para brasileiros.

Em consonância com a afirmação de Bolinger (1986) de que o ritmo do inglês apresenta características acentuais e silábicas, este trabalho aponta para a importância da proposição de exercícios de percepção dos *pitch accents* (nível acentual) e das sílabas plenas (nível silábico), bem como dos processos de redução e elisão segmental. No tocante aos aspectos melódicos, o estudo depreendeu evidências a favor de atividades que sensibilizem os aprendizes sobre o *downstep*, recorrente nas descrições da língua inglesa e nos dados desta tese, e os alertem sobre os sentidos de ordem pragmática que são veiculados pela transferência de certos tons produtivos no português, caso do L+H*.

Para o ensino dos padrões melódicos e das características das sílabas plenas e reduzidas aqui referidas, o recurso a instrumentais de visualização do sinal da fala podem ser úteis. Como professor de Fonética do Inglês e de Língua Inglesa, tenho introduzido esses instrumentais e obtido resultados favoráveis.

A música é, também, uma forma eficaz de introduzir a conceituação de entoação. Ao cantar uma música, o aluno percebe a alternância dos tons graves e agudos, que acusticamente correspondem às frequências baixas e altas.

Por fim, deve-se pensar em exercícios de alternâncias rítmicas (sílabas plenas e reduzidas), percepção de tons e de avaliação dos efeitos de sentido desses tons a partir de materiais discursivamente relevantes, tais como diálogos extraídos de filmes e séries em língua inglesa.

Cada língua e dialeto tem a sua própria musicalidade. Aprender uma língua estrangeira é aprender uma nova melodia.

REFERÊNCIAS BIBLIOGRÁFICAS

ABAUURRE, M. B. (1993). Fonologia: a gramática dos sons. *Revista Letras*. Santa Maria, UFSM, 5, p. 9-24.

ABERCROMBIE, D. (1967). *Elements of General Phonetics*. Edinburgh University Press Edinburgh.

AGUENA, A. M. (2006). *Contribuições da fonética acústica para a formação do professor de inglês: um estudo de caso*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

ALLEN, D.; HAWKINS, S. (1978). The development of phonological rhythm. In: BELL, A.; HOPPER, J. (eds.). *Syllables and Segments*. p. 173-5. Amsterdam: North-Holland.

ANDRADE, E. C. (2009). *Análise de unidades de ensino de inglês como L2 com ênfase nos aspectos orais*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

BÄNZIGER, T.; SCHERER, K. R. (2004). The role of intonation in emotional expressions. *Speech and Communication*, 46, p. 252-67.

BARBOSA, P. A. (2000). Syllable timing in Brazilian Portuguese: uma crítica a Roy Major. *D.E.L.T.A*, 16 (2), p. 369-402.

BARBOSA, P. A. (2006). *Incursões em torno do ritmo da fala*. Campinas: Pontes.

BARBOSA, M. (2013). *Questões de pronúncia e inteligibilidade em contexto de inglês como língua estrangeira*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

BARRY, W. J.; B.ANDREEVA, M.; RUSSO, S.; DIMITROVA, E. T.; KONTANTINOVA, (2003). Do rhythm measures tell us anything about language type?. *Proceedings of the 15th ICPHS*, Barcelona. p. 2693-6.

BECKMAN, M. E. (1992). Evidence for speech rhythm across languages. In: TOHKURA, Y.; VATIKIOTIS-BATESON, E.; SAGISAKA, Y. (eds.) *Speech perception, production and linguistic structure*, p. 457-63 Amsterdam: IOS Press.

BENQUEREL, A. P.; D'ARCY, J. (1986). Time-warping and the perception of rhythm in speech. *Journal of phonetics*, 14 (2), p. 231-46.

BLACK, B. (2011). Rhythm, Meter, and Scansion Made Easy disponível em: <http://gs.riverdale.k12.or.us/~bblack/meter.html>.

BOERSMA, P.; WEENINK, D. (2005). *Praat: doing phonetics by computer*. University of Amsterdam. Amsterdam, The Netherlands.

BOLINGER, D. (1958). A theory of pitch accent in English. *Word*, 13, p. 246-55.

BOLINGER, D. (1964). Around the edge of language: intonation. *Harvard Educational Review*, 34 (2), p. 282-93, Harvard College.

BOLINGER, D. (1978). Intonation Across Languages. In: GREENBERG, J.H. *Universals of Human language*, vol. 1. Stanford: Stanford University press. Stanford, California.

BOLINGER, D. (1981). *Two kinds of vowels, two kinds of rhythm*. Indiana University Linguistic Club.

BOLINGER, D. (1986). *Intonation and its parts*. Stanford: Stanford University Press. Stanford, California.

BOLINGER, D. (1998). Intonation in American English. In: HIRST, D.; DE CRISTO, A. (eds.). *Intonation Systems: a survey of twenty languages*, p. 47 Cambridge: Cambridge University Press.

BONGAERTS, T.; VAN SUMMEREN, C.; PLANKEN, B.; SCHILS, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language, *Studies in Second Language Acquisition*, 19, p. 447-65.

BRAZIL, D. (1975). *Discourse Intonation*. Birmingham: English language research monologs.

BREND, R. (1975). Male-female intonation patterns in American English. In: THORNE, B.; HENLEY, N. (eds.). *Language and Sex: Difference and Dominance*, p 84-7. Rowley, Mass.: Newbury House.

BROWMAN, C. P.; GOLDSTEIN, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In: KINGSTON, J.; BECKMAN, M. E. (eds.). *Papers in Laboratory phonology 1: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press.

BUUREN, L. V. (2000). Teaching rules for English rhythm/timing. *Colloquium of the British Association of Academic Phoneticians*. Glasgow.

CAMARGO, Z. C. (2002). *Análise da qualidade vocal de um grupo de indivíduos disfóricos: uma abordagem interpretativa e integrada de dados de natureza acústica*. Tese de doutorado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

CASTILHO, F. (2004). *Questões de pronúncia em língua inglesa: fatores desencadeantes da percepção do sotaque do falante brasileiro de língua inglesa em relação à produção dos sons fricativos alveolares*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

CAULDWELL, R. (2002). The Functional irrhythmicality of spontaneous speech: a discourse view of speech rhythm. *Apples – Journal of Applied Language Studies*, 2, p. 1-24.

CHUN, D. (2002). *Discourse Intonation in L2 – From theory and research to practice*. Philadelphia: John Benjamin's Publishing Company.

CLASSE, A. (1939). *The Rhythm of English Prose*. Oxford: Blackwell.

COUPER-KUHLEN, E. (1993). *English speech rhythm: Form and function in everyday verbal interaction*. Amsterdam: John Benjamins Publishing Company.

CRISTINO, L. (2007). *Bilinguismo e Code Switching: um estudo de caso*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

CRISTÓFARO-SILVA, T. (2001). *Fonética e fonologia do português: roteiro de estudos e guia de exercícios*. 4 ed. São Paulo: Contexto.

Cruz-Ferreira, M. (2004). Portuguese and English intonation in contrast. *Language in Contrast* 4 (2): 213-232

CRUZ-FERREIRA, M. (2010). *Age, decay, and missed opportunities*. blog *Being multilingua.blogspot.com.br*, extraído em 20 de fevereiro de 2010.

CRYSTAL, D. (1985). *A dictionary of linguistics and phonetics*. New York: Basil Blackwell.

CRYSTAL, D. (2010). The Past, Present and Future of English Rhythm. *Speakout!*, 18, p. 8-13.

DASHER, R.; BOLINGER, D. (1982). On pre-accentual lengthening, *Journal of the International Phonetic Association*. 12, p. 58-69.

DAUER, R. M. (1983). Stress timing and Syllable timing re-analyzed. *Journal of Phonetics*, 11, p. 51-62.

DAUER, R. M. (1987). Phonetic and Phonological components of language rhythm. In: *Proceedings of the 11th international congress of phonetic sciences*, v. 5, p. 477-50. Tallin: Estônia.

DAUER, R. M. (1993). *Accurate English: A complete course in pronunciation*. New Jersey: Prentice-Hall, Inc.

DELLWO, V.; VAGNER, P. (2003). Relationships between speech rate and rhythm. In: *Proceedings of the ICPHS*.

DERWING, T. M., ROSSITER, M. J.; MUNRO, M. J.; THOMSON, R. I. (2004). L2 fluency: Judgments on different tasks. *Language Learning*, 54, p. 655-79.

FABER, D. (1986). Teaching the rhythms of English: A new theoretical base. *IRAL International Review of Applied Linguistics* 24 (3), p. 205-16.

FEAR, B. D.; CUTLER, A.; BUTTERFIELD, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America*, 97, p. 1893-904.

FLEGE, J. (1986). Effects of equivalence classification on the production of foreign language speech sounds. In: JAMES, A.; LEATHER, J. (eds.) *Sound Patterns in Second Language Acquisition*. Dordrecht: Holand.

FLEGE, J. (1995). Second-language Speech Learning: Findings and Problems. In: STRANGE, W. *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, Timonium, MD: York press.

FLEGE, J.; LIU, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, p. 527-52.

FLEMMING, E.; JOHNSON, E. (2007). Rosa's roses: reduced vowels in American English. *Journal of International phonetics*, 37 (1), p. 83-96.

FLETCHER, J.; HARRINGTON, J. (1999). Lip and jaw coarticulation. In: HARDCASTLE, W. J. *Coarticulation – Theory data and techniques*, Cambridge University Press.

FOKES, J., BOND, Z.; STEINBERG, M. (1884). Patterns of English word stress by native and non-native speakers. In: VAN DEN BROECKE, M. P. R.; COHEN, A. (eds.). *Proceedings of the Tenth International Congress of Phonetic Sciences*, p. 682-6. Dordrecht: Foris.

FONAGY, P. (1993). As Funções Modais da Entoação. In: *Caderno de Estudos Linguísticos*. Campinas, 25, p. 25-65.

- FRAISSE, P. (1974). *Psychologie du rythme*. Paris: Presses universitaires de France.
- FRY, D. B. (1958). *Experiments in the perception of stress Language and Speech*, University College, London. p. 126-58.
- GARCIA, O. e OTHEGUY, R. (1989) *English across cultures: cultures across English: a reader in cross-cultural communication*. Berlin. Walter de Gruyter & Co.
- GILBERT, J. B. (2001). *Clear speech: Pronunciation and listening comprehension in American English*. Cambridge: Cambridge University Press.
- GRANT, L. (2004). *Well Said*. Boston: Heinle & Heinle.
- GUT, U. (2003). Non native speech rhythm. In: *German 15th ICPHS*. Barcelona, p. 2437-40.
- HIRST, D.; DI CRISTO, A. (1998). *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press.
- HOUSEHOLDER, F. W. (1971). *Linguistic Speculations*. New York: Cambridge University Press.
- KELLY, M. H.; BOCK, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human perception and Performance* 14, 3, p. 389-403.
- KENT, R. (1992). *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group.
- KIM, H.; COLE, J. (2007). *The stress foot as a unit of planned timing: evidence from shortening in the prosodic phrase*. Department of Linguistics University of Illinois at Urbana-Champaign, U.S.A.
- KOHLER, K. (2006). Paradigms in Experimental Prosodic Analysis: From Measurement to Function. In: *Methods in Empirical Prosody Research*. Berlin/New York: Walter de Gruyter (publisher).
- KOO H. S. (1997). A study using acoustic measurements and perceptual judgement to identify prosodic characteristics of English as spoken by Koreans. *Speech Sciences*, 2, 95-108.
- LADD, D. R. (1994). *Intonational phonology*. Cambridge: Cambridge University Press.

LADD, D. R. (1996). Naturalness and Spontaneous Speech. In: SAGISAKA, Y., CAMPBELL, N.; HIGUSHI, N. (eds.). *Computing Prosody. Computational Models for Processing Spontaneous Speech*. New York: Springer-Verlag.

LADD, D. R.; SILVERMAN, K. E. A. (1984). Vowel intrinsic pitch in connected speech. *Phonetica*, 41, 31-40.

LADEFOGED, P. (2003). *Phonetic data analysis: an introduction to instrumental phonetic fieldwork*. Oxford: Blackwells.

LAVER, J. (2002). *Principles of Phonetics*. Cambridge University Press.

LEHISTE, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts: MIT Press.

LEITE, D. R. (2009). *Estudos prosódicos sobre a manifestação do foco*. Dissertação de mestrado em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais.

LENNEBERG, E. (1971). The importance of temporal factors in behavior. In: HORTON, D. L.; JENKINS, J. (eds.). *Perception of Language*, p. 174-84. Columbus, Ohio: Charles E. Merrill publishing.

LIBERMAN, E.; PRINCE, A. (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry*, 8, p. 249-336.

LIEFF, C.; NUNES, Z. (1993). English Pronunciation and the Brazilian Learner: How to cope with language transfer. In: *Speak out!* – newsletter of the IATEFL Phonology Special Interest Group, 12.

LORANT, R. (1920). Free Prose. In: *The Century Magazine*.

LUCENTE, L. (2008): *DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos: ênfase no foco e na fala espontânea*. Dissertação de Mestrado, Instituto de Estudos da Linguagem, Unicamp. Campinas.

LUJAN, A. B. (2004). *The American Accent Guide*. Salt Lake City, Utah: Lingual Arts.

MADUREIRA, S. (2002). An acoustic study of sequences of words with adjacent primary-stressed syllables: does stress shift occur in Brazilian Portuguese?. *Cadernos de Estudos Linguísticos (UNICAMP)*, Campinas, 43, p. 109-26.

MADUREIRA, S. (2004). Expressividade da fala. In: KYRILLOS, L. R. (org.). *Expressividade: da teoria à prática*. Rio de Janeiro: Revinter, p. 15-25.

MADUREIRA, S. (2006). *Workshop de Pronúncia de Inglês como L2*. Programa de Pós Graduação em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

MAIDMENT, J. A. (1981). Abstract de Laver (1980) – The phonetic description of voice quality. Cambridge University press. *Journal of the International Phonetic Association*, 11.

MALMBERG, B. (1988). *A Fonética*. Lisboa: Edição Livros do Brasil.

MASSINI, G. (1991). *A duração no estudo do acento e do ritmo em português*. Dissertação de Mestrado. Instituto de Estudos da Linguagem, Universidade de Campinas.

MAUAD, S. A. (2007). *Questões de prosódia: uma investigação, com apoio de instrumentais de análise fonético-acústica, dos padrões entoacionais de falantes bilíngües brasileiros e norte-americanos*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

MAUAD, S. A. (2010a). An acoustic investigation into the prosodic features of an American English and Brazilian Portuguese idiom. *Revista The ESPEcialist*, 31, 2. p.145-166.

MAUAD, S. A. (2010b). *Ensino/Aprendizagem de Pronúncia de Inglês como Língua Estrangeira*. Oficina ministrada no laboratório da PUC-SP, no dia 27 de outubro de 2010.

MAUAD, S. A.; PEREZ, L.; SILVA, A.; MADUREIRA, S. (2011). Núcleo Básico: Key to English. *Coleção técnica interativa*, v. 1, Centro Paula Souza. São Paulo: Fundação Padre Anchieta.

McCONNELL-GINET (1983). Intonation in a Man's World. In: THORNE; KRAMARAE; HENLEY. *Language, Gender & Society*, p. 369-77. Chicago: Chicago Linguistic Society.

McGREGOR, L. A. (2007). *An Examination of Comprehensibility in a High Stakes Oral Proficiency Assessment for Prospective International Teaching Assistants*. Tese de Doutorado. The University of Texas at Austin.

MEIRELES, A. R. (2009). Self-organizing rhythms in Brazilian Portuguese: speech rate as a system perturbation. *VDM Verlag*, Germany.

MEIRELES, A. R.; TOZETTI, J. P.; BORGES R. R. (2010). Speech rate and rhythmic variation in Brazilian Portuguese. *Proceedings of the Speech Prosody*, Illinois.

- MOYER, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21, p. 81-108.
- MOYER, A. (2004). *Age, Accent and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Multilingual Matters Ltd, p. 17-8.
- NOOTEBOOM, S. (1997). In *The Prosody of Speech: Melody and Rhythm*. In: LAVER, J. *The Handbook of Phonetic Science*. Blackwell Publishers LTD, p. 640.
- O'DELL, M.; NIEMINEN, T. (1999). Coupled Oscillator model of speech rhythm. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 2, 1075-8.
- PATEL, A. D. (2010). *Music, Language, and the Brain*. Oxford University Press, USA.
- PELEIAS, F. (2009). *A produção e percepção da fricativa interdental surda por aprendizes brasileiros de língua inglesa*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.
- PETERSON, G. E.; BARNEY, H. L. (1952). Control methods used in a study of the vowels. *The Journal of Acoustic. Society*, 24, p. 175-84.
- PIERREHUMBERT, J. (1980). *The phonology and phonetics of English intonation*. PhD thesis, MIT.
- PIERREHUMBERT, J.; HIRSCHBERG, J. (1990). The meaning of intonation contours in interpretation of discourse. In: COHEN *et al* (eds.) *Intentions in Communication*. MIT Press.
- PIKE, K. L. (1945). *The intonation of American English*. Ann Harbour: University of Michigan Press.
- POLACZEK, M. (2003). *Compreensão oral em língua estrangeira: aspectos psicolinguísticos, fatores fonético-fonológicos, e questões de ensino aprendizagem*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.
- PRATOR, C. H. (1985). *Manual of American English pronunciation*. 4th ed. Orlando: Holt, Rinehartand Wiston, Inc.
- RIETVELD; A. C. M.; KOOPMANS-VAN BEINUM, F. J. (1987). Vowel Reduction and Stress. *Speech Communication*, 6, p. 217-29.
- ROACH, P. (1982). On the distinction between stress-timed and syllable-timed languages. In: CRYSTAL, D. (ed.) *Linguistic controversies*. London, E. Arnold.

ROACH, P. (1998). *English Phonetics and Phonology*. 2nd ed. Cambridge University Press.

ROCCA, P. (2003). *A tecnologia da fala aplicada ao ensino de entoação da língua inglesa para falantes nativos da língua portuguesa*. Tese de doutorado em Linguística aplicada e estudos da linguagem. Pontifícia Universidade Católica de São Paulo.

SCHLÜTER, J. (2005). *Rhythmic Grammar: The Influence of Rhythm on Grammatical Variation and Change in English*. (*Topics in English Linguistics*). Walter de Gruyter.

SEARLE, J. (1962). *Speech Acts*. Cambridge: Cambridge University Press.

SILVA, A. F. (2009). *Contrastes entre estratégias de falantes bilíngües na produção de um diálogo e um monólogo em inglês*. Dissertação de mestrado em Linguística Aplicada e Estudos da Linguagem. Pontifícia Universidade Católica de São Paulo.

SILVERMAN, K. (1986). F0 segmental clues depend on intonation: The case of the rise after voiced stops. *Phonetica*, 43, 76-91.

SILVERMAN, K.; BECKMAN, M.; PITRELLI, J.; OSTENDORF, M.; WIGHTMAN, C.; PRICE, P.; PIERREHUMBERT, J.; HIRSCHBERG, J. (1992). ToBI: a standard for labeling English prosody. *Proceedings of the 1992 International Conference of Spoken Language Processing*, p. 867-70.

STEELE, S. A. (1986). Interaction of vowel f0 and prosody. *Phonetica*, 43, 92-105.

TENANI, L. (2006). Considerações sobre a relação entre processos de sândi e ritmo. *Estudos da Linguagem*, 3, p. 105-22.

TENCH, P. (1992). *Studies in systemic phonology*. London: Pinter Publisher Limited.

TESCHNER, R.; WHITLEY, M. S. (2004). *Pronouncing English: a stress-based approach*. Washington D.C: Georgetown University Press.

TRUBETZKOY, N. S. (1949). *Principes de Phonologie*, tr. J. Cantineau. Paris: Klincksieck.

UDOFOT ; I. (2003). Stress and Rhythm in the Nigerian Accent of English. *Journal of Language Teaching, Linguistics and Literature*.

UNDERHILL, A. (1994). *Sound Foundations*, Oxford: Heinemann Publishers Ltd.

- VOGEL, I. BUNNEL, H.T. HOSKINS, S. (1995). The Phonetic and Phonology of the Rhythm Rule. In: CONNELL, B. e ARVANITI, A. *Phonology and Phonetic Evidence. Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press.
- WAUGH, L.; SCHUUNEVELD, C. H. (1980). *The Melody of Language: intonation and Prosody*. Baltimore: University Park Press.
- XU, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, p. 55-105.
- XU, Y. (2010). In defense of lab speech. *Journal of Phonetics*, 38, p. 329-36.
- XU, Y.; WANG, M. (2009). Organizing syllables into groups – Evidence from F0 and duration patterns in Mandarin. *Journal of Phonetics*, 37, p. 502-20.
- XU, Y.; XU, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33, p. 159-97.