



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

LEONARDO ANTONIO SILVA TEIXEIRA

ANÁLISE DO DESENVOLVIMENTO DO RITMO DE INGLÊS-L2 DE APRENDIZES
BRASILEIROS

FORTALEZA

2021

LEONARDO ANTONIO SILVA TEIXEIRA

ANÁLISE DO DESENVOLVIMENTO DO RITMO DE INGLÊS-L2 DE APRENDIZES
BRASILEIROS

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Linguística.
Área de concentração: Descrição e Análise Linguística.

Orientador: Prof. Dr. Ronaldo Manguiera Lima Jr.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

T267a Teixeira, Leonardo Antonio Silva.

Análise do desenvolvimento do ritmo de inglês-L2 de aprendizes brasileiros / Leonardo Antonio Silva Teixeira. – 2021.

137 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística, Fortaleza, 2021.

Orientação: Prof. Dr. Ronaldo Manguiera Lima Júnior.

1. aquisição fonológica. 2. prosódia. 3. métricas rítmicas. 4. parâmetros acústicos. 5. instrução explícita. I. Título.

CDD 410

LEONARDO ANTONIO SILVA TEIXEIRA

ANÁLISE DO DESENVOLVIMENTO DO RITMO DE INGLÊS-L2 DE APRENDIZES
BRASILEIROS

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Linguística.
Área de concentração: Descrição e Análise Linguística.

Aprovada em: 27/09/2021

BANCA EXAMINADORA

Prof. Dr. Ronaldo Manguiera Lima Júnior. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dra. Pâmela Freitas Pereira Toassi
Universidade Federal do Ceará (UFC)

Prof. Dr. Leônidas José da Silva Júnior
Universidade Estadual da Paraíba (UEPB)

Aos meus pais, Graça e Antonio.

AGRADECIMENTOS

A todos e todas que constroem o Programa de Pós-Graduação em Linguística da Universidade Federal do Ceará, por contribuírem para o avanço da ciência brasileira, sobretudo em tempos em que o estudo e a pesquisa são formas de resistência ao negacionismo e à barbárie.

À professora Ms. Maria Manolisa Nogueira Vasconcellos, com quem tive o privilégio de trabalhar e aprender como aluno e bolsista durante a graduação, e cujos ensinamentos sobre a linguagem e a vida foram fundamentais no percurso que trilhei até aqui.

À professora Dra. Vlândia Maria Cabral Borges, pelas valiosas sugestões enquanto parecerista dessa dissertação em andamento e por ter sempre me estimulado a “alçar voos mais altos”.

À Professora Dra. Elisângela Nogueira Teixeira, pelos valiosos ensinamentos não só de cunho teórico, mas também sobre o próprio fazer acadêmico-científico.

Ao professor Dr. Willem Leo Marie Wetzels, pelas pertinentes observações em minha qualificação.

Ao professor Dr. Leônidas José da Silva Júnior, pela solicitude com que aceitou compor minhas bancas de qualificação e defesa, e pela disponibilidade em colaborar com o desenvolvimento deste trabalho.

À Professora Dra. Pâmela Freitas Pereira Toassi, por ter gentilmente aceitado compor esta banca de defesa e pelo tempo concedido à leitura deste trabalho.

Ao professor Dr. Ronaldo Manguiera Lima Júnior, pela excelente e cuidadosa orientação, por todas as oportunidades de aprendizagem que me proporcionou, e pelo estímulo a ir sempre além.

À minha família, aos meus amigos e às minhas amigas, que souberem entender meus momentos de ausência.

Ao Valdiney, pela compreensão e companheirismo, mesmo nas horas de incerteza.

Precious, precious silver and gold and pearls in oyster's flesh
Drop down we two to serve and pray to love
Born again from the rhythm, screaming down from heaven
Ageless, ageless, I'm there in your arms

Jeff Buckley

RESUMO

O objetivo deste estudo é descrever e discutir o desenvolvimento do ritmo do inglês-L2 por aprendizes brasileiros por meio de métricas rítmicas e de parâmetros prosódico-acústicos que caracterizam a produção oral desses aprendizes em diferentes etapas do desenvolvimento da L2. Este estudo adota uma perspectiva dinâmica de desenvolvimento de L2 (DE BOT; LOWIE; VESPOOR, 2007; DE BOT, 2008; LARSEN-FREEMAN, 1997; LARSEN-FREEMAN; CAMERON, 2008; LOWIE; VERSPOOR, 2015) e uma abordagem multidimensional e gradiente para o ritmo da fala (CUMMING, 2010; FUCHS, 2016; LOUKINA et. al., 2011; SILVA JR.; BARBOSA, 2019b). Foram analisados quatro corpora de produção oral. Cinco aprendizes brasileiros de inglês-L2 foram gravados lendo um texto em inglês no início do curso de Letras-Inglês, e mais uma vez quatro semestres depois, após terem cursado duas disciplinas de fonologia da língua inglesa. Eles também foram gravados lendo uma versão desse texto traduzida para o português. Além dos aprendizes, cinco falantes nativos de inglês norte-americano foram gravados lendo o mesmo texto em inglês. Sendo assim, o estudo também busca avaliar possíveis efeitos da instrução explícita de pronúncia sobre o desenvolvimento do ritmo do inglês-L2 dos aprendizes. Os dados foram segmentados manualmente para a extração das métricas rítmicas no PRAAT (BOERSMA; WEENINK, 2021) em unidades vocálicas (V), consonantais (C), vogal-vogal (VV), sentenças (S) e unidades prosódicas superiores – *chunks* (CH), e foi utilizado o script Metrics&AcousticsExtractor (SILVA JR.; BARBOSA, 2019) para a extração de 30 métricas rítmicas e 12 parâmetros acústicos. Os dados foram tratados estatisticamente via R (R CORE TEAM, 2021) por meio da implementação de modelos de regressão de efeitos mistos, adotando língua e semestre como variáveis preditoras, e métricas rítmicas e parâmetros acústicos como variáveis de resposta. Os resultados posicionaram o português-L1 e o inglês-L1 em espaços rítmicos distintos, conforme previsto pela literatura; na dimensão duracional, as métricas posicionaram o inglês-L2 da primeira gravação distante tanto do inglês-L1 como do português-L1; na dimensão de f_0 , os parâmetros acústicos posicionaram o inglês-L2 da primeira gravação mais próximo ao português-L1. Em ambas as dimensões, o inglês-L2 da gravação posterior ficou mais próximo ao inglês-L1, sugerindo um percurso de desenvolvimento em direção à língua alvo. Os resultados também indicam efeitos positivos do ensino explícito de pronúncia.

PALAVRAS-CHAVE: aquisição fonológica; prosódia; métricas rítmicas; parâmetros acústicos; instrução explícita.

ABSTRACT

The aim of this study is to describe and discuss the development of L2 English rhythm by Brazilian learners through rhythmic metrics and prosodic-acoustic parameters that characterize the oral production of these learners at different stages of L2 development. This study adopts a dynamic perspective of L2 development (DE BOT; LOWIE; VESPOOR, 2007; DE BOT, 2008; LARSEN-FREEMAN, 1997; LARSEN-FREEMAN; CAMERON, 2008; LOWIE; VERSPOOR, 2015) and a multidimensional and gradient approach to speech rhythm (CUMMING, 2010; FUCHS, 2016; LOUKINA et. al., 2011; SILVA JR.; BARBOSA, 2019b). Four corpora of oral production were analyzed. Five Brazilian learners of English-L2 were recorded reading a text in English at the beginning of their college studies in English Language Teaching, and again four semesters later, after having taken two English phonology courses. They were also recorded reading a version of the text translated into Portuguese. Besides the learners, five native speakers of North American English were recorded reading the same text in English. Thus, this study also seeks to evaluate possible effects of explicit pronunciation instruction on the development of the rhythm of English by the learners. Data were manually segmented into vowel units (V), consonant (C), vowel-vowel (VV), sentences (S) and higher prosodic units - chunks (CH) for the extraction of rhythmic metrics in PRAAT (BOERSMA; WEENINK, 2021), and the script Metrics&AcousticsExtractor (SILVA JR.; BARBOSA, 2019) was used to extract 30 rhythmic metrics and 12 acoustic parameters. Data were then statistically treated via R (R CORE TEAM, 2021) through the implementation of mixed-effects regression models, adopting language and semester as predictor variables, and rhythmic metrics and acoustic parameters as response variables. Results placed Brazilian Portuguese and English-L1 in different rhythmic spaces, as predicted by the literature; in the durational dimension, the metrics positioned the English-L2 of the first recording far from both English-L1 and Brazilian Portuguese; in the f_0 dimension, however, the acoustic parameters placed the English-L2 of the first recording closer to Brazilian Portuguese. In both dimensions, the English-L2 of the subsequent recording was closer to English-L1, suggesting a developmental route towards the target language. The results also suggest positive effects of the explicit teaching of pronunciation.

KEY-WORDS: phonological acquisition; prosody; rhythmic metrics; acoustic parameters; explicit instruction.

LISTA DE FIGURAS

Figura 1 – Distribuição do inglês (EN), holandês (DU), polonês (PO), espanhol (SP), italiano (IT), francês (FR), catalão (CA) e japonês (JA) em um plano cartesiano com ΔC e %V nos eixos, em que as barras representam o erro- padrão	29
Figura 2 – Duração de vogais sucessivas em duas línguas hipotéticas A e B.....	31
Figura 3 – Forma de onda parcial, espectrograma de banda larga e quatro camadas respectivamente, segmentadas e rotuladas como: 1) sílabas fonéticas (VV); 2) vogais e/ou consoantes (V/C); 3) sentenças (S); 4) <i>chunks</i> (CH) produzidas por um falante não nativo.....	52
Figura 4 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔC e %V.....	57
Figura 5 – <i>Boxplot</i> de %C para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB.....	58
Figura 6 – Dados do presente estudo e dos de Ramus, Nespore e Mehler (1999) para ΔC e %V, em que Eng = inglês, Pol = polonês, Dut = holandês, Spa = espanhol, Fre = francês, Ita = italiano, Cat = catalão, Jap = japonês	59
Figura 7 – Dados do presente estudo em meio a todos os dados resenhados e obtidos por Arvaniti (2012) para ΔC e %V, em que Eng = inglês, Ger = alemão, Gre = grego, Spa = espanhol, Ita = italiano, Kor = coreano	60
Figura 8 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔC e ΔV	61
Figura 9 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔV e %V.....	62
Figura 10– <i>Boxplots</i> de ΔVC e ΔS para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB.....	63
Figura 11 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) PB no plano formado pelos eixos VarcoC e VarcoV	65
Figura 12– Dados do presente estudo em meio a todos os dados resenhados e obtidos por Arvaniti (2012) para VarcoC e VarcoV, em que Eng = inglês, Ger = alemão, Gre – grego, Spa = espanhol, Ita = italiano, Kor = coreano.....	67
Figura 13– <i>Boxplots</i> das médias de VarcoVC e VarcoS para Inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	68
Figura 14– Inglês-L1, inglês-L2(1), inglês-L2(4) e PB no plano formado pelos eixos	

rPVI-C e nPVI-V	70
Figura 15– Dados do presente estudo e dos de Grabe e Low (2002) para rPVI-C e nPVI-V.....	72
Figura 16– Dados do presente estudo em meio a todos os dados resenhados e obtidos por Arvaniti (2012) para rPVI-C e nPVI-V, em que Eng = inglês, Ger = alemão, – Gre – grego, Spa = espanhol, Ita = italiano, Kor = coreano.....	73
Figura 17– <i>Boxplots</i> das médias de rPVI-V e nPVI-C para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	74
Figura 18– <i>Boxplots</i> das médias de rPVI-VC e nPVI-S para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	75
Figura 19– <i>Boxplots</i> das médias de rPVI-S e nPVI-S para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	75
Figura 20– <i>Boxplots</i> das médias de RR-V e RR-C para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	78
Figura 21– <i>Boxplots</i> das médias de RR-VC e RR-S para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	79
Figura 22– <i>Boxplots</i> das médias de VI-V e VI-C para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	81
Figura 23– <i>Boxplots</i> das médias de VI-VC e VI-S para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	81
Figura 24– <i>Boxplots</i> das médias de YARD-V e YARD-C para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	84
Figura 25– <i>Boxplots</i> das médias de YARD-VC e YARD-S para inglês-L1, inglês-L2(1), inglês-L2(4) e PB.....	84
Figura 26– Distâncias euclidianas no eixo ΔC (eixo y) para inglês-L1-L2(1), inglês-L1-L2(4).....	86
Figura 27– <i>Boxplots</i> da mediana de f_0 e média da primeira derivada de f_0 para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB.....	89
Figura 28– Medianas de f_0 para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por <i>chunk</i>	90
Figura 29– <i>Boxplots</i> para médias de pico de f_0 e f_0 mínimo para Inglês-L1, inglês-L2(1), inglês-L2 (4) e PB.....	92
Figura 30– <i>Boxplot</i> de pico de f_0 para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por <i>chunk</i>	93

Figura 31 – <i>Boxplot de f_0 mínimo para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por chunk</i>	94
Figura 32 – <i>Boxplots para médias de desvio-padrão de f_0 e assimetria de f_0 para inglês-L1 inglês-L2 (1), inglês-L2 (4) e PB</i>	98
Figura 33 – <i>Boxplots para $\sigma\Delta 1-f_0$ e $sk\Delta 1-f_0$ para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB</i>	99
Figura 34 – <i>Boxplot da taxa de f_0 para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB</i>	102
Figura 35 – <i>Boxplot da 24ª fase espectral (emph) para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB</i>	103
Figura 36 – <i>Ênfase espectral para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por chunk</i>	104
Figura 37 – <i>Duração média dos picos silábicos (μ_{dur}- Sil) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB</i>	107
Figura 38 – <i>Taxa de elocução (SR) para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB</i>	108

LISTA DE QUADROS

Quadro 1 – Métricas rítmicas baseadas na duração	27
Quadro 2 – Métricas rítmicas baseadas em f_0 , intensidade, altura e sonoridade	36
Quadro 3 – Métricas rítmicas e parâmetros acústicos utilizados neste estudo	53

LISTA DE TABELAS

Tabela 1 – Médias absolutas de %V, %C, ΔV , ΔC , ΔVC e ΔS e desvio-padrão para PB, inglês-L1, inglês-L2 (1) e inglês-L2.....	56
Tabela 2 – Coeficientes, intervalos de confiança (95%) e valores p para os dois modelos de regressão linear com efeitos mistos ajustados para AC e %V	59
Tabela 3 – Coeficientes, intervalos de confiança (95%) e valores de p um modelo de regressão linear com efeitos mistos ajustados para ΔV	63
Tabela 4 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para ΔVC e ΔS .	64
Tabela 5 – Médias absolutas de VarcoV, VarcoC, VarcoVC e VarcoS e desvio-padrão para PB, inglês-L1, inglês-L2 (1) e inglês-L2.....	65
Tabela 6 – Coeficientes, intervalos de confiança (95%) e valores p para os dois modelos de regressão linear com efeitos mistos ajustados para VarcoC e VarcoV	66
Tabela 7 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VarcoVC e VarcoS.....	69
Tabela 8 – Médias absolutas de PVIs brutos e normalizados e desvio-padrão para inglês-L1, inglês-L2 (1) e inglês-L2.....	70
Tabela 9 – Coeficientes, intervalos de confiança (95%) e valores p para os dois modelos de regressão linear com efeitos mistos ajustados para rPVI-C e nPVI-V	71
Tabela 10 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para nPVI-C e rPVI-V.....	74
Tabela 11 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para rPVI-VC e nPVI-VC.....	76
Tabela 12 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois	77

modelos de regressão linear com efeitos mistos ajustados para rPVI-S e nPVI-S.....	
Tabela 13 – Médias absolutas e desvio-padrão de RR-V, RR-C, RR-VC e RR-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2.....	77
Tabela 14 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para RR-V e RR-C.....	79
Tabela 15 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para RR-VC e RR-S.....	80
Tabela 16 – Médias absolutas e desvio-padrão (entre parênteses) de, VI-C, VI-VC e VI-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2.....	80
Tabela 17 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VI-V e VI-C.....	82
Tabela 18 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VI-VC e VI-S.....	82
Tabela 19 – Médias absolutas de YARD-V, YARD-C, YARD-VC e YARD-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2.....	83
Tabela 20 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para YARD-V e YARD-C.....	85
Tabela 21 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para YARD-VC e YARD-S.....	85
Tabela 22 – Distâncias euclidianas entre inglês-L1 e interlínguas, e entre as L1s.....	87
Tabela 23 – Médias absolutas dos parâmetros acústicos para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4)	88

Tabela 24 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para f_{0med} e $\mu\Delta 1-f_0$	91
Tabela 25 – Médias de pico de f_0 e f_0 mínimo, e escopo de f_0 em semitons para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.....	95
Tabela 26 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para f_{0peak} e f_{0min}	97
Tabela 27 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para σf_0 e $\sigma\Delta 1-f_0$..	100
Tabela 28 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para f_{0sk} e $sk\Delta 1-f_0$	100
Tabela 29 – Médias de desvio-padrão de f_0 (σf_0) e primeira derivada do desvio-padrão de f_0 ($\sigma\Delta 1-f_0$) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.....	101
Tabela 30 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para f_0-R	103
Tabela 31 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para ênfase espectral.....	105
Tabela 32 – Médias ênfase espectral (emph) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.....	106
Tabela 33 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para ênfase espectral.....	107
Tabela 34 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para taxa de elocução.....	109
Tabela 35 – Médias da taxa de elocução (SR) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.....	109

SUMÁRIO

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	O conceito de ritmo	20
2.2	O ritmo como isocronia	21
2.3	O ritmo como gradiência	23
2.4	Métricas rítmicas de base duracional	27
2.5	Métricas rítmicas com base em f_0 e intensidade.....	36
2.6	O ritmo do português brasileiro	40
2.7	O ritmo do inglês	44
2.8	O ritmo em pesquisas sobre o desenvolvimento de L2	46
3	MÉTODO	50
3.1	Contexto	50
3.2	Participantes	50
3.3	Materiais	51
3.4	Procedimentos	51
4	RESULTADOS	55
4.1	Métricas rítmicas	55
4.1.1	<i>Métricas de proporção (%) e desvio-padrão (Δ).....</i>	56
4.1.2	<i>Varcos</i>	64
4.1.3	<i>PVIs</i>	69
4.1.4	<i>RRs.....</i>	77
4.1.5	<i>VIIs</i>	80
4.1.6	<i>YARDS</i>	83
4.1.7	<i>Distâncias euclidianas</i>	85
4.2	Parâmetros acústicos	87
4.2.1	<i>Mediana de f_0 e média da primeira derivada de f_0 ($\mu\Delta 1- f_0$).....</i>	89
4.2.2	<i>Pico de f_0 e f_0 mínimo.....</i>	91
4.2.3	<i>Desvio-padrão de f_0 (σf_0), assimetria (skewness) de f_0 e respectivas derivadas ($\sigma\Delta 1- f_0$ e $sk\Delta 1- f_0$)</i>	97
4.2.4	<i>Taxa de f_0 (f_0-R).....</i>	102
4.2.5	<i>Ênfase espectral.....</i>	103
4.2.6	<i>Duração média dos picos silábicos ($\mu dur- Sil$).....</i>	106

4.2.7	<i>Taxa de elocução (SR)</i>	108
5	DISCUSSÃO	111
6	CONCLUSÃO	120
	REFERÊNCIAS	122
	ANEXO A – TEXTO EM INGLÊS	131
	ANEXO B – TEXTO EM PORTUGUÊS	132

1 INTRODUÇÃO

No âmbito dos estudos em desenvolvimento de línguas não nativas (L2¹), parece haver maior ênfase em pesquisas que envolvem aspectos segmentais em comparação a aspectos prosódicos (THOMSON; DERWING, 2015; LI; POST, 2014). A menor ênfase em aspectos prosódicos também se reflete em modelos de desenvolvimento de L2, tais como o Modelo de Aquisição da Fala - *Speech Learning Model* (FLEGE, 1995) e sua versão revisada - *Revised Speech Learning Model* (FLEGE; BOHN, 2021), bem como o Modelo de Assimilação Perceptual - *Perceptual Assimilation Model* (BEST, 1995) e sua versão para L2 - *Perceptual Assimilation Model of Second Language Speech Learning* (BEST; TYLER, 2007), que, ao enfatizarem aspectos segmentais, oferecem poucos subsídios à compreensão do desenvolvimento da prosódia por aprendizes de L2. Esse universo de pesquisa pouco explorado contrasta com evidências que demonstram a persistência de padrões prosódicos atípicos da L2 mesmo em usuários de nível avançado (LI; POST, 2014).

Moreno (2000) assevera que a manifestação desses padrões prosódicos atípicos pelo falante não nativo afeta mais a compreensibilidade (i.e. o nível de esforço necessário para a compreensão do enunciado falado) que realizações fonéticas atípicas no nível segmental. Celce-Murcia, Brinton e Goodwin (2010) afirmam ainda que os mal-entendidos que podem emergir a partir da mobilização de padrões prosódicos inadequados podem causar prejuízos ainda mais sérios à comunicação que os advindos do nível segmental:

Ainda segundo Moreno (2000), um bom domínio da prosódia em L2 facilita a compreensão por parte dos interlocutores mesmo quando os aprendizes cometem erros no nível lexical, estilístico e gramatical. Mas o que é prosódia? De acordo com Barbosa (2012), a definição negativa de prosódia cunhada inicialmente por estruturalistas e funcionalistas como fatos de fala não pertinentes ao quadro fonemático cedeu lugar a um termo guarda-chuva que abrange questões linguísticas como acento, entonação e ritmo, paralinguísticas como marcadores discursivos e sociais, e fatores extralinguísticos como as emoções instanciadas nos atos enunciativos. Esses fatores, combinados a variáveis sociais e biológicas “moldam nossa enunciação imprimindo a” “o que se fala” um “modo de falar” que é dirigido intencionalmente ou não ao ouvinte.” (BARBOSA, 2012, p. 14).

Dentre os aspectos prosódicos passíveis de investigação, o ritmo é ainda o menos explorado (WHITWORTH, 2002; CUMMING, 2010; GUT, 2012). Há evidências de que o

¹ Neste trabalho não faremos distinção entre segunda língua e língua estrangeira consoante Lima Jr., Alves e Kupske (2021).

ritmo pode influenciar o processo de comunicação de maneira global, afetando níveis de grau de acento, inteligibilidade (SILVA JR.; BARBOSA, 2019b) e compreensibilidade (MUNRO; DERWING, 2001; ORDIN; POLYANSKAYA, 2015). Outrossim, o ritmo pode oferecer aos aprendizes pistas acústicas que os guiam na segmentação da cadeia da fala em constituintes prosódicos que realizam múltiplas funções nos campos linguístico, paralinguístico e expressivo (BARBOSA, 2012). Dessa forma, o ritmo figura como importante construto a ser explorado em pesquisas sobre desenvolvimento de L2.

A escolha desse objeto de estudos também é orientada fundamentalmente pelo entendimento de que o ensino-aprendizagem de L2 pode beneficiar-se do fazer empírico, que tem o potencial de informar tanto acerca dos processos básicos subjacentes à cognição humana quanto das intervenções pedagógicas que se mostram necessárias diante de um determinado conjunto de dados. Ainda de acordo com Levis *et al.* (2016), o ensino-aprendizagem do componente fonético-fonológico no contexto de desenvolvimento de L2 necessita de melhor conexão entre práticas pedagógicas orientadas por pesquisas e pesquisas orientadas às práticas pedagógicas. Sabe-se ainda que, para um comando satisfatório em L2, o domínio da pronúncia é imprescindível (LIMA JR., 2010; ALVES; MAGRO, 2011; DERWING, 2018). Autores como Celce-Murcia, Brinton e Goodwin (2010) asseveram a necessidade de falantes não nativos possuírem um determinado nível de pronúncia para não terem sua habilidade de comunicação oral comprometida, sendo a pronúncia componente essencial da competência comunicativa (MORLEY, 1994). Ora, todos os domínios que compõem a competência comunicativa são mediados pela oralidade, sendo o domínio fonético- fonológico “a base de todas as unidades linguísticas superiores” (PENNINGTON, 1996, apud LIMA JR. 2010). Assim, “a falta de conhecimento fonético-fonológico da L2 aprendida pode, portanto, prejudicar as tentativas de comunicação, deflagrando, assim, a necessidade de haver foco no ensino da pronúncia na sala de L2.” (LIMA JR., 2010, p. 751).

A realidade física e psicológica do ritmo tem sido objeto de inúmeras pesquisas ao longo dos últimos 50 anos. Evidências da realidade psicológica do ritmo foram encontradas em experimentos que, ao manipular o sinal da fala por meio de filtros passa-baixa e ressíntese para neutralizar a informação segmental, resultaram na percepção de semelhança entre línguas consideradas de mesma categoria rítmica e diferença entre línguas consideradas de categorias rítmicas distintas (NAZZI *et al.* 1998; RAMUS; MEHLER, 1999; RAMUS; NESPOR; MEHLER, 2003). Dados empíricos também sugerem que bebês são sensíveis às classes rítmicas, isto é, são capazes de perceber diferenças interlinguísticas em função do ritmo (MEHLER *et al.* 1987; 1988). O trabalho de psicolinguistas também sugere que a aquisição

fonológica dos bebês é baseada nas diferenças rítmicas, pois parece haver correlação entre tipologia rítmica e segmentação da cadeia da fala (RAMUS; NESPOR; MEHLER, 1999), o que sugere que o ritmo pode desempenhar um papel importante nas primeiras fases de aquisição de primeira língua (doravante L1), auxiliando na detecção dos limites entre palavras (MORGAN, 1996), e, posteriormente, no processo de aprendizagem de leitura (HOLLIMAN *et al.* 2010). Há ainda indícios de que o ritmo permanece operando enquanto pista acústica na detecção de palavras mesmo na fase adulta (CUTLER; BUTTERFIELD 1992; GILBERT *et al.* 2011). Apesar dos avanços, problemas de natureza empírica e teórica perduram e a realidade do ritmo ainda permanece por ser compreendida inteiramente em suas múltiplas dimensões.

Em termos perceptuais, o ritmo caracteriza-se pela maneira como as línguas naturais “fluem”, ou, ainda, segundo Platão (apud NESPOR, SHUKLA; MEHLER, 2011), é a ordem no movimento. A percepção de diferentes categorias rítmicas associadas às diferentes línguas e grupos de línguas tem motivado analogias que buscam captar o efeito impressionístico provocado por seus diferentes modos de “fluir”. Assim, o ritmo de línguas como o inglês é comparado ao sinal de código Morse e o ritmo de línguas como o espanhol é comparado ao som de metralhadora (LLOYD JAMES, 1940 apud ABERCROMBIE 1967). A analogia busca captar a impressão causada pelo tipo de recorrência periódica dos elementos que marcariam o ritmo dessas línguas. Tanto o código Morse como o inglês teriam uma ênfase/alongamento em alguns segmentos intervalados por outros mais curtos, como o código Morse. Já o espanhol teria uma sequência de segmentos com pulsação mais regular, como o som de metralhadora. Essa perspectiva mostrou-se empiricamente equivocada por meio de testes de percepção conduzidos por Abercrombie (1967).

Ainda na década de 40, Pike (1945) propôs que as diferenças rítmicas entre as línguas poderiam ser explicadas a partir do tipo de unidade recorrente no sinal acústico das línguas naturais: em línguas como o inglês, essas unidades seriam os grupos interacentuais (entre o início de uma sílaba tônica e o início da próxima), e em línguas como o espanhol, essas unidades seriam as sílabas. Mais tarde, Abercrombie (1967) propôs a Hipótese da Isocronia a fim de formalizar essa diferença perceptual, buscando fazer uma classificação tipológica entre as línguas do mundo. De acordo com essa proposta, as unidades rítmicas seriam unidades isócronas, ou seja, de mesma duração. Assim, a diferença rítmica entre as línguas poderia ser explicada por qual unidade detém regularidade duracional. Línguas como o inglês seriam classificadas como línguas de ritmo acentual, por terem intervalos isócronos; e línguas como o espanhol teriam um ritmo silábico, pois suas unidades isócronas seriam as sílabas. Posteriormente, Ladefoged (1975) acrescentou a tipologia de ritmo moraico, para dar conta de

línguas como o japonês, que teria a mora como unidade isócrona (LADEFOGED 1975; LADEFOGED; JOHNSON, 2014).

Sendo assim, a tipologia rítmica, cuja nomenclatura perdura até hoje, foi concebida categoricamente, ou seja, cada língua natural deveria figurar em uma ou outra tipologia rítmica. Entretanto, o paradigma isócrono mostrou-se empiricamente insustentável: quando se considera, por exemplo, que o desvio-padrão de medidas de duração de sílabas e pés silábicos em línguas consideradas de ritmo silábico e acentual, que mede a variabilidade das durações das unidades supostamente isócronas, não é capaz de satisfatoriamente separar tais línguas em categorias distintas (DAUER, 1983). Diversos estudos conduzidos nos anos 70, 80 e 90 demonstraram a implausibilidade da hipótese isócrona, uma vez que intervalos de mesma duração não são encontrados no sinal acústico (CUMMING, 2010), o que suscitou propostas alternativas para explicar a variabilidade do ritmo nas línguas do mundo. Entre elas, surgiu o Modelo Unidimensional de Ritmo (DAUER, 1987), em que o ritmo é concebido como um efeito global em função da estrutura fonológica das línguas naturais, e as categorias *mais acentual* e *menos acentual* ocupam, cada uma, um extremo de um *continuum* rítmico.

A partir de meados dos anos 90, uma segunda onda de estudos em ritmo linguístico emerge seguindo o paradigma de variabilidade e gradiência, e não mais um categórico e isócrono. Nesse paradigma, a classificação rítmica é buscada em função das características duracionais dos intervalos de referência (vocálicos, consonantais, silábicos etc.), que podem ser computados por meio de índices estatísticos denominados métricas rítmicas.

Há ainda abordagens que se utilizam de outros correlatos de proeminência, tais quais frequência fundamental (f_0) e intensidade, já que a duração não é o único correlato de proeminência. Frequência fundamental e intensidade, ou seus respectivos correlatos *pitch* e altura, também contribuem com a percepção da proeminência e as línguas diferem quanto ao uso desses mecanismos para expressar acento lexical, acento frasal e foco (FUCHS, 2016).

Não há consenso, porém, em relação às métricas e parâmetros acústicos que sustentam a percepção de diferentes categorias rítmicas. Conforme Gut (2012), diferentes estudos às vezes geram resultados conflitantes em métricas distintas. Há pouco progresso também em relação à correlação entre os fatores que parecem comprovar a realidade física do ritmo e sua dimensão psicológica. Além disso, as medidas de ritmo parecem variar consoante aos estímulos, às idiossincrasias do falante, aos examinadores, aos métodos de segmentação da cadeia da fala e aos textos. Tamanha complexidade, inconsistência de resultados e dificuldade de obtenção de dados robustos têm suscitado inclusive questionamentos sobre a natureza do ritmo enquanto construto teórico válido (ARVANITI, 2012).

Este trabalho busca contribuir com as pesquisas no âmbito do ritmo por meio da aplicação de métricas de base duracional e de parâmetros acústicos (cf. Quadro 3) aos dados de quatro corpora de produção oral: português brasileiro (doravante PB) -L1, inglês-L1 e inglês-L2 em duas etapas distintas de desenvolvimento, tendo em vista que poucos trabalhos no âmbito do ritmo investigam aprendizes de L2, e em menor número ainda são os trabalhos que investigam falantes de português-L1 aprendizes de inglês-L2. Constata-se ainda uma lacuna para estudos que consideram o desenvolvimento do ritmo em L2 em função do ensino explícito de pronúncia, e em uma perspectiva longitudinal, sendo esta a primeira investigação do desenvolvimento do ritmo por falantes de PB da variedade dialetal de Fortaleza aprendizes de inglês-L2 correlacionada ao ensino explícito de pronúncia em diferentes etapas de desenvolvimento. Tais considerações demonstram o caráter oportuno deste estudo, em um campo ainda pouco explorado do componente fonético-fonológico no contexto de desenvolvimento de L2. Outrossim, argumentamos consoante Lima Jr. (2016), que o processo de desenvolvimento fonológico de uma L2, por ser intrinsecamente dinâmico e não-linear, deve ser estudado por meio da análise de dados individuais e longitudinais de aprendizes.

O objetivo geral deste estudo é descrever o desenvolvimento do ritmo do inglês-L2 por aprendizes brasileiros por meio de métricas rítmicas e de parâmetros acústicos que caracterizam a produção oral desses aprendizes em diferentes etapas do desenvolvimento da L2, em uma perspectiva multidimensional. Este estudo também elenca como objetivos específicos: (i) posicionar o ritmo do PB, do inglês norte-americano, e do inglês-L2 de brasileiros no espaço rítmico; (ii) avaliar as métricas rítmicas e os parâmetros acústicos comumente utilizados na caracterização rítmica das línguas quando aplicados à interlíngua; (iii) avaliar o efeito do ensino explícito de pronúncia no desenvolvimento do ritmo do inglês-L2 por aprendizes brasileiros; e (iv) avaliar a influência do ritmo do PB-L1 no desenvolvimento do ritmo do inglês-L2 de aprendizes brasileiros. Para tanto, analisamos dados de produção oral de alunos brasileiros de graduação em Letras-Inglês lendo um texto em português e lendo um texto em inglês em dois momentos distintos, antes e após cursarem duas disciplinas sobre a fonologia do inglês. Também foram analisados dados de falantes nativos de inglês norte-americano lendo o mesmo texto.

Buscamos, por meio da presente análise, responder às seguintes perguntas: (i) Como as métricas e os parâmetros acústicos empregados posicionam o inglês-L1 norte-americano, o inglês-L2 de brasileiros, e o PB-L1 em seus espaços rítmicos? (ii) Qual é a influência do ritmo do PB-L1 no desenvolvimento do inglês-L2 dos aprendizes? (iii) Qual é o efeito do ensino explícito de pronúncia no desenvolvimento do ritmo do inglês-L2 dos

aprendizes? A hipótese de trabalho básica é de que o PB-L1, o inglês-L1 e o inglês-L2 são ritmicamente diferentes, o que será capturado pelas métricas rítmicas e pelos parâmetros acústicos empregados. Além disso, haverá diferenças rítmicas entre o inglês-L2 dos falantes nas duas diferentes etapas do desenvolvimento cujas gravações foram analisadas. Espera-se também que as métricas rítmicas da primeira gravação dos aprendizes sejam mais dissimilares ao inglês-L1 devido à ausência de instrução explícita e aos processos fonológicos do PB, que tendem a ser transferidos para o inglês-L2.

Com vistas a atingir os objetivos elencados neste projeto de pesquisa, adotaremos a perspectiva proposta por autores como Cumming (2010), Loukina et. al. (2011), Fuchs (2016) e Silva Jr. e Barbosa (2019b), que argumentam em favor de um modelo multidimensional de ritmo, com uma gama de correlatos acústicos e métricas que o informam. Assim, assumimos que a alternância de unidades proeminentes e menos proeminentes em uma língua só pode ser totalmente capturada se diferentes correlatos e dimensões de proeminência forem considerados. Essa perspectiva corrobora também a definição de ritmo adotada neste trabalho, que foi proposta por Barbosa e Bailly (1994), e que concebe o ritmo como a sensação causada pela sucessão de diferentes graus de proeminência e não proeminência silábica alternados (SILVA JR.; BARBOSA, 2019).

Por fim, espera-se que este estudo contribua com a construção de uma metodologia multidimensional de pesquisa em ritmo, que seja capaz de captar não apenas diferenças interlinguísticas mas também interindividuais no contexto do desenvolvimento de L2, com implicações também pedagógicas para o ensino-aprendizagem do componente fonético-fonológico.

A fim de cumprir com os objetivos elencados, esta dissertação organiza-se em seis capítulos. O primeiro capítulo apresenta a introdução do estudo realizado, com um panorama geral do território de pesquisa em que se insere, a justificativa para a sua realização, os objetivos a serem cumpridos, as questões que o orientam e as hipóteses levantadas. O segundo capítulo apresenta a fundamentação teórica deste estudo, abordando o conceito de ritmo, as principais ondas de pesquisa sobre o ritmo na Linguística, métricas rítmicas, o ritmo do PB, o ritmo do inglês e o ritmo no âmbito de pesquisas em desenvolvimento de L2. No terceiro capítulo, destinado à descrição dos materiais e métodos utilizados nesta pesquisa, caracterizamos o contexto da pesquisa, os participantes, os materiais e os procedimentos empregados na coleta, segmentação e análise de dados. O quarto capítulo apresenta os resultados obtidos e o quinto capítulo é destinado à discussão destes resultados em relação à literatura. No sexto capítulo, a conclusão da pesquisa é apresentada.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são apresentadas as bases teóricas que orientam esta pesquisa. A seção 2.1 introduz o conceito de ritmo e a definição de trabalho adotada neste estudo para esse construto. Nas subseções seguintes o percurso histórico do ritmo enquanto objeto de pesquisa a partir da década de 40 é apresentado de forma não exaustiva (subseções 2.2 e 2.3). Também são descritas as principais métricas rítmicas disponíveis na literatura (seções 2.4 e 2.5). Em seguida, as características rítmicas do PB e do inglês são apresentadas e discutidas (seções 2.6 e 2.7 respectivamente). Na seção 2.8 são pontuadas algumas investigações sobre o ritmo no âmbito específico do desenvolvimento de L2.

2.1 O conceito de ritmo

A etimologia da palavra *ritmo* é elucidativa para o construto teórico que será apresentado nas próximas seções: origina-se da palavra grega $\rhoυθμοϲ$, que, por sua vez, deriva do verbo $ρεί$, que significa “fluir”. Remontando ainda à Antiguidade Clássica, Platão, citado por Nespor, Shukla e Mehler (2011), define ritmo como a ordem no movimento. Há, portanto, ordem no aparente caos de coisas como as ondas, o cantar dos pássaros, os batimentos cardíacos e os ciclos circadianos, sendo o ritmo parte fundamental da vida (CUMMINGS, 2010).

Na espécie humana, a produção e a percepção do comportamento rítmico manifestam-se na música e na linguagem. Para Fitch, Hauser e Chomsky (2005), o fato de o comportamento rítmico manifestar-se em outros domínios cognitivos (como a música e a dança), sugere que essa capacidade não é específica da linguagem (em seu sentido estrito) e pode ser uma característica compartilhada com outras espécies. Por outro lado, Pinker e Jackendoff (2005) argumentam que as propriedades rítmicas da linguagem são exclusivamente humanas e específicas da linguagem, uma vez que observações informais sugerem que nenhum outro primata pode ser facilmente treinado para movimentar-se rítmicamente ao som de batidas regulares, o que, segundo os autores, representa uma característica elementar da maneira como os seres humanos reagem aos padrões rítmicos e se manifesta espontaneamente nas crianças.

No dicionário Michaelis (2021), as seguintes definições gerais de ritmo podem ser encontradas:

Sucessão de tempos fortes e fracos que ocorrem alternadamente, com intervalos regulares, em uma frase musical, um verso etc; movimento regular e periódico no curso de qualquer processo; cadência. (MICHAELIS, 2021).

Já no âmbito específico da linguística, Crystal (2000) define o ritmo da seguinte forma:

A aplicação do sentido geral do termo na fonologia se refere a uma regularidade percebida nas unidades proeminentes da fala. estas regularidades podem ser expressas em termos dos seguintes padrões: sílabas acentuadas x não-acentuadas, extensão da sílaba (longa X breve) ou *pitch* (alto X baixo) – ou uma combinação destas variáveis. (CRYSTAL, 2000, p. 230).

Tanto na definição geral quanto na definição específica supracitadas, o ritmo associa-se às noções de alternância, proeminência e regularidade. Neste trabalho, adotaremos uma perspectiva multidimensional para o ritmo, situando-o como função da distribuição de elementos de proeminência no sinal acústico, que envolve várias dimensões acústicas e pode sofrer influência da língua nativa do falante (CUMMING, 2010; LOUKINA *et al.*, 2011; FUCHS, 2016; SILVA JR. & BARBOSA, 2019). Assim, assume-se que a alternância de unidades mais e menos proeminentes em uma língua só pode ser totalmente capturada se diferentes correlatos e dimensões de proeminência forem considerados. É essa perspectiva não categórica, gradiente e multidimensional que orienta a presente investigação, corroborando também a definição de ritmo proposta por Barbosa e Bailly (1994) que concebe o ritmo como a sensação causada pela sucessão de diferentes graus de proeminências, e que será adotada neste trabalho.

Há pelos menos três ondas de pesquisas em ritmo na linguística. A primeira delas baseava-se na Hipótese da Isocronia, segundo a qual os elementos proeminentes no sinal acústico (i.e., as sílabas ou os grupos acentuais) teriam igual duração e as línguas do mundo deveriam pertencer a uma ou outra categoria rítmica. A segunda onda de estudos em ritmo passou a incorporar uma perspectiva mais gradiente e variável, investigando o ritmo a partir de índices estatísticos denominados métricas rítmicas que computam a variabilidade de intervalos de referência (vocálicos, consonantais, silábicos, etc, cf. seção 2.3) a partir do correlato de proeminência da duração. Por fim, uma terceira onda de pesquisas, em que o presente estudo se insere, busca investigar o ritmo a partir também de outros correlatos de proeminência que não a duração, como intensidade e frequência fundamental - f_0 .

2.2 O ritmo como isocronia

O modo de “fluir” da fala humana varia nas diferentes línguas do mundo. Essa ideia

está presente na clássica metáfora de Lloyd James (1940), citado por Abercrombie (1967), que se utiliza da dicotomia código Morse x metralhadora para ilustrar a diferença do ponto de vista impressionístico entre inglês e espanhol, respectivamente. Pike (1945) observa que essa diferença poderia ser explicada remetendo aos elementos recorrentes em ambas as línguas: no inglês, os intervalos interacentuais, e, no espanhol, as sílabas. Essa observação deu origem à dicotomia clássica entre línguas de ritmo silábico e línguas de ritmo acentual. Mais tarde, Abercrombie (1967) acrescentou o conceito de isocronia à teoria das diferenças tipológicas entre as línguas do mundo, inaugurando a Hipótese da Isocronia. De acordo com essa posição teórica, as unidades rítmicas são unidades isócronas, ou seja, de mesma duração. Assim, as diferenças entre línguas de ritmo acentual e de ritmo silábico podem ser explicadas pelas diferentes unidades que apresentam regularidade duracional nessas línguas. No inglês, portanto, as unidades isócronas seriam os intervalos interacentuais e no espanhol as unidades isócronas seriam as sílabas. Há ainda as línguas de ritmo moraico, como o Japonês, que possuem a mora como unidade isócrona (LADEFOGED 1975; LADEFOGED; JOHNSON, 2014). Nesse sentido, o ritmo era concebido categoricamente, ou seja, cada língua natural deveria figurar em uma ou outra tipologia rítmica:

Estas categorias são mutualmente exclusivas porque, se intervalos interacentuais contém um número diferente de sílabas, essas sílabas devem ser comprimidas ou estendidas para manter o distanciamento igual entre os acentos de uma língua de ritmo acentual; ao contrário, em uma língua de ritmo silábico, a distância entre as sílabas acentuadas variaria dependendo do número de sílabas intervenientes.² (DAUER, 1983, p. 51, tradução nossa)

Entretanto, diversos experimentos demonstraram a implausibilidade da Hipótese da Isocronia. Conforme Dellwo (2006),

A atribuição de classes rítmicas a línguas particulares revelou-se basear-se meramente em intuições, pois um vasto número de tentativas de encontrar correlatos acústicos para a isocronia nas línguas dessas respectivas classes, conduzidas principalmente nas décadas de 1970 e 1980 permaneceram sem sucesso.³ (DELLWO, 2006, p. 231, tradução nossa)

² No original: *These categories are mutually exclusive, since if interstress intervals contain different numbers of syllables, these syllables must be either crushed together or spread out in order to maintain equal spacing of stresses in a stress-timed language; whereas, in a syllable-timed language the distance between stressed syllables would vary depending on the number of intervening syllables.*

³ No original: *The attribution of rhythm-classes to particular languages turned out to be based solely on intuitions since a vast amount of attempts to find acoustic correlates for isochrony in languages of the respective classes carried out mainly in the 1970s and 80s remained without success.*

Dauer (1983) comparou dados fonéticos da leitura de textos do inglês, tailandês, espanhol, italiano e grego e concluiu, por meio da análise das médias e do desvio-padrão dos intervalos interacentuais, que essas unidades eram tão variáveis em línguas consideradas de ritmo acentual (inglês) quanto em línguas consideradas de ritmo silábico (espanhol) e nas demais línguas supracitadas, não classificadas até então.

Borzone de Manrique e Signorini (1983) atestaram uma alta variabilidade na duração dos intervalos silábicos do espanhol, concluindo que isso ocorre devido a diversos fatores: acentuação, estrutura silábica, precedência de pausas e taxa de elocução. Além de desencorajar a hipótese de isossilabidade no espanhol, essas autoras também encontraram indícios de relativa regularidade nos intervalos interacentuais dessa língua, com durações agrupando-se em um intervalo entre 0.2 e 0.8 segundos, considerado por Allen (1975) apud Borzone de Manrique e Signorini (1983) como característico de línguas de ritmo acentual. A isocronia também pressupõe que a duração dos intervalos entre as acentuações deveria ser diretamente proporcional ao número de sílabas contidas no pé silábico e, para uma suposta regularidade entre pés silábicos, seria necessário que tanto pés polissilábicos quanto pés monossilábicos fossem mais ou menos simétricos, ocasionando a redução silábica para os primeiros e a extensão da duração silábica nos últimos, o que nem sempre ocorre (FUCHS, 2016).

Portanto, pode-se dizer que os intervalos interacentuais, silábicos ou moraicos não representam um correlato fonético que caracteriza o ritmo, pois não possuem duração regular. Que características no sinal da fala seriam então responsáveis pelo efeito código Morse (línguas de ritmo acentual) e metralhadora (línguas de ritmo silábico), já que a isocronia, por si só, não parece dar conta deste efeito em termos experimentais? Que fatores estabelecem ordem nesse nível? Qual é a realidade física do ritmo? Foram perguntas como essas que, diante da insuficiência de dados empíricos indicadores da plausibilidade da isocronia, levaram pesquisadores a buscarem outras abordagens para o ritmo.

2.3 O ritmo como gradiência

Dauer (1983) questiona a dicotomia entre acentual e silábico ao sugerir que as diferenças rítmicas entre as línguas se baseiam na proeminência do acento. Isso significa que em línguas como o inglês, a atribuição de acento carrega muito mais proeminência do que em línguas como o francês ou o italiano. Nesse sentido, conforme assevera Arvaniti (2012), Dauer dissocia o ritmo da noção de tempo e sugere que as línguas variam quanto ao ritmo em termos

do menor ou maior grau de proeminência do acento. Dauer (1983) propõe então que o ritmo é função da combinação de três fatores: estrutura silábica, redução vocálica e acento. De acordo com essa proposta, as línguas de ritmo mais acentual tendem a apresentar: maior variabilidade na estrutura silábica, e, conseqüentemente, maior variabilidade em termos de duração de intervalos consonantais; correlação entre acentuação e peso silábico; e redução vocálica de sílabas não acentuadas, com vogais curtas ou mesmo ausentes. O oposto parece caracterizar línguas de ritmo menos acentual: menor variabilidade de estruturas silábicas (por exemplo, maior prevalência do tipo CV), acento e peso silábico independentes e menor redução vocálica. A combinação dessas características causaria a impressão rítmica. Tais propriedades são independentes e cumulativas, podendo ocorrer em diferentes níveis de gradiência, e não de modo categórico.

Essas considerações foram refinadas em um modelo unidimensional de ritmo (DAUER, 1987), em que as categorias mais acentual e menos acentual ocupam, respectivamente, um extremo de um *continuum*. Nesse sentido, o ritmo é um efeito global que envolve os seguintes componentes e subcomponentes: a) extensão (*length*), que consiste em duração silábica e vocálica, estrutura silábica e quantidade silábica como componentes; b) *pitch*, que se distingue nas línguas por meio de diferentes padrões entonacionais e tonais; c) qualidade vocálica e consonantal; e d) função linguística do acento. Nesse modelo fonológico de ritmo, as línguas naturais podem ser avaliadas em termos rítmicos por meio da atribuição de scores do tipo (+), 0 ou (-), em que uma maior quantidade de (+) indicaria maior tendência ao ritmo acentual. A gradiência do modelo permite acomodar desde línguas prototipicamente acentuais ou silábicas a línguas que compartilham tanto características de ambos os extremos, como o Catalão [(-) complexidade de estrutura silábica/ (+) redução vocálica] e o Polonês [(+ complexidade de estrutura silábica/ (-) redução vocálica)] (Ramus, Nespor e Mehler, 1999). Esse modelo representa um avanço, pois rompe com a ideia binária de ritmo e admite um *continuum*. Outras abordagens para o ritmo, como a Hipótese de Controle/Compensação, os Modelos Dinâmicos de Ritmo, com a utilização de osciladores, e o advento das métricas rítmicas preveem uma gradiência real sem dissociar o ritmo do fator tempo.

A Hipótese Controle/Compensação começa a ser desenhada também como uma alternativa à dicotomia categórica silábico-accentual. Bertinetto e Fowler (1989) sugerem, com base no trabalho seminal de Fowler (1977), que as diferenças rítmicas poderiam basear-se em mecanismos de controle de alguns fenômenos inerentes ao ritmo presentes em todas as línguas naturais, como a alternância de estruturas coordenativas especializadas para a produção em separado (mas parcialmente sobreposta) de vogais e consoantes, ou a regulação cíclica da

produção das vogais. O refinamento progressivo dessa proposta (BERTINETTO; BERTINI, 2008, 2010) culminou na Hipótese Controle/Compensação (CC), que sugere que as línguas podem diferir em termos de como os gestos articulatórios são acopladas no fluxo articulatório. Uma língua-controle ideal (doravante CTL) seria aquela em que todos os segmentos recebem o mesmo grau de esforço articulatório e tenderiam idealmente à mesma duração. Processos fonológicos de reforço do arranjo silábico canônico, com a prevalência do tipo CV, seriam exemplos de mecanismos implementadores de controle. Em uma língua-compensação ideal (doravante CPS), em contrapartida, os segmentos recebem um esforço articulatório muito variável e tendem idealmente a não ter a mesma duração, com altos níveis de coarticulação. A redução silábica seria um processo de compensação. Como nenhum desses tipos ideais é plausível, e as línguas naturais fazem uso em maior ou menor grau de mecanismos de controle e compensação, essa hipótese prevê gradiência. Bertinetto e Bertini (2008) propuseram uma métrica rítmica denominada CCI, que é o índice de controle-compensação (*control-compensation index*), e computa a média das diferenças entre intervalos vocálicos e consonantais sucessivos dividido pelo número de segmentos no respectivo intervalo, multiplicado por 100. Ainda segundo Bertinetto e Bertini (2010), a Hipótese Controle/Compensação tem como objetivo descrever o comportamento intrasilábico, que por sua vez afeta (e é possivelmente afetado) pela alternância acentual. Neste sentido, há convergência entre a Hipótese Controle/Compensação e o Modelo Dinâmico de Ritmo que será descrito a seguir.

No Modelo Dinâmico de Ritmo discutido em Barbosa (2000, 2006) o ritmo é concebido como função da influência mútua entre um oscilador silábico, “que é um movimento alternante, em termos de picos de intensidade sonora ou de abertura mandibular” (BARBOSA, 2000, p. 13) e um oscilador acentual, cujo correlato articulatório consiste no deslocamento mais proeminente da mandíbula periodicamente. Assim, a oscilação mais rápida, vocálica, caracterizaria o ritmo silábico, e a oscilação mais lenta, o acentual. Entre os dois osciladores, a força de acoplamento, expressa por um valor numérico obtido por análises duracionais de dados de fala, determina a força de influência do oscilador acentual sobre o silábico. Quanto maior a força de acoplamento, maior será a influência do oscilador acentual sobre o silábico, e mais próxima será a língua de uma prototipicamente de ritmo acentual. Por outro lado, quanto menor a força de acoplamento, menor é a influência do oscilador acentual sobre o silábico, e mais próxima estará a língua em análise de uma prototipicamente de ritmo silábico.

Conforme Ordin e Polyanskaya (2015) o ritmo nesse modelo consiste em um fenômeno emergente do controle e da coordenação motora entre certos gestos articulatórios, no

caso, as sílabas fonéticas, ou unidades vogal-vogal (VV). Uma das consequências desse modelo é que ambos os ritmos coexistem numa mesma língua, sendo possível falar em maior ou menor grau de ritmo acentual ou silábico, configurando-se como uma noção rítmica gradiente. Barbosa (2000) pondera que a força de acoplamento também precisa ser observada em relação à taxa de elocução.

Outra perspectiva gradiente de estudo do ritmo surge com o advento das métricas rítmicas. A utilização de métricas rítmicas de maneira bidimensional inicia-se com o estudo seminal de Ramus, Nespore e Mehler (1999), que toma por base uma série de experimentos que abordam o ritmo em seu aspecto perceptual, e no processo de desenvolvimento e processamento da linguagem, citando vários estudos: a) Mehler *et al.* (1996), que pontua que os bebês percebem as vogais, e que a percepção é centrada nas vogais porque estas possuem mais energia e duram mais do que as consoantes, carregam acento e sinalizam a acentuação silábica; b) Bertoncini *et al.* (1988), que sugere que os bebês prestam mais atenção às vogais do que às consoantes; c) Bertoncini e Mehler, (1981); Bertoncini *et al.* (1995); Bijeljic-Babic, Bertoncini, & Mehler, (1993); Van Ooyen *et al.* (1997), que sugerem que os bebês são capazes de perceber o número de sílabas, independentemente da estrutura ou peso silábico. Nesse sentido, os bebês percebem a fala como uma sucessão de vogais que variam em termos de duração e intensidade e que se alternam com períodos de ruídos não analisados. Com base nessa conclusão, a segmentação da fala em consoantes e vogais poderia, segundo Ramus, Nespore e Mehler (1999), dar conta da tipologia rítmica, oferecer subsídios à compreensão da percepção de língua observada nos bebês e elucidar como o ritmo pode ser extraído do sinal da fala.

As diferenças fonológicas explicitadas pelo modelo de Ramus, Nespore e Mehler (1999) formam a base para uma abordagem do ritmo com base em aspectos duracionais, em que as diferenças na duração dos intervalos de referência (vocálicos, consonantais ou silábicos) podem ser mensuradas por meio de índices estatísticos denominados métricas rítmicas. As métricas propostas por Ramus, Nespore e Mehler (1999), ΔC , desvio-padrão da duração dos intervalos consonantais, e %V, proporção da duração dos segmentos vocálicos no enunciado, e todas as métricas que foram propostas posteriormente a partir desse modelo por outros pesquisadores (cf. Quadro 1 – seção 2.4) podem prever gradiência no ritmo porque as línguas podem comportar-se de forma mais silábica em relação a uma métrica, e mais acentual em relação a outra. Algumas dessas métricas serão discutidas na seção 2.4.

2.4 Métricas rítmicas de base duracional

Métricas rítmicas de base duracional podem dar suporte à perspectiva gradiente de ritmo a partir da quantificação de diferenças duracionais entre as línguas, que podem ser computadas tomando como referência intervalos vocálicos, consonantais e silábicos. De acordo com Fuchs (2016), essas diferenças, isto é, a variabilidade, podem ser computadas globalmente, considerando como todos os intervalos de referência do enunciado diferem entre si, ou localmente, a partir do somatório das diferenças dos intervalos adjacentes. Ainda segundo esse autor, “se unidades longas e curtas alternam, isso será capturado por métricas locais. Por outro lado, se unidades curtas e longas se agrupam em diferentes partes do enunciado, isso será refletido nas métricas globais”⁴ (FUCHS, 2016, p. 41, tradução nossa).

As métricas de base duracional também podem tomar como referência outros tipos de intervalo, como os segmentos vozeados e desvozeados, e variar em relação a técnicas de normalização e tratamentos estatísticos empregados. O quadro a seguir resume um grande número de métricas de base duracional disponíveis na literatura:

Quadro 1 – Métricas rítmicas baseadas na duração⁵

Métrica	Descrição	Referência
ΔV	Desvio-padrão da duração de intervalos vocálicos	Ramus, Nespor e Mehler (1999)
ΔC	Desvio-padrão da duração de intervalos consonantais	Ramus, Nespor e Mehler (1999)
Δ Obstruent	Desvio-padrão da duração de intervalos de obstruentes	Steiner (2004, 2005)
%V	Porcentagem da duração total do enunciado composto de intervalos vocálicos	Ramus, Nespor e Mehler (1999)
%n, %l	Porcentagem da duração total do enunciado composto de nasais (n) e laterais (l)	Steiner (2004, 2005)
%Sonorant	Porcentagem da duração total do enunciado composto de intervalos sonoros.	Steiner (2004, 2005)
%Voiced	Porcentagem da duração total do enunciado composto de intervalos vozeados.	Dellwo <i>et al.</i> (2007)
VarcoV	Coeficiente de variação da duração de intervalos vocálicos (i.e. desvio-padrão da duração do intervalo vocálico dividido pela média), multiplicado por 100	White and Mattys (2007a)
VarcoC	Coeficiente de variação da duração de intervalos consonantais (i.e. desvio-padrão da duração do intervalo consonantal dividido pela média), multiplicado por 100	Dellwo (2006)
VarcoVC	Coeficiente de variação da duração combinada de intervalos vocálicos e consonantais sucessivos	Liss <i>et al.</i> (2009)

⁴ No original: (...) *if long and short units alternate, this will be captured well by local metrics. On the other hand, if short and long units cluster at different places in the utterance, this will be reflected by global metrics.*

⁵ Parte das referências do quadro 1 constam apenas para direcionar o leitor aos autores originais dessas propostas. Nem todas essas métricas serão investigadas neste trabalho.

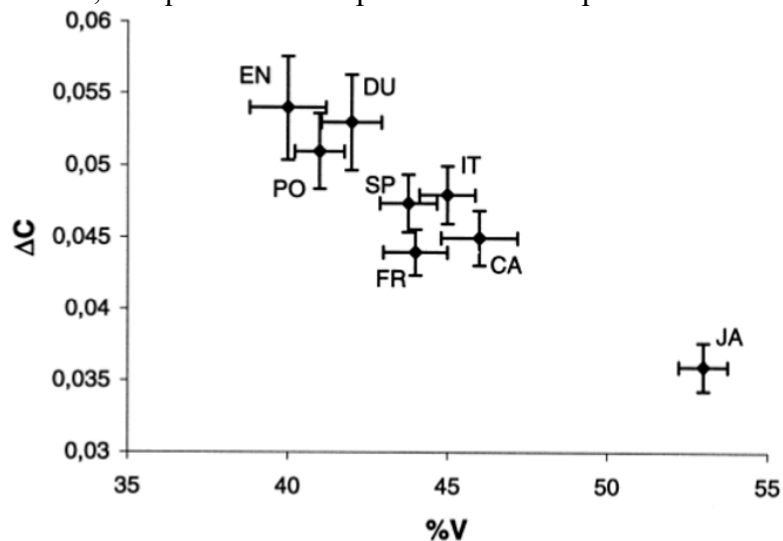
VarcoS	Coeficiente de variação das sílabas (i.e. desvio-padrão da duração do intervalo silábico dividido pela média), multiplicado por 100	Ratheke Smith (2011)
VarcoUV	Coeficiente de variação da duração de intervalos vozeados e desvozeados (i.e. desvio-padrão da duração do intervalo vozeado/desvozeado dividido pela média), multiplicado por 100	Dellwo <i>et al.</i> (2007)
nPVI-V	Índice normalizado de variabilidade pareada para intervalos vocálicos. Média das diferenças entre intervalos vocálicos sucessivos dividido pelo seu somatório, multiplicado por 100	Low, Grabe e Nolan (2000)
rPVI-C	Índice de variabilidade pareada para intervalos consonantais. Média das diferenças entre intervalos consonantais sucessivos.	Low, Grabe e Nolan (2000)
nPVI-VC	Índice normalizado de variabilidade pareada para as durações combinadas de intervalos vocálicos e consonantais sucessivos. Média das diferenças entre intervalos vocálicos e consonantais sucessivos dividido pelo seu somatório, multiplicado por 100	Liss <i>et al.</i> (2009)
nPVI-S	Índice normalizado de variabilidade pareada para sílabas. Média das diferenças entre intervalos silábicos adjacentes, dividido pelo seu somatório, multiplicado por 100	Gibbon and Gut (2001)
RR	Taxa de Ritmo: Índice normalizado de variabilidade pareada para sílabas. Média dos quocientes acoplados das durações de sílabas adjacentes, em que a duração da mais curta é dividida pela duração da mais longa e multiplicada por 100	Gibbon and Gut (2001)
VI	Índice de variabilidade pareada para sílabas. Média das diferenças entre sílabas sucessivas em que a duração de cada sílaba é normalizada por meio da divisão pela média de todas as sílabas	Deterding (1994, 2001)
YARD	Índice de variabilidade pareada para sílabas. Média das diferenças entre sílabas sucessivas em que a duração de cada sílaba é normalizada por meio de transformada z.	Wagner and Dellwo (2004)
Taxa Silábica	Média da taxa de duração de todos os pares silábicos em que a primeira sílaba contém uma vogal plena e a segunda sílaba contém uma vogal reduzida ou deletada	Gut (2003a)
CCI	Índice de variabilidade pareada para vogais ou consoantes. Média das diferenças entre intervalos vocálicos e consonantais sucessivos dividido pelo número de segmentos no respectivo intervalo, multiplicado por 100	Bertinetto and Bertini (2008)
Taxa de elocução	Média do número de unidades articulatórias (segmentos, sílabas, unidades CV) por unidade de tempo	Dellwo (2008)
nPVI-Son	Índice de variabilidade pareada para intervalos sonoros. Média das diferenças entre intervalos sonoros dividido pelo somatório, multiplicado por 100	Fuchs (2016)
rPVI-Obs	Índice de variabilidade pareada para intervalos obstruentes. Média das diferenças entre intervalos de obstruentes sucessivos.	Fuchs (2016)
VarcoSon	Coeficiente de variação da duração dos intervalos sonoros (i.e. desvio-padrão da duração de intervalos sonoros dividido pela média), multiplicado por 100	Fuchs (2016)
VarcoObs	Coeficiente de variação da duração dos intervalos sonoros (i.e. desvio-padrão da duração de intervalos obstruentes dividido pela média), multiplicado por 100	Fuchs (2016)
nPVI-Voiced	Índice de variabilidade pareada para intervalos vozeados. Médias das diferenças entre intervalos vozeados dividido pelo somatório, multiplicado por 100	Fuchs (2016)
rPVI-UV	Índice de variabilidade pareada para intervalos vozeados/desvozeados. Média das diferenças entre intervalos vozeados/desvozeados sucessivos	Fuchs (2016)
VarcoVoiced	Coeficiente de variação da duração dos intervalos vozeados (i.e. desvio-padrão da duração dos intervalos vozeados dividido pela média), multiplicado por 100.	

Fonte: Fuchs (2016, p. 39)

A utilização de métricas rítmicas de maneira bidimensional inicia-se com o estudo seminal de Ramus, Nespor e Mehler (1999). Embora tenha havido propostas de índices

estatísticos para computar o ritmo anteriores a esses autores, como em Fowler (1977) e Deterding (1994), o caráter inovador da proposta de Ramus, Nespore e Mehler (1999) residuiu no fato de terem concebido a noção de espaço rítmico bidimensional gradiente, tomando por base uma série de experimentos psicolinguísticos que forneciam evidências de que os bebês percebem a fala como uma sucessão de vogais que variam em termos de duração e intensidade, e que se alternam com períodos de ruídos não analisados. Na proposta de Ramus, Nespore e Mehler (1999), o desvio-padrão da duração de intervalos consonantais (ΔC) e a proporção da duração total do enunciado composto de intervalos vocálicos (%V) são capazes de diferenciar línguas de ritmo silábico, acentual e moraico, visível se plotadas em um gráfico com ΔC e %V em cada eixo. Ramus, Nespore e Mehler (1999) investigaram, em sua proposta, dados do inglês, polonês e holandês, como representantes da classe ritmo acentual; francês, espanhol, italiano e catalão, representando línguas de ritmo silábico; e japonês, como tendo ritmo moraico. Como pode ser visto na figura 1, reproduzida da publicação original, as métricas propostas foram capazes de agrupar as línguas de cada tipo rítmico dentro de um espaço gradiente⁶.

Figura 1 – Distribuição do inglês (EN), holandês (DU), polonês (PO), espanhol (SP), italiano (IT), francês (FR), catalão (CA) e japonês (JA) em um plano cartesiano com ΔC e %V nos eixos, em que as barras representam o erro-padrão.



Fonte: Ramus, Nespore e Mehler (1999, p. 273)

Os resultados sugerem que ΔC e %V estão relacionados à estrutura silábica das línguas e correlacionam-se negativamente, com valores mais altos de ΔC para línguas que possuem um inventário mais amplo de tipos silábicos, com conseqüente redução de %V nesses casos. O posicionamento, de um lado, do inglês, do holandês e do polonês, línguas com mais

⁶ A diferença entre os três ritmos foi significativa em uma ANOVA e também nos pós-testes pareados de Tukey.

de 15 tipos silábicos, e, do outro, do japonês, que possui quatro tipos silábicos, nos respectivos extremos do plano $\Delta C \times \%V$ evidencia esse contraste e oferece suporte à utilização dessas métricas na categorização de tipologias rítmicas (RAMUS; NESPOR; MEHLER, 1999). Dellwo (2006) endossa a hipótese de trabalho desses autores, sugerindo que o comportamento distinto dessas línguas pode ser explicado com base em suas estruturas silábicas: línguas de ritmo acentual permitem arranjos consonantais mais complexos, o que justifica uma maior variabilidade em termos de duração consonantal e, conseqüentemente, ΔC mais elevado; línguas de ritmo silábico possuem $\%V$ elevado por apresentarem menor incidência de redução vocálica.

Uma outra métrica proposta por Ramus, Nespor e Mehler (1999), o desvio-padrão da duração de intervalos vocálicos (ΔV), mostrou-se, segundo esses autores, menos transparente e conclusiva no comparativo entre as línguas observadas. Segundo os autores, isso acontece porque a variabilidade na duração dos segmentos vocálicos é influenciada pela combinação de fatores fonológicos tais quais redução vocálica, alongamento vocálico contrastivo, alongamento vocálico em contextos específicos e vogais longas.

Grabe e Low (2002) tentaram replicar os resultados de Ramus, Nespor e Mehler (1999) e chegaram a resultados diferentes, o que os levou a concluir que as métricas globais propostas por esses autores não apresentavam robustez suficiente para caracterizar o ritmo. Em um outro estudo, Low, Grabe e Nolan (2014) demonstraram que $\%V$ não foi capaz de diferenciar variedades do inglês consideradas ritmicamente distintas: o inglês britânico (doravante BrE) e o inglês singapuriano (doravante SingE) não apresentaram diferenças em termos de $\%V$. Ainda segundo Dellwo e Wagner (2003), $\%V$ e ΔC podem ser afetados pela taxa de elocução porque em produções orais com taxa de elocução mais elevada, os intervalos consonantais tendem a ser mais curtos, afetando diretamente o escopo de variação das durações.

Como meio de atenuar os efeitos da taxa de elocução, Dellwo (2006) sugeriu que o desvio-padrão da duração dos intervalos consonantais, ΔC , fosse normalizado, dividindo-o pela duração média dos intervalos consonantais. Essa métrica foi chamada de *VarcoC* (*variation coefficient of consonants*), por se tratar do coeficiente de variação da duração dos intervalos consonantais:

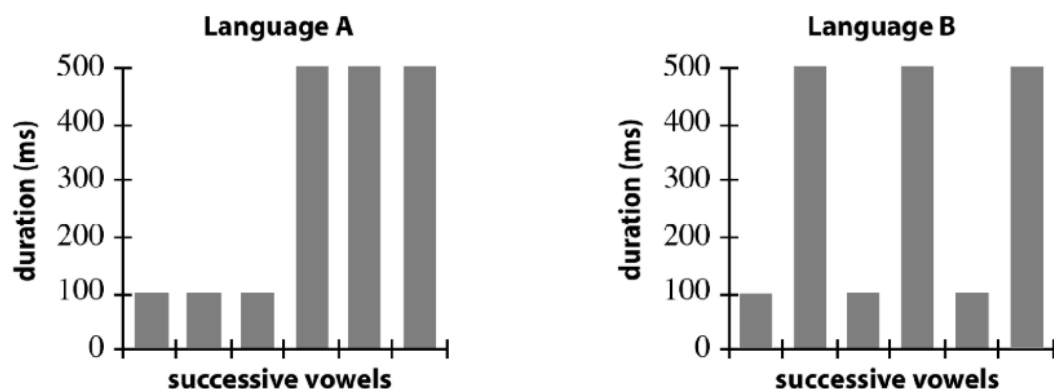
$$\text{VarcoC} = \frac{\Delta C}{\text{média}(C)}, \quad (1)$$

Dellwo (2006) realizou um experimento com VarcoC, envolvendo 12 falantes de alemão e 7 falantes de inglês e francês, lendo um texto de aproximadamente 80 sílabas com cinco diferentes taxas de elocução. Como resultado, a influência da taxa de elocução foi neutralizada para o francês e minimizada para o inglês e o alemão.

O mesmo procedimento passou a ser aplicado também a vogais – VarcoV (WHITE; MATTYS, 2007), a intervalos vocálicos e consonantais sucessivos – VarcoVC (LISS *et al.*, 2009), a sílabas – VarcoS (RATHCKE; SMITH, 2011), e a intervalos vozeados e desvozeados – VarcoUV (DELLWO; FOURCIN; ABBERTON, 2007) a fim de buscar as melhores métricas para a classificação rítmica das línguas.

O que as métricas apresentadas até aqui possuem em comum é o fato de tomarem por base o cálculo de desvio-padrão da duração dos intervalos de referência, sendo consideradas globais porque buscam explicitar como todos esses intervalos diferem entre si no enunciado. Por isso, Low, Grabe e Nolan (2000) argumentam que o cálculo do desvio-padrão da duração de intervalos de referência pode não ser suficiente para distinguir o ritmo de duas línguas ou variedades de uma mesma língua. Para tanto, utilizam-se da representação da duração de segmentos vocálicos consecutivos em duas línguas hipotéticas A e B, conforme figura 2:

Figura 2 – Duração de vogais sucessivas em duas línguas hipotéticas A e B



Fonte: Low, Grabe e Nolan (2000, p. 382)

Em ambas as línguas hipotéticas A e B, o desvio-padrão em relação à média, que é de 300ms, é de 200ms. No entanto, a língua A parece aproximar-se mais de uma língua prototipicamente silábica, devido à tendência de regularidade entre os intervalos vocálicos consecutivos. Mesmo havendo claramente dois blocos de durações distintas, o que pode ser reflexo de uma taxa de elocução mais baixa no início do enunciado, o efeito silábico poderia ser mensurado caso as diferenças entre as sílabas consecutivas fossem computadas. Por outro

lado, a língua B apresenta alternância entre intervalos vocálicos longos e breves, o que seria esperado de uma língua com padrão rítmico acentual. Uma métrica global como ΔV não daria conta dessas diferenças locais, que parecem emergir como importantes indícios para uma categorização rítmica adequada.

Esses autores sugerem uma métrica local como meio de captar as diferenças entre intervalos de referência adjacentes, denominada PVI – *pairwise variability index*, ou, índice de variabilidade pareada, que consiste na média das diferenças entre intervalos de referência sucessivos. Low, Grabe e Nolan (2000) propõem o nPVI-V, cujo ‘V’ indica se tratar da métrica local de intervalos vocálicos, e cujo ‘n’ indica a normalização para atenuar os efeitos da taxa de elocução, novamente realizada pela divisão da métrica pela média de duração dos mesmos intervalos. Todo o cálculo pode ser multiplicado por 100 para ficar na escala de 0 a 100, comumente utilizada nas comparações de métricas:

$$nPVI(V) = 100 \times \frac{\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right|}{m-1} \quad (2)$$

onde m é o número de intervalos vocálicos

e d_k é a duração do k -ésimo intervalo vocálico

Em teoria, valores menores de nPVI-V são esperados para línguas de ritmo silábico, refletindo menor variabilidade entre os intervalos vocálicos consecutivos; e valores maiores são esperados para línguas de ritmo acentual, em decorrência de uma maior variabilidade resultante da alternância entre vogais longas e breves ou reduzidas. Essa hipótese foi testada e corroborada pelos autores em um estudo comparativo entre o inglês britânico e o inglês singapuriano, em que o nPVI-V se mostrou capaz de explicar as diferenças rítmicas entre essas línguas. Outra métrica local, o índice bruto de variabilidade pareada para as consoantes, rPVI-C, também proposta por esses autores, computa a correlação entre padrões rítmicos e variabilidade na duração de intervalos consonantais em diferentes línguas. Ao contrário de nPVI-C, a normalização não é aplicada a essa métrica⁷, tendo em vista que a taxa de elocução não parece afetar esse coeficiente (LOW; GRABE; NOLAN, 2000).

Grabe e Low (2002) examinaram o nPVI-V e o rPVI-C de 18 línguas, incluindo línguas consideradas prototipicamente silábicas, acentuais ou moraicais, e línguas não investigadas até então. As línguas distribuíram-se em zonas diferentes em um plano gradiente

⁷ Por isso o ‘r’ de “raw” (valor bruto).

de possíveis categorias rítmicas. No entanto, conforme pontua Fuchs (2016), menos da metade das comparações das diferenças entre as línguas atingiu significância, o que poderia ser atribuído à quantidade de participantes, apenas um por língua.

Duas outras métricas locais com base em PVI foram propostas por Gibbon e Gut (2001) para o estudo da variabilidade duracional entre sílabas adjacentes: nPVI-S e RR, que é a Taxa de Ritmo, um índice normalizado de variabilidade pareada para sílabas. A primeira é calculada da mesma maneira que nPVI-V, considerando a sílaba como intervalo de referência. Valores mais baixos são esperados para línguas de ritmo silábico e valores mais altos desse índice indicam ritmo acentual. RR diferencia-se de nPVI-S porque não aplica a mesma técnica de normalização. No nPVI-S, a diferença entre dois intervalos silábicos adjacentes é dividida pela média entre esses dois intervalos. Já RR divide a duração da sílaba mais longa pela mais curta, cumulando em um numerador sempre positivo que será dividido por sua vez pelo número total de sílabas menos 1 e multiplicado por 100. Conforme Fuchs (2016), um RR equivalente a 100 corresponderia a um inventário silábico perfeitamente isócrono, enquanto valores mais baixos de RR são esperados para línguas de ritmo acentual. Gibbon e Gut (2001) mediram o RR de falantes de Ibibio, BrE e inglês nigeriano (doravante NgE). Os resultados confirmaram Ibibio como de maior ritmo silábico, mas posicionaram BrE como de ritmo mais silábico que NgE, o que contraria os resultados esperados. Um outro coeficiente de variação silábica também foi proposto por Ackermann and Hertrich (1994) e Deterding (1994). Trata-se do VI (*variability index*), um coeficiente de variabilidade pareada entre sílabas adjacentes que emprega o seguinte método de normalização: a duração de cada sílaba é dividida pela duração média de todas as sílabas do enunciado antes de as diferenças entre as sílabas adjacentes serem computadas. Deterding (1994) aplicou essa métrica ao estudo comparativo entre o ritmo de BrE e SingE e detectou diferenças do ponto de vista rítmico entre as duas variedades do inglês.

Ainda no âmbito dos intervalos silábicos, Wagner e Dellwo (2004) sugeriram o YARD – *Yet Another Rhythm Determination*, uma métrica duracional idealizada para cumprir com os seguintes critérios: ser mais que uma medida de complexidade silábica (segundo esses autores, esse é o caso de ΔC e $\%V$), possuir robustez em relação à taxa de elocução e capturar a natureza sequencial dos padrões rítmicos. O método de normalização empregado nessa métrica é o de z-transformada e consiste na equivalência da média de duração silábica a 0 e o desvio-padrão a 1. As durações normalizadas das sílabas (z_i) da duração absoluta (x_i) são mensuradas como unidades de desvio-padrão (σ):

$$z_i = \frac{x_i - \bar{x}}{\sigma}$$

(3)

As durações normalizadas de cada sílaba são então submetidas ao cálculo do índice YARD, conforme a equação abaixo. Nota-se que o YARD é uma variação das medidas de PVI, diferindo apenas quanto à técnica de normalização:

$$\text{YARD} = \sum_{i=1}^{n-1} / \text{sylldur}(i) - \text{sylldur}(i+1) / \sqrt{n-1} \quad (4)$$

O experimento conduzido por Wagner e Dellwo (2004) envolveu uma sentença traduzida em quatro línguas (inglês, francês, alemão e italiano), lidas com cinco taxas de elocução diferentes. Inglês e alemão registraram os maiores índices de variabilidade (1.20 e 1.18, respectivamente), e o francês e o italiano os menores (0.7 e 0.46, respectivamente), corroborando o posicionamento esperado em relação à literatura sobre ritmo.

Uma outra variante do PVI, o índice normalizado de variabilidade pareada de intervalos vocálicos e consonantais consecutivos, nPVI-VC, extrapola os limites silábicos ao calcular as diferenças entre pares de intervalos vocálicos e intervalos consonantais adjacentes. Liss *et al.* (2009) também propuseram uma métrica global normalizada – VarcoVC para o cálculo do coeficiente de variação entre as durações combinadas de intervalos vocálicos e consonantais sucessivos. O experimento envolveu pacientes acometidos por disartria, condição neurológica que causa uma alteração na pronúncia ou articulação da fala. Das métricas propostas, VarcoVC figurou como o melhor preditor da condição disártrica, sendo eficaz conjuntamente com outras métricas - VarcoV, VarcoC, %V and ΔV na discriminação dos pacientes, com 80% de eficácia, e na identificação do tipo de disartria em questão.

Uma das dificuldades atribuída à manipulação de métricas silábicas é o próprio processo de silabificação, já que não há consenso entre os autores sobre os métodos de silabificação a serem empregados (FUCHS, 2016). Assim, no nível puramente fonético, o método de silabificação poderia considerar a sonorização. Os núcleos silábicos são então divididos em zonas de intensidade, onde os limites das codas e *onsets* silábicos se encontram, e essas zonas constituem uma sílaba foneticamente determinada.

Sonoridade é uma propriedade gradiente, e o consenso entre a maioria dos autores é sobre a seguinte hierarquia, dos segmentos mais sonoros aos menos sonoros: vogais, aproximantes, laterais, nasais, fricativas e plosivas (FUCHS, 2016). Nesse sentido, Steiner (2005) propôs um conjunto de métricas rítmicas baseadas em maior ou menor sonorização de intervalos: %n (proporção de consoantes nasais) e %l (proporção de consoantes laterais) separaram línguas consideradas de ritmos diferentes. O contraste entre intervalos sonoros e

obstruentes também se mostrou promissor: tanto o desvio-padrão da duração de intervalos de obstruentes (Δ Obs) quanto o desvio-padrão da duração de intervalos de sonoras (Δ Son) foram capazes de separar línguas de tipologias rítmicas distintas e diferenciar minimamente línguas com a mesma tipologia rítmica (FUCHS, 2016). Esse autor também propõe que o contraste entre obstruentes e sonoras poderia ser computado com base nas métricas de PVI (n-PVI-Son e r-PVI-Obs) e desvio-padrão normalizado para taxa de elocução (VarcoSon e VarcoObs).

Uma alternativa à silabificação ou ao contraste entre vogais e consoantes é o contraste entre intervalos vozeados e desvozeados, que apresenta inúmeras vantagens metodológicas: menor complexidade na discriminação dos intervalos, maior contraste entre as unidades, menor propensão a erros no processo de segmentação, possibilidade de segmentação automática, com conseqüente economia de tempo, e menor variabilidade entre transcrições. As métricas rítmicas que fazem uso da distinção vozeamento/desvozeamento são: %Voiced, VarcoUV, %Voiced. Em um estudo realizado por Dellwo *et al.* (2007), %Voiced e VarcoUV correlacionaram-se às categorias rítmicas da seguinte forma: inglês e alemão com menor índice de %Voiced e maior índice de VarcoUV, e italiano e espanhol com maior índice de %Voiced e menor índice de VarcoUV (FUCHS, 2016).

Diante da diversidade de métricas rítmicas e das vantagens e desvantagens apresentadas por cada uma delas, a pergunta sobre quais métricas são mais viáveis e realistas para uma tipologia rítmica é inevitável. De acordo com Fuchs (2016), a robustez é um atributo ideal de uma métrica rítmica, ou seja, a não sensibilidade a fatores como taxa de elocução, idiossincrasias do falante e do analista e peculiaridades do texto. Ainda segundo Wiget *et al.* (2010), as métricas rítmicas devem ser utilizadas pelos pesquisadores consoante algumas observações: (i) %V, VarcoV e nPVI-V são mais robustas em relação à variação na taxa de elocução e relativamente em relação aos materiais (estímulos), participantes e transcritores e discriminam bem entre as línguas. É recomendável também utilizar %V em conjunção com VarcoV ou nPVI-V; (ii) métricas não normalizadas para taxa de elocução e métricas consonantais não discriminam bem entre línguas e mostram pouca consistência entre transcritores, falantes e materiais; (iii) os estudos que se pretendem representativos de uma língua precisam ter um número significativo de participantes; (iv) duas estratégias são recomendadas para atenuar a sensibilidade dos indicadores de ritmo aos materiais selecionados: a seleção cuidadosa dos materiais, que devem ser representativos das propriedades métricas e fonológicas da língua, ou a utilização de uma grande quantidade de sentenças; e (v) aplicar princípios norteadores de segmentação entre todos os participantes. Neste estudo, serão utilizadas as métricas duracionais de proporção para segmentos vocálicos e consonantais, e

desvio-padrão, coeficientes de variabilidade (Varcos), índices brutos e normalizados de variabilidade pareada (nPVI e rPVI), taxa de ritmo (RR), índice de variabilidade (VI) e YARD aplicadas aos segmentos vocálicos e/ou consonantais e às unidades VV.

É importante ressaltar, no entanto, que a abordagem e a investigação do ritmo com base em aspectos exclusivamente duracionais têm dado lugar a uma terceira onda de estudos acerca do ritmo, que emprega outros correlatos de proeminência, tais quais frequência fundamental (f_0) e intensidade, já que seus respectivos correlatos perceptuais, *pitch* e altura, também contribuem com a sensação de proeminência e as línguas naturais diferem quanto ao uso desses mecanismos para expressar acento lexical, acento frasal e foco (FUCHS, 2016). Nesse sentido, outras abordagens para o ritmo foram propostas com base em métricas que se utilizam desses parâmetros acústicos.

2.5 Métricas rítmicas com base em f_0 e intensidade,

De acordo com Fuchs (2016), a correlação entre ritmo e diferenças silábicas justifica a inclusão desses correlatos como bases para medidas de ritmo:

Se o ritmo for compreendido com base nas diferenças em proeminência entre as sílabas, então é sensato incluir variação de f_0 e intensidade na mensuração de ritmo. Além disso, a influência de f_0 e intensidade na percepção de proeminência mostra-se menos dependente da língua do que as diferenças interlinguísticas na produção sugerem. Isso justifica o uso de medidas de variação em f_0 e intensidade independentes das línguas.⁸ (FUCHS, 2016, p. 72, tradução nossa)

O quadro a seguir resume um conjunto de métricas rítmicas baseadas em f_0 , intensidade, que foram compiladas por Fuchs (2016) e são baseadas em correlatos de proeminência outros que não a duração:

Quadro 2 – Métricas rítmicas baseadas em f_0 e intensidade⁹

Métrica	Descrição	Referência
nPVI-V (avgInt)	Índice de variabilidade pareada para a variação da intensidade entre intervalos vocálicos. Média das diferenças entre a raiz do valor quadrático médio (RMS) da amplitude de intervalos vocálicos sucessivos.	Low (1998)

⁸ No original: *If rhythm is understood to be based on differences in prominence between syllables, then it is sensible to include variation in f_0 and intensity in the measurement of rhythm. In addition, the influence of f_0 and intensity on the perception of prominence appears to be less language dependent than cross-linguistic differences in production suggest. This justifies the use of language-independent measures of variation in f_0 and intensity.*

⁹ As referências do quadro 2 constam apenas para direcionar o leitor aos autores originais dessas propostas. Essas métricas não serão investigadas neste trabalho.

nPVI-V(AI)	Índice de variabilidade pareada para a variação da intensidade e da duração entre intervalos vocálicos. Média das diferenças entre a integral da amplitude de intervalos vocálicos sucessivos.	Low (1998)
ΔS (avgInt)	Desvio-padrão da média de intensidade silábica	He (2012)
nPVI-S (avgInt)	Coefficiente da variação da intensidade média em sílabas (i.e., desvio-padrão da intensidade média dividido pela média), multiplicado por 100	He (2012)
nPVI-V(f_0)	Índice de variabilidade pareada para a variação na intensidade média. Média das diferenças entre a intensidade média de sílabas adjacentes, dividido por seu somatório, multiplicado por 100	Cumming (2010, 2011)
δS	Média da mudança de sonoridade, indicativo do quão dominantes são os obstruentes	Galves <i>et al.</i> (2002)
\bar{S}	Média da sonoridade	Galves <i>et al.</i> (2002)
nPVI-V (peakInt)	Índice de variabilidade pareada para a variação da intensidade entre intervalos vocálicos. Média das diferenças entre a raiz do valor quadrático médio (RMS) do pico de amplitude de intervalos vocálicos sucessivos.	Fuchs (2016)
nPVI-V (avgLoud)	Índice de variabilidade pareada para a variação do pico de volume entre intervalos vocálicos. Média das diferenças entre a integral do volume de intervalos vocálicos sucessivos	Fuchs (2016)
nPVI-V (peakLoud)	Índice de variabilidade pareada para a variação entre o pico de volume entre intervalos vocálicos. Média das diferenças entre o pico de volume de intervalos vocálicos sucessivos.	Fuchs (2016)
nPVI-V (dur+avgLoud)	Índice de variabilidade pareada para a combinação entre variação de duração e média de volume entre intervalos vocálicos. Média quadrática das diferenças normalizadas entre duração e volume de intervalos vocálicos sucessivos.	Fuchs (2016)
nPVI-V (dur+peakLoud)	Índice de variabilidade pareada para a combinação entre variação de duração e pico de volume entre intervalos vocálicos. Média quadrática das diferenças normalizadas entre duração e pico de volume de intervalos vocálicos sucessivos.	Fuchs (2016)
nPVI-V(LI)	Índice de variabilidade pareada para variação de volume e duração entre as integrais de volume de intervalos vocálicos sucessivos.	Fuchs (2016)

Fonte: Fuchs (2016, p. 78)

Importa compreender o que cada um desses correlatos de proeminência representa em termos físicos. Segundo Crystal (2000, p. 148), a intensidade “marca uma situação de relevo ou proeminência de uma sílaba da palavra.” Esse correlato de proeminência está relacionado à pressão acústica do sinal, isto é, à quantidade de energia que a onda sonora transmite, e

correlaciona-se positivamente com a amplitude: o aumento na amplitude corresponde ao aumento da intensidade, conforme pontua Lehiste (1970):

O afastamento máximo de uma partícula de sua posição de repouso é chamado de amplitude da vibração. A amplitude de uma onda não depende da frequência da onda; intensidade, por outro lado, sim. Se a frequência da oscilação de uma partícula do ar é dobrada, a partícula deve viajar o dobro da distância no mesmo intervalo de tempo.¹⁰ (LEHISTE, 1970, p. 112).

Na proposta de Low (1998), citado por Fuchs (2016), o coeficiente da variação da intensidade média em sílabas, nPVI-V(avgInt), figura como uma métrica local para computar a variabilidade de amplitude entre intervalos vocálicos sucessivos utilizando-se da raiz do valor quadrático médio (RMS) da amplitude de intervalos vocálicos sucessivos. Outros exemplos de métricas com base em intensidade foram propostas por He (2012), como o desvio-padrão da média de intensidade silábica, $\Delta S(\text{avgInt})$, que é uma versão de nPVI-V(avgInt) normalizada para a taxa de elocução e o VarcoS(avgInt), que quantifica a variação na média de intensidade silábica, e nPVI-S(avgInt), que se constitui como uma métrica local normalizada para a variabilidade pareada da intensidade de sílabas adjacentes.

Essas métricas, no entanto, não constituem, segundo Fuchs (2016), medidas psicológicas realistas. Como sons mais intensos são percebidos como mais altos, o que também depende de f_0 e distribuição de energia espectral, métricas como o índice de variabilidade pareada para a variação do pico de volume entre intervalos vocálicos, o nPVI-V(avgLoud), e nPVI-V(peakLoud), o índice de variabilidade pareada para a variação entre o pico de volume entre intervalos vocálicos, são propostas por Fuchs (2016) como potenciais índices estatísticos para captar a variabilidade pareada entre as médias de altura e de picos de altura de sílabas adjacentes. Uma integral de altura, o índice de variabilidade pareada para variação de volume e duração entre as integrais de volume de intervalos vocálicos sucessivos, ou nPVI-V(LI), também foi sugerida por Fuchs (2016) como meio de captar os efeitos de combinação entre duração e altura na variabilidade pareada de segmentos vocálicos. Uma outra métrica divisada por Fuchs (2016), nPVI-V(dur+loud), poderia mensurar em que medida duração e altura se equilibram do ponto de vista perceptual. A eficiência dessas métricas ainda precisa ser testada, tendo em vista que o autor apenas as sugeriu.

¹⁰ No original: *The maximum displacement of the particle from its rest position is called the amplitude of the vibration. The amplitude of a wave does not depend on the frequency of the wave; intensity, however, does. If the frequency of oscillation of a particle of air is doubled, the particle must travel twice as long a distance in the same length of time.*

Outro correlato de proeminência que pode ser útil na determinação do ritmo é a frequência fundamental (f_0). Segundo Barbosa (2012), f_0 é o correlato acústico da frequência de vibração das pregas vocais e seu correlato perceptual é o *pitch*, que está associado às noções de grave e agudo. Cumming (2012, p. 256), com base em experimentos prévios, pontua que “ f_0 e duração são interdependentes na percepção de grupos rítmicos na ritmicidade da fala e da sentença”¹¹. Portanto, há correlação entre f_0 e duração e a quantificação dessa relação pode constituir-se como pista acústica interessante na investigação do ritmo. Com vistas a mensurar a correlação entre duração e f_0 , nPVI-V (f_0), que corresponde ao índice de variabilidade pareada para a variação na intensidade média, foi proposto por Cumming (2010) para se obter o índice normalizado de variabilidade pareada para a variação de f_0 entre segmentos vocálicos adjacentes e também para intervalos silábicos. A autora aplicou essa métrica ao estudo dos padrões rítmicos do alemão suíço, do francês suíço e do francês metropolitano e detectou valores significativamente mais baixos para as duas variedades do francês.

Silva Jr. e Barbosa (2019b) também se utilizam de parâmetros acústicos com base em f_0 no estudo dos padrões rítmicos de aprendizes de inglês-L2 falantes de PB-L1 em comparação com falantes de AmE-L1. Foram utilizados parâmetros acústicos como: f_0 mínimo ($f_{0min.}$), taxa de elocução (SR) e média e desvio-padrão da primeira derivada de f_0 . Os índices desses parâmetros registraram valores significativamente mais elevados para os falantes de AmE-L1, explicitando que o contorno de f_0 para falantes de PB-L1 permaneceu quase monotônico.

Finalmente, a noção de sonoridade, que é definida por Crystal (2000, p.244) como “a altura global de um som em relação a outros do mesmo *pitch*, acento e duração”, que é perceptualmente vinculada à sensação de continuidade, serve de base para as métricas de média de sonoridade (\bar{S}), que representa a soma dos *scores* de sonoridade de todos os intervalos dividida pelo número de intervalos, e média da mudança de sonoridade (δS), que mede a dominância das obstruções na cadeia de fala. Essas métricas foram propostas por Gut (2002) e replicam os resultados de Ramus, Nespore e Mehler (1999), performando melhor em alguns casos (FUCHS, 2016).

Portanto, há vantagens na utilização de parâmetros de proeminência outros que não a duração para a estudo de ritmo. Além disso, um modelo multidimensional de ritmo precisa levar em consideração uma diversidade de métricas e parâmetros acústicos “como meio de

¹¹ No original: *f0 and duration are interdependent in the perception of rhythmic groups in speech and sentence rhythmicity.*

capturar a alternância entre unidades mais proeminentes e menos proeminentes (como sílabas e intervalos vocálicos) na língua”¹² (FUCHS, 2016, p. 80). Essa estratégia também dialoga com a perspectiva gradiente, não categórica e multidimensional de ritmo adotada neste estudo.

2.6 O ritmo do português brasileiro (PB)

Algumas línguas do mundo (ou variedades de uma mesma língua) são consideradas pertencentes a uma ou outra categoria rítmica de maneira relativamente consensual (e quase categórica). É o caso do BrE, do espanhol e do japonês, línguas consideradas exemplares prototípicos dos ritmos acentual, silábico e moraico, respectivamente. No entanto, o PB já foi considerado uma língua com tendência ao ritmo acentual (CAGLIARI, 1981; MAJOR, 1985; MIGLIORINI; MASSINI-CAGLIARI, 2010), silábico (ABURRE-GNERRE, 1981; SILVA JR., 2012; AVERY & EHRLICH, 2012;), e misto (BARBOSA, 2000; BISOL, 2000; FROTA; VIGÁRIO, 2000; FROTA; VIGÁRIO; MARTINS, 2000; LIMA JR.; GARCIA, 2017; VOGEL; ATHANASOPOULOU; GUZZO, 2018). As diferenças explicitam perspectivas teórico-metodológicas distintas e um inventário de processos fonológicos e dados fonético-acústicos que podem configurar-se tanto como favorecedores do ritmo acentual quanto do silábico, a depender do recorte epistemológico empreendido.

Cagliari (1981) parte de uma perspectiva isócrona que caracteriza o ritmo como “um tipo de simetria, uma harmonia resultante de certas combinações e proporções regulares” (CAGLIARI, 1981, p. 123), para classificar o PB como ritmicamente acentual. No entanto, o autor tece considerações acerca do fenômeno de variação linguística – diafásica e diatópica, como fenômeno modulador do ritmo em PB. Como exemplo situacional que evidencia variação diafásica do ritmo, o autor cita a súplica, em que o ritmo silábico é instanciado no ato de fala com objetivo semântico (e, sob nosso olhar, pragmático). O autor também cita a variedade gaúcha como exemplo da ocorrência do ritmo silábico em PB que exemplifica um tipo de variação diatópica do PB quanto ao ritmo.

Major (1981) realizou um experimento com falantes nativos do PB lendo frases-veículo do tipo “Repita a palavra _____ de novo”, com a inserção de palavras monossilábicas, dissilábicas e trissilábicas e, logo em seguida, logatomas. O autor realizou a medição dos intervalos interacentuais e concluiu que havia uma tendência à isocronia para essas unidades, já que a variação aferida entre os intervalos não era significativa do ponto de vista perceptual,

¹² No original: *in order to capture the alternation of prominent and less prominent units (such as syllables or vocalic intervals) in a language.*

e ressaltou uma tendência ainda maior ao ritmo acentual na fala casual. Complementarmente, Major teceu observações fonológicas acerca do PB com base nos dados aferidos para justificar essa classificação:

Acredito que o português possui tendência ao ritmo acentual pelas seguintes razões: (1) as durações dos intervalos interacentuais não são diretamente proporcionais ao número de sílabas; (2) muitas diferenças em durações interacentuais não são perceptíveis; (3) a duração silábica é inversamente proporcional ao número de sílabas na palavra; (4) na fala casual, sílabas átonas são apagadas, o que provoca o efeito de equalização do número de sílabas em cada grupo acentual; e (5) processos de redução (de sílabas átonas), que reduzem a duração, tem efeito de suporte ao ritmo acentual, isto é, alteamento, monotongação e mudanças de silabicidade¹³. (MAJOR, 1981, p. 350, tradução nossa)

No entanto, conforme Barbosa (2000), as primeiras quatro razões apontadas por Major (1981) não parecem representar atributos específicos do PB ou de qualquer outra língua considerada de ritmo acentual ou silábico, mas antes características universais do ritmo de qualquer língua natural. Sobre a última razão, conforme pontua Silva Jr. (2012), processos como a monotongação de /ey/ e /ou/, por exemplo, também podem atuar em favor do ritmo silábico, como em “peixe” → [‘pe.ʃi] e “pouco” → [‘po.ku], em que a monotongação favorece a isocronia silábica.

Migliorini e Massini-Cagliari (2010) consideram que as inconsistências entre os estudos acerca do ritmo com base em análises acústicas encorajam a utilização do nível fonológico como pressuposto para uma categorização rítmica. A análise consistiu na determinação do papel dos processos fonológicos de reforço e redução silábica na categorização do ritmo do PB. Dentre os processos fonológicos de reforço, isto é, acréscimo de segmentos otimizando o arranjo silábico canônico CV, as autoras analisaram a epêntese, e o alongamento e fortalecimento da vogal. Já a análise dos processos de redução, ou seja, perda de material fonético, englobou redução vocálica, redução dos ditongos nasais átonos, síncope em proparoxítonas, e processos de sândi. De acordo com as autoras, o ritmo se constitui no nível pós-lexical e, portanto, somente os processos fonológicos que operam nesse nível são determinantes da categoria rítmica de uma língua. Os processos de redução supracitados ocorrem todos no nível pós-lexical e favorecem o ritmo acentual, segundo as autoras. Sobre o

¹³ No original: *I believe for the following reasons to have shown that Portuguese has a tendency toward stress-timing: (1) interstress durations are not directly proportional to the number of syllables; (2) many differences in interstress durations are not perceptible; (3) syllable duration is inversely proportional to the number of syllables in a word; (4) in casual speech unstressed syllables delete, which has the effect of equalizing the number of syllables in each stress group; and (5) shortening processes (of unstressed syllables), which reduce duration, have the effect of aiding stress-timing, i.e. "raising", "monophthongization", and "syllabicity shifts.*

sândi do PB, as autoras também o consideram favorecedor da implementação do ritmo acentual, porque atuam no sentido de reduzir o material fonético entre as sílabas tônicas.

Em relação aos estudos que apontam maior tendência de ritmo silábico no PB, Aburre-Gnerre (1981) considera que a fala formal do PB apresenta predominância de ritmo silábico, com aplicação de processos fonológicos considerados tipicamente silábicos nesse estilo, dentre os quais a epêntese, o apagamento de consoantes nasais finais em sílabas do tipo CVN, o enfraquecimento ou apagamento do flap ou [s] final, a monotongação e a harmonia silábica, em que “vogais médias fechadas pretônicas realizam-se como vogais médias abertas, harmonizando-se em altura com a vogal aberta da sílaba seguinte” (ABURRE-GNERRE, 1981, p. 27). Esses fenômenos fonológicos evidenciam o ritmo silábico para a fala formal do PB porque favorecem a estrutura silábica CV, que é predominante em línguas com padrão rítmico silábico. A autora salienta ainda que em estilos menos formais, haveria uma tendência ao ritmo acentual, devido ao aumento da taxa de elocução e aos processos de redução vocálica, que ocasionariam o agrupamento de segmentos consonantais ao redor de núcleos acentuados. Nesse sentido, a autora dialoga com a perspectiva de Major (1981), segundo a qual flutuações rítmicas são (também) condicionadas estilisticamente.

Silva Jr. (2013), em sua tese de doutorado, investiga o choque acentual e as estratégias rítmico-prosódicas de resolução do choque acentual empregados por falantes de PB-L1 na realização fonética do inglês-L2. Foram analisados os parâmetros acústicos de duração, altura e intensidade. Verificou-se que, diferentemente dos falantes de inglês-L1 (grupo controle), os falantes de PB-L1 tendem a não retrair o acento no contexto do choque acentual tanto na produção do inglês-L2 quanto do PB-L1. A estratégia utilizada por esses falantes é a inserção de uma batida silenciosa, o que denota a influência rítmica do PB na realização fonética do inglês-L2. O autor conclui que, sob uma perspectiva fonética, os falantes do PB utilizam-se de estratégias de resolução de choque acentual típicas de línguas de ritmo silábico. Nessa mesma linha, o autor cita Avery & Ehrlich (2012), que diferenciam o Português Europeu (doravante PE) do PB pelo fato de o primeiro apresentar redução vocálica. Para esses autores, a ausência de uma vogal equivalente ao *schwa* do inglês interfere nas possibilidades de redução vocálica do PB.

Barbosa (2000) empreende uma análise quantitativa de dados fonético-acústicos do PB com base no Modelo Dinâmico de Ritmo. Os dados analisados do PB demonstraram a existência de enunciados tendendo mais ao ritmo silábico em uma mesma faixa de variação da taxa de elocução. No comparativo com outras línguas, como o espanhol e o finlandês, o PB aparece semelhante em termos rítmicos, ocupando uma posição fronteira entre ritmo acentual

e silábico, com tendência ligeiramente mais silábica. Em taxas de elocução mais rápida, o PB figura como mais acentual que o grego e o italiano, porém mais silábico em relação ao thai e ao BrE. Portanto, os padrões rítmicos do PB flutuam a depender da taxa de elocução.

Frota e Vigário (2000) e Frota, Vigário e Martins (2002) também empreenderam uma análise do ritmo do PB com base em dados fonéticos. Com base em características entonacionais e na aplicação das métricas %V e ΔC , propostas por Ramus, Nespor e Mehler (1999), as autoras chegaram à conclusão de que o PB e o Português Europeu (PE) são línguas rítmicamente diferentes e apresentam pistas rítmicas mistas no sinal acústico.

Lima Jr. e Garcia (2017) realizaram um estudo preliminar acerca do ritmo do PB com base em dados fonéticos. Os autores investigaram as diferenças nas durações das sílabas acentuadas e não acentuadas na realização fonética do PB e do AmE e concluíram que há indícios de ritmo misto no PB, o que favorece a redução de sílabas não acentuadas por aprendizes falantes de PB-L1 na realização fonética do inglês-L2.

Na mesma direção, Vogel, Athanasopoulou e Guzzo (2018) compararam dados fonéticos do Espanhol Europeu (doravante EE) e do PB e concluíram que, diferentemente do Espanhol Europeu, os falantes do PB utilizam-se do correlato de proeminência da duração como pista acústica para marcar o contraste entre sílabas acentuadas e não acentuadas, independentemente de tratar-se de palavra ocupando posição de proeminência frasal. Além disso, os falantes do EE ao utilizarem-se do correlato da duração para atribuir ênfase a um determinado termo do enunciado, o fazem sem diferenciar o contraste entre sílabas acentuadas e não acentuadas, isto é, ambas as sílabas são alongadas. Por contraste, o PB discrimina entre sílabas acentuadas e não acentuadas presentes em palavras que ocupam posição de proeminência frasal, registrando uma média duracional duas vezes maior para as sílabas acentuadas nesta situação. Diante desses dados, as autoras sugerem que o PB apresenta propriedades fonético-fonológicas que tendem ao ritmo acentual ou pelo menos a não silabidade rítmica.

Finalmente, Bisol (2000) defende uma tipologia rítmica mista para o PB com base em uma análise fonológica sob o enfoque da Fonologia Métrica. Segundo a autora, processos fonológicos como redução vocálica, atribuição de acento, haplogia, degeminação e elisão denotam um ritmo misto associado ao PB, com forte tendência ao ritmo silábico. A autora enfatiza ainda o papel do troqueu silábico, presente em frases neutras, como importante elemento no sistema e no ritmo do PB.

Do exposto, conclui-se que o ritmo do PB tem sido analisado tanto sob uma perspectiva fonética quanto fonológica, chegando a resultados muitas vezes conflitantes. Neste

trabalho, adotaremos uma perspectiva fonética em relação à caracterização do ritmo do PB, em uma abordagem quantitativa e multidimensional. Argumentamos, consoante Frota, Vigário e Martins (2002), que uma abordagem puramente fonológica do ritmo não parece explicar como falantes conseguem discriminar línguas ritmicamente distintas quando expostos à fala filtrada, composta de pistas exclusivamente prosódicas (RAMUS; NESPOR; MEHLER, 1999) ou como os bebês também o conseguem (NAZZI *et al.* 1998). Esses estudos indicam que as diferenças rítmicas entre as línguas podem estar codificadas em pistas acústicas presentes no sinal de fala e podem ser investigadas a partir de dados fonéticos.

2.7 O ritmo do inglês

O ritmo do inglês tem sido exaustivamente investigado desde as primeiras propostas acerca das tipologias rítmicas, sendo categorizado como prototipicamente acentual. No modelo fonológico de Dauer (1987), o inglês ocupa a posição mais acentual no *continuum*, pois marca quase todos os scores do modelo com (+). No componente extensão, sílabas e especialmente vogais tônicas são mais longas que as átonas (+duração) e a língua apresenta grande variedade de tipos silábicos (+estrutura silábica). Já em *pitch*, sílabas acentuadas representam pontos de mudança no contorno entonacional (+entonação). No aspecto qualidade, o inglês apresenta um vasto inventário de segmentos vocálicos (+vogais), e segmentos consonantais são mais precisamente articulados em sílabas acentuadas, podendo possuir alofones reduzidos ou serem neutralizados em sílabas átonas (+consoante). Finalmente, o componente acento em inglês possui função distintiva e seu deslocamento de uma sílaba para outra pode acarretar mudanças semânticas (+função do acento). De acordo com esse modelo, aspectos fonológicos do inglês tendem a favorecer a implementação do ritmo acentual.

Os diversos estudos que utilizam métricas rítmicas duracionais também têm classificado o inglês como uma língua de ritmo acentual e, em certa medida, confirmado a influência de alguns desses aspectos fonológicos. Ramus, Nespor e Mehler (1999) classificam o inglês como de ritmo acentual uma vez que há uma correlação negativa para o inglês entre ΔC e %V (cf. Quadro 1), com maior ΔC indicando maior variabilidade na duração dos segmentos consonantais e menor %V denotando processos como redução e apagamento vocálico em sílabas átonas.

Em estudos empregando métricas globais normalizadas, como o VarcoC (cf. Quadro 1), o inglês apresentou maior índice de variabilidade que o francês (Dellwo, 2006), agrupando-se no plano formado pelos eixos VarcoC e %V juntamente com o alemão, outra

língua considerada de ritmo acentual. Em estudos empregando métricas locais, com base em PVI (cf. Quadro 1), pode-se citar Grabe e Low (2002), em que o BrE apresentou valores mais altos de nPVI-V e rPVI-C, agrupando-se com línguas como o holandês e o alemão, também de ritmo acentual. Esses estudos exemplificam a categorização rítmica do inglês como acentual com base em dados fonéticos e índices quantitativos baseados no correlato de duração.

Entretanto, o emprego experimental de métricas rítmicas nem sempre chegou a resultados consistentes e isso também se aplica ao inglês. Arvaniti (2012) utilizou as métricas rítmicas de ΔC , %V, PVIs and Varcos com o inglês, o alemão, o grego, o italiano, o coreano e o espanhol na análise da fala em três modos de elocução distintos (fala espontânea; leitura de narrativa e leitura de sentenças), com manipulações aplicadas aos materiais no sentido de maximizar ou minimizar a complexidade silábica. As classificações foram inconsistentes e as métricas variaram entre os participantes e mostraram-se sensíveis ao tipo de texto e à estrutura silábica. Em relação ao inglês, VarcoC, por exemplo, não distinguiu essa língua em uma categoria diferente do italiano ou do coreano. Esse estudo sugere que as escolhas metodológicas podem impactar os resultados métricos e que a manipulação da estrutura silábica pode influenciar os resultados independentemente da tipologia rítmica, demonstrando que as métricas são medidas sensíveis ao “ruído” nos dados.

Além disso, conforme Barbosa (2000), há que se considerar a tipologia rítmica em relação às variedades dialetais sob análise porque essas variedades podem ter características rítmicas diferentes. Conforme Fuchs (2016), a atribuição do ritmo acentual ao inglês parece estar associada às variedades pertencentes ao chamado *inner-circle*. Por outro lado, evidências empíricas advindas de uma série estudos listados pelo autor (FUCHS, 2016, p. 88, 89) sugerem que as variedades dialetais do inglês pertencentes ao chamado *outer-circle* demonstram tendência ao ritmo silábico.

Deterding (1994), por exemplo, utilizou-se do VI para mensurar as diferenças rítmicas entre BrE e o SingE e constatou que há maior variabilidade em relação às durações silábicas do BrE, o que indica maior tendência ao ritmo silábico no SingE. Essa tendência foi corroborada por Low, Grabe e Nolan (2000), que encontraram evidências métricas de ritmo mais silábico na variedade SingE, com menor variedade de duração entre vogais adjacentes e padrões formânticos F1 e F2 em vogais reduzidas mais periféricas que no BrE. Fuchs (2016) também menciona Mesthrie (2008), que realiza um inventário fonológico das variedades do inglês no continente africano e das regiões sul e sudeste da Ásia, concluindo que essas variedades possuem tendência ao ritmo silábico.

Ademais, línguas de contato entre o inglês e uma língua de ritmo silábico parecem emergir como silábicas e normalmente o BrE e o AmE são utilizados como referenciais comparativos. Outrossim, essas variedades compartilham de certos atributos fonológicos, como o alongamento de vogais tônicas, o que pode posicioná-las como intermediárias no *continuum* silábico-acental, desencorajando uma classificação categórica. O presente estudo pode acrescentar à literatura sobre o ritmo manifestado nas diversas variedades da língua inglesa já que é possível considerar o inglês-L2 falado por aprendizes brasileiros como uma variedade do chamado *expanding circle*.

Há evidências também de uma tendência ao ritmo silábico em variedades do inglês pertencentes ao *inner-circle*. Ferragne (2008), citado por Fuchs (2016), compara algumas variedades do inglês falado nas diversas ilhas britânicas por meio da medição de métricas como o nPVI-V, VarcoV, %V e rPVI-C (cf. Quadro 1), e demonstra que esses dialetos apresentam variação rítmica em comparação com o padrão acental prototípico associado ao BrE. As variedades de Ulster e Cornwall, por exemplo, apresentam menor variabilidade na duração dos intervalos vocálicos em comparação com outras variedades. Já White e Mattys (2007) apresentam evidências de variabilidade ainda menor desses índices nas variedades de Bristol e da região de Welsh Valleys.

Portanto, o ritmo do inglês é tradicionalmente classificado como acental, mas diferenças dialetais e uma perspectiva de gradiência adotada por linhas de investigação mais contemporâneas permitem conceber flutuações rítmicas nessa língua.

2.8 O ritmo em pesquisas sobre o desenvolvimento de L2

No âmbito do desenvolvimento de L2, os estudos acerca do ritmo giram em torno de três questões fundamentais: a) O ritmo de L2 pode ser adquirido?; b) Qual é o percurso de desenvolvimento do ritmo em L2?; c) Qual é a influência do ritmo da L1 sobre o ritmo da L2?

Adams (1979) realizou um estudo abrangente acerca do desenvolvimento do ritmo do inglês por aprendizes com diferentes L1s. A autora atribuiu o grau de acento detectado no ritmo da produção desses aprendizes a fatores como: contraste duracional insuficiente entre sílabas acentuadas e não acentuadas; ausência de mecanismos fonológicos de juntura, isto é, transição entre os sons, em que um som é introduzido (ou apagado) no final de uma palavra dependendo da palavra seguinte; pausas inadequadas; e posicionamento incorreto de acento. Esses fatores poderiam indicar transferência de padrões rítmicos da L1 para a L2. Vale ressaltar que Adams utilizou-se de uma abordagem multidimensional de ritmo, por meio da investigação

de diferentes correlatos de proeminência – f_0 , duração e intensidade, antecipando uma tendência mais contemporânea de pesquisas sobre o ritmo.

Outros estudos debruçaram-se sobre o desenvolvimento do ritmo em L2 por meio das métricas rítmicas previamente discutidas, variando em função do grau de similaridade e dissimilaridade entre L1 e L2, das medidas utilizadas, e dos procedimentos de coleta e análise de dados. Lin e Wang (2005), por exemplo, realizaram uma investigação acerca dos padrões rítmicos manifestados na interlíngua de aprendizes de inglês-L2 falantes de mandarim-L1, com o objetivo de confirmar a tipologia silábica da L1, atestar as diferenças rítmicas entre mandarim-L1 e inglês-L1, e compreender o processo de desenvolvimento dos padrões rítmicos manifestados pela interlíngua. Para tanto, aplicaram as métricas de Ramus, Nespor e Mehler (1999), realizando as medições de ΔC e %V das respectivas L1s e interlíngua. Os resultados confirmaram as diferenças rítmicas entre as L1s tanto em termos de proporção de segmentos vocálicos, com o mandarim apresentando maior proporção, quanto de variação na duração dos intervalos consonantais, com maior variabilidade para o inglês.

Já o comparativo entre ritmo da interlíngua dos aprendizes e o inglês apresentou resultados surpreendentes. Demonstrou-se não haver diferença significativa entre os padrões rítmicos do inglês-L1 e inglês-L2, exceto pela proporção de segmentos vocálicos na situação de leitura de texto, sugerindo que o inglês-L2 de aprendizes falantes de mandarim-L1 aproxima-se mais ritmicamente da L2 que da L1, o que poderia indicar uma correlação entre desenvolvimento do ritmo da L2 e nível de proficiência. O posicionamento do mandarim como língua de ritmo silábico também foi constatado por Ding e Xu (2016) em um estudo comparativo entre mandarim-L1, inglês-L1 e inglês-L2. Os resultados sugerem que a aplicação de métricas duracionais que tomam por referência os segmentos vocálicos figuram como melhores indicadores do ritmo silábico manifestado na realização fonética do inglês-L2 por falantes de mandarim-L1.

Whitworth (2002), por sua vez, realizou um estudo com crianças bilíngues falantes de inglês e alemão para investigar, entre outras questões, se haveria o desenvolvimento de padrões rítmicos diferentes para ambas as línguas, mesmo tratando-se, nesse caso, de línguas consideradas ritmicamente similares, e de que maneira essas produções fonéticas relacionam-se às produções de falantes monolíngues da L1 e da L2. A análise utilizou-se do rPVI-C e nPVI-V (cf. Quadro 1). Os resultados sugeriram que as crianças bilíngues são capazes de detectar diferenças rítmicas sutis no *input* linguístico e são capazes de produzir padrões rítmicos diferenciados, que não necessariamente coincidem com aqueles apresentados por monolíngues, pelo menos do ponto de vista acústico. Os resultados também sugerem que crianças bilíngues

passam por uma fase de produção fonética com ritmo mais silábico, mesmo quando expostas a línguas ritmicamente similares e acentuais. O mesmo parece ocorrer em bilíngues expostos a línguas ritmicamente diferentes (BUNTA; INGRAM, 2007; LLÉO; RAKOW; KEHOE, 2007). Essas últimas conclusões dão suporte a outros autores para os quais o percurso do desenvolvimento do ritmo em L2 segue padrões universais ortogonais à L1, indo do silábico ao acentual, independentemente da similaridade rítmica entre L1 e L2.

Essa hipótese foi testada por Ordin e Polyanskaya (2015), que analisaram os padrões rítmicos na realização fonética do inglês-L2 por aprendizes falantes de alemão-L1, uma língua ritmicamente similar ao inglês; e francês-L1, ritmicamente dissimilar, com o objetivo de investigar o desenvolvimento rítmico como função do desenvolvimento da L2. Foram utilizadas métricas rítmicas com base em intervalos vocálicos (nPVI-V, VarcoV), consonantais (nPVI-C e VarcoC), e silábicos (nPVI-S e VarcoS), assim como o cálculo da média de duração silábica (meanS) de cada sentença elicitada por cada falante nos contextos de fala espontânea e não controlada. Análises de Variância (ANOVA) foram implementadas com L1 e nível de proficiência como variáveis preditoras e meanS como variável de resposta. Constatou-se efeito significativo mais robusto entre nível de proficiência e meanS, e não houve interação significativa entre L1 e nível de proficiência, demonstrando taxa de mudança nas durações silábicas similar em ambos os grupos de aprendizes. Outros modelos de regressão implementados nesse estudo demonstraram que tanto a L1 como o nível de proficiência podem exercer efeito sobre as métricas rítmicas.

As investigações sobre o desenvolvimento do ritmo em L2 apresentadas até aqui utilizaram-se de um desenho transversal. Como exemplo de estudo longitudinal do ritmo, White e Mok (2019) investigaram o desenvolvimento rítmico do inglês-L2 de imigrantes de Hong Kong falantes de cantonês-L1 em países falantes de inglês-L1. O estudo também considerou o grau de utilização da L1 e da L2 durante o tempo de estadia no país estrangeiro. Foram utilizadas as métricas de taxa de elocução, que corresponde à média do número de unidades articulatórias (segmentos, sílabas, unidades CV) por unidade de tempo (DELLWO, 2008), Varco-C e PVI-V. Demonstrou-se que o ritmo manifestado na L2 poderia tornar-se mais acentual dentro de certas condições sociais e fonéticas: acomodação da L2, configuração dialetal da população do local de estadia e a demografia da cidade. Tais fatores demonstraram interferir no maior ou menor uso da L2 pelos participantes, afetando o desenvolvimento do ritmo acentual em maior ou menor escala no curso do tempo.

Como exemplo de estudo envolvendo aprendizes falantes de português do Brasil (PB-L1) e inglês-L2, Silva Jr. e Barbosa (2019b) realizaram uma investigação sobre a

realização do ritmo do inglês-L2 por falantes de PB-L1 na produção de fala semi-espontânea. Foram analisadas métricas e parâmetros acústicos com base em duração e f_0 , e a correlação desses índices com grau de acento. A hipótese de trabalho era que o ritmo do inglês-L2 apresentaria taxa de elocução mais baixa, devido à falta de fluência; menor variabilidade silábica, pela influência de padrões mais regulares de intervalos silábicos e grupos acentuais; e menor variabilidade de f_0 , em decorrência da atenção mais direcionada a aspectos segmentais que suprasegmentais. Concluiu-se que a produção do inglês-L2 por falantes de PB apresentou grande extensão de vocábulos funcionais, como artigos e preposições, e menor extensão de palavras de conteúdo, como substantivos e verbos nas sentenças, o que explica a menor variação silábica. Essa dificuldade de reduzir apropriadamente vocábulos funcionais desfavorece a implementação de maior grau de ritmo acentual por falantes de PB, característico do inglês.

Neste estudo, partimos da análise realizada por Silva Jr. e Barbosa (2019b), expandindo-a com a inclusão de dados do PB-L1 dos aprendizes, falantes da variedade dialetal de Fortaleza, e dados do inglês-L2 coletados em dois pontos distintos de desenvolvimento.

3 MÉTODO

Nesta seção, são apresentados o contexto, os participantes, os materiais e os procedimentos empregados neste estudo.

3.1 Contexto da pesquisa

Este estudo insere-se em um projeto de pesquisa mais amplo, denominado “Aquisição do ritmo do inglês-L2 por brasileiros”, que é coordenado pelo professor Dr. Ronaldo Manguera Lima Jr. (UFC) e financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico. Esse projeto tem por objetivo geral investigar o desenvolvimento do ritmo do inglês-L2 por alunos universitários brasileiros de Letras-Inglês que tenham estudado a L2 exclusivamente em salas de aula no Brasil. As análises de natureza acústica executadas no âmbito do projeto têm também o objetivo secundário de avaliar os modelos de descrição do ritmo das línguas e suas aplicações para interlínguas.

Assim, esta pesquisa baseou-se em dados de fala oriundos de um banco previamente constituído. Portanto, os procedimentos relacionados à escolha dos participantes, dos métodos de coleta e dos estímulos estão descritos a seguir com o objetivo de situar o leitor em relação à natureza dos dados analisados e não foram executados pelo presente autor. Ainda em relação aos estímulos, aspectos tais como distribuição de sílabas acentuadas e não acentuadas, distribuição segmental, complexidade silábica e choque acentual não foram controlados em nenhum dos estímulos, já que a seleção e construção das sentenças não possuía o objetivo explícito de analisar o ritmo, mas constituir um banco de dados que permitisse o estudo fonético de múltiplos fenômenos tanto no nível segmental quanto no nível prosódico, no contexto do desenvolvimento de inglês-L2. Já os procedimentos de segmentação, extração das métricas rítmicas e parâmetros acústicos, e tratamento estatístico foram conduzidos no âmbito da realização desta pesquisa, com vistas a atingir seus objetivos.

3.2 Participantes

Os participantes que compõem o grupo experimental perfazem um total de cinco aprendizes, estudantes universitários do curso de Letras-Inglês, sendo quatro homens e uma mulher, com idade entre 18 e 24 anos. Os critérios para participação na pesquisa incluíam: a) não possuir trancamento ou reprovação no curso; b) não ter viajado para país falante de língua

inglesa; e c) nunca ter tido contato contínuo com falante nativo de língua inglesa. Todos os alunos ingressantes foram convidados a contribuir com gravações semestrais a partir do início do segundo semestre de estudos, porém alguns aprendizes foram desistindo de participar das coletas, ou não se encaixavam mais nos critérios, de modo que houve apenas cinco participantes na última gravação, os quais integram este estudo. O grupo controle é composto por cinco falantes de inglês-L1, canadenses, com idades entre 23 e 34 anos, sendo um homem e quatro mulheres. A coleta de dados foi autorizada pelo Comitê de Ética em Pesquisa (CAAE: 40985414.1.0000.5054) e os participantes assinaram um Termo de Consentimento Livre Esclarecido.

3.3 Materiais

Integram este estudo quatro corpora de produção oral: inglês-L1, PB-L1, inglês-L2 (1) e inglês-L2 (4). Os dados do inglês-L2 foram obtidos por meio de gravações dos aprendizes brasileiros lendo o primeiro parágrafo do texto diagnóstico de Celce-Murcia, Brinton e Goodwin (2010) em dois momentos distintos, antes e depois de os alunos cursarem as disciplinas Fonologia Segmental da Língua Inglesa, no terceiro semestre da graduação; e Fonologia Suprasegmental da Língua Inglesa, no quarto semestre. Trata-se da primeira e da quarta gravação realizadas, por isso são chamadas aqui de inglês-L2 (1) e inglês-L2 (4).

Os dados do inglês-L1 foram obtidos por meio da leitura do mesmo texto pelos integrantes do grupo controle. Os dados do português-L1 são provenientes da leitura realizada pelos cinco aprendizes brasileiros de uma versão traduzida para o português do mesmo texto (Anexos A e B). Ambas as versões do texto contêm sete sentenças, incluindo orações simples e complexas, declarativas afirmativas, e interrogativas abertas e fechadas.

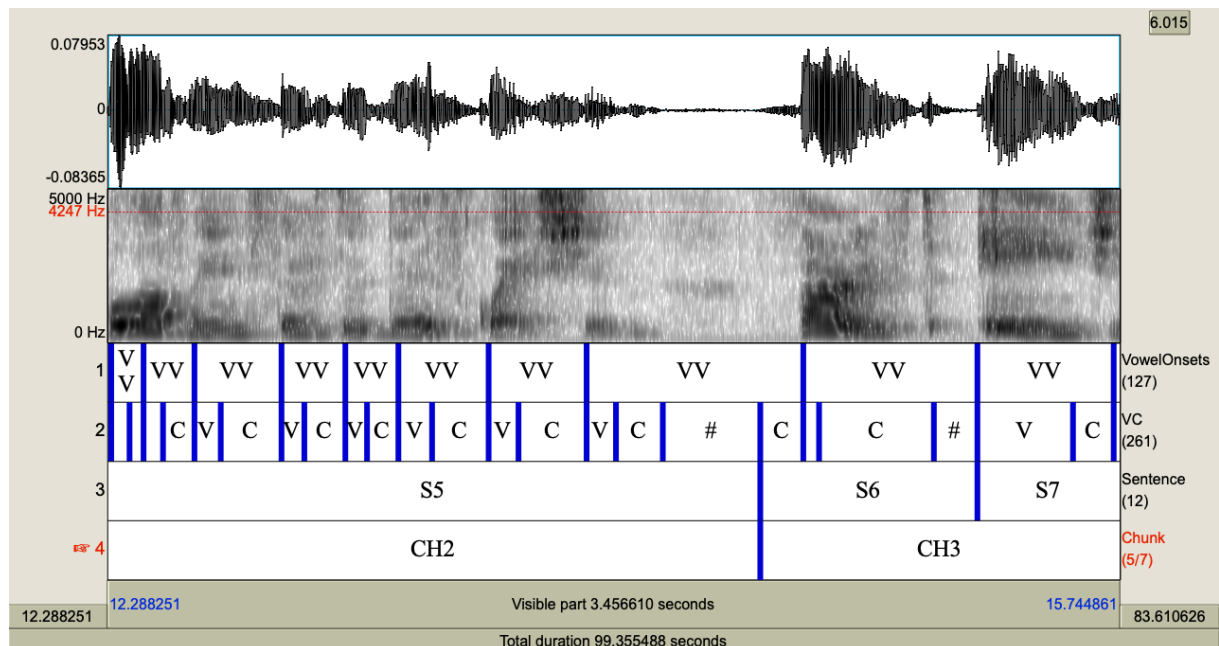
3.4 Procedimentos

Os participantes foram gravados individualmente em uma sala silenciosa, uma vez por semestre por quatro semestres consecutivos. As gravações foram realizadas com um microfone de lapela cardioide Shure MX150B, conectado a um gravador Zoom H4n. O áudio foi capturado em mono, com taxa de amostragem de 44.1kHz e posteriormente salvo em formato wav. Para a análise acústica, os dados foram segmentados manualmente no PRAAT (BOERSMA; WEENINK, 2020), em vogais (V), consoantes (C), sílabas fonéticas (VV),

sentenças (S), unidades sintático-prosódicas (*chunks*) e pausas (#) para a extração das métricas rítmicas e parâmetros acústicos.

Conforme pode ser observado na figura 3, a unidade VV corresponde ao intervalo entre o *onset* acústico de uma vogal e o *onset* da vogal adjacente, integrando fones de duas sílabas distintas (BARBOSA, 2007). Segundo Barbosa, a unidade VV “evidencia a cadência da posição da vogal, que caracteriza essencialmente o ritmo da fala” (BARBOSA, 2007, p. 81). As sentenças foram segmentadas segundo critérios sintático-prosódicos, com o objetivo de diferenciar fronteiras tonais fortes e fracas: frases verbais formadas por cópula; elementos na posição de *complementizer*; orações declarativas; orações interrogativas abertas e fechadas; orações complexas de transposição substantiva; e orações subordinadas. Já as unidades sintático-prosódicas superiores ou *chunks* encerram interpausas mais longas se comparadas às pausas existentes entre as sentenças, sendo o período a unidade de referência, com exceção das orações interrogativas com função retórica e suas respectivas respostas, que foram tratadas como um só *chunk* em ambas as línguas. A aplicação desse protocolo de segmentação resultou em 10 sentenças e 5 *chunks* por falante em cada corpus analisado.

Figura 3 – Forma de onda parcial, espectrograma de banda larga e quatro camadas, respectivamente, segmentadas e rotuladas como: 1) sílabas fonéticas (VV); 2) vogais e/ou consoantes (V/C); 3) sentenças (S); 4) *chunks* (CH) produzidas por um falante não nativo.



Fonte: elaborado pelo autor.

Para a extração das métricas e parâmetros acústicos, foi utilizado um script para PRAAT, o Metrics&AcousticsExtractor (SILVA JR.; BARBOSA, 2019a). Consoante Silva Jr.

e Barbosa (2019b), a categoria *métricas rítmicas* será utilizada neste trabalho para referir-se às métricas com base em duração, ao passo que o termo *parâmetros acústicos* diz respeito aos índices estatísticos que computam outros correlatos de proeminência do sinal acústico que não a duração, como f_0 , sonoridade e intensidade. O quadro 3 especifica as métricas e parâmetros acústicos que foram utilizados neste estudo e os segmentos sobre os quais incidem, conforme procedimentos metodológicos utilizados em Silva Jr. e Barbosa (2019b).

Quadro 3 – Métricas rítmicas e parâmetros utilizados neste estudo

MÉTRICAS		PARÂMETROS ACÚSTICOS	
Parâmetro	Segmento	Parâmetro	Segmento
Proporção (%)	V, C	Mediana de f_0	S, CH
Desvio-padrão (σ)	V,C (V ou C), VV	Pico de f_0	S, CH
Coefficiente de variação (Varco)	V,C (V ou C), VV	f_0 mínimo	S, CH
Índice bruto de variabilidade pareada (r-PVI)	V,C (V ou C), VV	Desvio-padrão de f_0 (σf_0)	S, CH
Índice normalizado de variabilidade pareada (n-PVI)	V,C (V ou C), VV	Assimetria (<i>skewness</i>) de f_0	S, CH
Taxa de Ritmo (RR)	V,C (V ou C), VV	Média da primeira derivada de f_0 P($\mu\Delta 1-f_0$)	S, CH
Índice de Variabilidade (VI)	V,C (V ou C), VV	Primeira derivada do desvio-padrão de f_0 . ($\sigma\Delta 1-f_0$)	S, CH
<i>Yet another rhythm determination (YARD)</i>	V,C (V ou C), VV	Primeira derivada da assimetria de f_0 ($sk\Delta 1-f_0$)	
		Taxa de Elocução (SR)	VV, S, CH
		Taxa de f_0 (f_0 -R)	S, CH
		Enfãse Espectral	S, CH
		Média normalizada da duração do pico silábico (μ_{dur} -Sil)	VV, S, CH
		Duração média de pausas (μ_{dur} -#)	S, CH

Fonte: Silva Jr. e Barbosa (2019b, p. 44)

Os dados obtidos por meio da extração das métricas e parâmetros acústicos foram tratados estatisticamente via R (R CORE TEAM, 2020) por meio da implementação de modelos de regressão de efeitos mistos para verificar as métricas e parâmetros acústicos que apresentam maior robustez e maior capacidade de predição e distinção entre inglês-L1, inglês-L2 e PB-L1, e entre os aprendizes brasileiros de inglês-L2 em diferentes estágios de desenvolvimento,

levando-se em consideração as variáveis preditoras (língua, semestre) e variáveis de resposta (métricas rítmicas e parâmetros acústicos).

4 RESULTADOS

Esta seção subdivide-se em duas subseções principais. Na subseção 4.1 são apresentados os resultados para as métricas rítmicas de base duracional e na subseção 4.2 são apresentados os resultados obtidos para os parâmetros acústicos.

A subseção 4.1 subdivide-se em outras sete subseções: métricas de proporção e desvio-padrão (4.1.1); coeficientes de variabilidade (*varcos*) (4.1.2); índices brutos e normalizados de variabilidade pareada - PVI (4.1.3); taxas de ritmo - RRs (4.1.4); índices de variabilidade - VIs (4.1.5); índices de variabilidade pareada com transformada *z* (*z-score*) - YARDs (4.1.6); e distâncias euclidianas (4.1.7).

A subseção 4.2, por sua vez, subdivide-se nas seguintes subseções: mediana de f_0 e média da primeira derivada de f_0 ($\mu\Delta 1-f_0$) (4.2.1); pico de f_0 e f_0 mínimo (4.2.2); desvio-padrão de f_0 (σf_0), assimetria (*skewness*) de f_0 e respectivas derivadas ($\sigma\Delta 1-f_0$ e $sk\Delta 1-f_0$) (4.2.3); taxa de f_0 (f_0 -R) (4.2.4); ênfase espectral (4.2.5); duração média dos picos silábicos (μdur - Sil) (4.2.6); e taxa de elocução (SR) (4.2.7).

4.1 Métricas rítmicas

As métricas rítmicas de base duracional tomam por base o correlato de proeminência da duração como referência para computar os padrões rítmicos das línguas naturais. Esses índices podem ser globais ou locais, a depender de como as diferenças entre os intervalos de referência são calculadas, se em relação a todos os intervalos de mesmo tipo do enunciado ou em relação ao intervalo adjacente. Também podem ser brutos ou normalizados para a taxa de elocução.

Neste trabalho, sete métricas duracionais foram analisadas: proporção (%), desvio-padrão (Δ), coeficiente de variabilidade (*variability coefficient* - *varco*), índice de variabilidade pareada (*pairwise variability index* - PVI), taxa de ritmo (*rhythm rate* - RR), índice de variabilidade (*variability index* - VI), e YARD (*yet another rhythm determination*) (cf. Quadro 1). Essas métricas foram calculadas a partir dos seguintes intervalos de referência: vogais (V), consoantes (C), vogais e consoantes consecutivas (VC) e sílabas fonéticas (VV), de modo que a combinação de cada métrica com um intervalo de referência gerou uma variável resposta. As métricas de proporção incidiram somente sobre vogais e consoantes. Foram ajustados modelos de regressão com efeitos mistos para cada uma das métricas a fim de verificar a significância dos resultados e os efeitos das (inter)línguas sobre as elas.

4.1.1 Métricas de proporção (%) e desvio-padrão (Δ)

Nesta subseção apresentamos os resultados para as métricas que computam proporção e desvio-padrão de intervalos de referência. Conforme assevera Fuchs (2016), métricas como %V e %C (cf. Quadro 1) representam a contribuição substancial do tipo de intervalo considerado na composição do material fonético analisado. Espera-se que uma maior diversidade de arranjos silábicos, permitindo agrupamentos consonantais mais longos, seja implementadora de um ritmo com maior tendência acentual, com %V mais baixo e %C mais alto. O contrário é esperado de línguas que sejam majoritariamente realizadas com sílabas do tipo CV. Já as métricas de desvio-padrão computam globalmente o grau de variação ou dispersão na duração dos intervalos de referência. A tabela 1 apresenta as médias das métricas %V, %C, ΔV , ΔC , ΔVC e ΔS e respectivos desvios-padrão (entre parênteses) para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4). As médias obtidas foram multiplicadas por 100 para facilitar o comparativo entre os resultados, exceto %V e %C por já possuírem essa característica intrínseca.

Tabela 1 – Médias absolutas de %V, %C, ΔV , ΔC , ΔVC e ΔS e desvio-padrão (entre parênteses) para PB, inglês-L1, inglês-L2 (1) e inglês-L2

Língua	%V	%C	ΔV	ΔC	ΔVC	ΔS
BP	48,56 (3,16)	51,44 (3,16)	40.08 (10,81)	46.48 (8,02)	105.5 (103,47)	133.4 (45,27)
Eng-L1	38,88 (4,96)	61,12 (4,96)	41.16 (11,82)	68.416 (14,55)	125.59 (33,88)	198.53 (77,44)
Eng-L2 (1)	36,24 (5,36)	63,76 (5,36)	51.81 (12,79)	105.192 (32,6)	136.16 (61,67)	217.46 (97,65)
Eng-L2 (4)	39,28 (4,72)	60,72 (4,72)	52.92 (19,73)	84.088 (36,51)	107.88 (29,45)	184.75 (56,72)

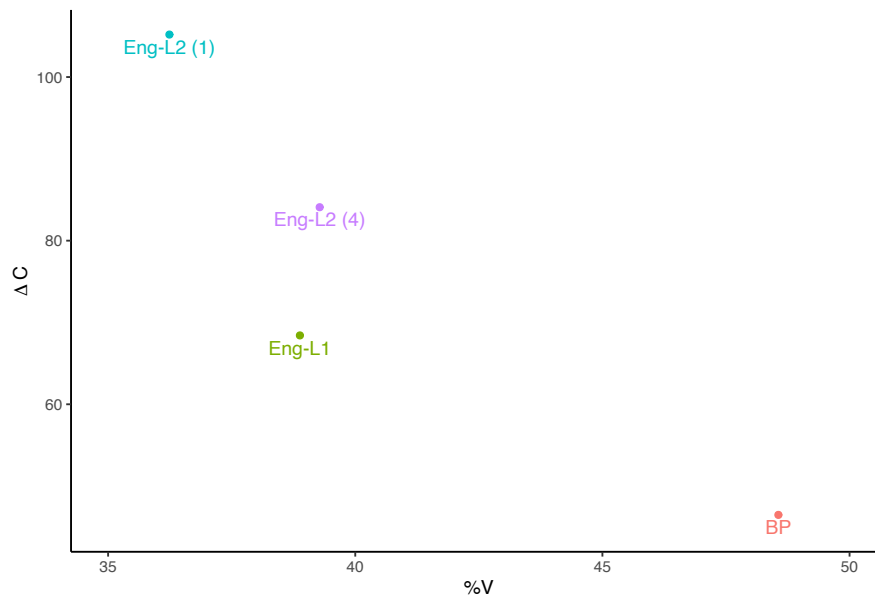
Fonte: elaborada pelo autor.

Conforme pode ser observado na tabela 1, o PB registrou os menores valores para todas as métricas, à exceção de %V, evidenciando uma tendência ao ritmo mais silábico. O contrário ocorreu com os dados do inglês-L2(1) que, à exceção de %V e ΔV , registraram os valores mais extremos para os índices métricos analisados. Já os dados do inglês-L2(4) aproximaram-se mais do inglês-L1, cujos valores figuram como intermediários entre inglês-L2(4) e PB para a maior parte das métricas de proporção e desvio-padrão. A seguir, parte desses resultados serão detalhados tomando por base o estudo seminal de Ramus, Nespor e Mehler (1999), bem como Arvaniti (2012). Para tanto, ΔC e %V são apresentadas de maneira bidimensional. Os dados de %C também são reportados, complementarmente aos dados de %V,

embora a literatura em geral não o faça, considerando que a proporção da duração total do enunciado composto de intervalos consonantais (%C) é meramente o complemento da duração total do enunciado composto de intervalos vocálicos (%V). Em seguida, os dados de ΔV são apresentados e plotados bidimensionalmente com ΔC e %V. Finalmente, os resultados obtidos para o desvio-padrão da duração dos intervalos vocálicos e consonantais sucessivos (ΔVC) e o desvio-padrão da duração das sílabas fonéticas (ΔS) são apresentados.

O gráfico a seguir (figura 4) apresenta a distribuição dos quatro corpora deste estudo em um plano cartesiano com ΔC e %V nos eixos, como realizado por Ramus, Nespore e Mehler (1999):

Figura 4 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔC e %V



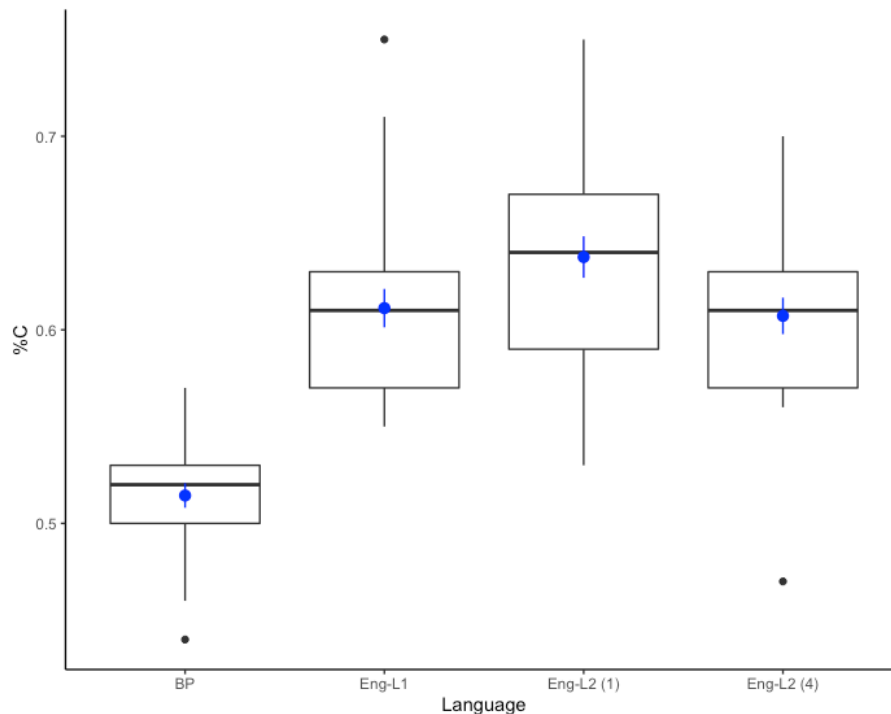
Fonte: Teixeira e Lima Jr. (2021)

No plano rítmico criado pelos eixos ΔC e %V, inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB ocuparam posições claramente distintas. Em relação aos dados de L1, inglês-L1 e PB ocuparam posições opostas em relação aos eixos ΔC e %V. O inglês-L1 apresentou maior desvio-padrão da duração de intervalos consonantais ($\Delta C_{\text{inglês-L1}} = 68,41$) em comparação ao PB ($\Delta C_{\text{PB}} = 46,48$); e o PB apresentou maior proporção do enunciado composto de intervalos vocálicos ($\%V_{\text{PB}} = 48,56$) em comparação ao inglês-L1 ($\%V_{\text{inglês-L1}} = 38,88$). Os dados de inglês-L2 se posicionaram distantes das duas línguas nativas, com o inglês-L2 (1) ocupando uma posição no plano rítmico diametralmente oposta ao PB e distante também do inglês-L1, registrando valores de ΔC bastante elevados ($\Delta C_{\text{inglês-L2(1)}} = 105,19$) e a menor proporção de

segmentos vocálicos ($\%V_{\text{inglês-L2(1)}} = 36,24$) dentre os quatro corpora analisados. O inglês-L2 (4) aproximou-se do inglês-L1 em relação ao eixo ΔC ($\Delta C_{\text{inglês-L2(4)}} = 84,08$) e registrou valores bem próximos ao inglês-L1 para o eixo $\%V$ ($\%V_{\text{inglês-L2(4)}} = 39,28$), indicando um percurso de desenvolvimento dos aprendizes rumo a padrões prosódicos mais próximos do inglês-L1.

Já a distribuição dos resultados de proporção do enunciado composto de intervalos consonantais ($\%C$) pode ser visualizada na figura 5. Naturalmente, o PB apresentou o comportamento oposto a $\%V$, com menor proporção do enunciado composto de intervalos consonantais ($\%C_{\text{PB}} = 51,44$) em comparação ao inglês-L1 ($\%C_{\text{inglês-L1}} = 61,12$). O inglês-L2(1) superou os valores para o inglês-L1 ($\%C_{\text{inglês-L2(1)}} = 63,76$) em comparação aos dados do inglês-L2(4) ($\%C_{\text{inglês-L2(4)}} = 60,72$).

Figura 5 – *Boxplot* de $\%C$ para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborado pelo autor.

Para as métricas de $\%V$ e ΔC , dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada métrica, e ambos revelaram efeitos significativos de cada uma das (inter)línguas sobre as métricas, conforme mostra o resumo dos modelos da tabela 2.

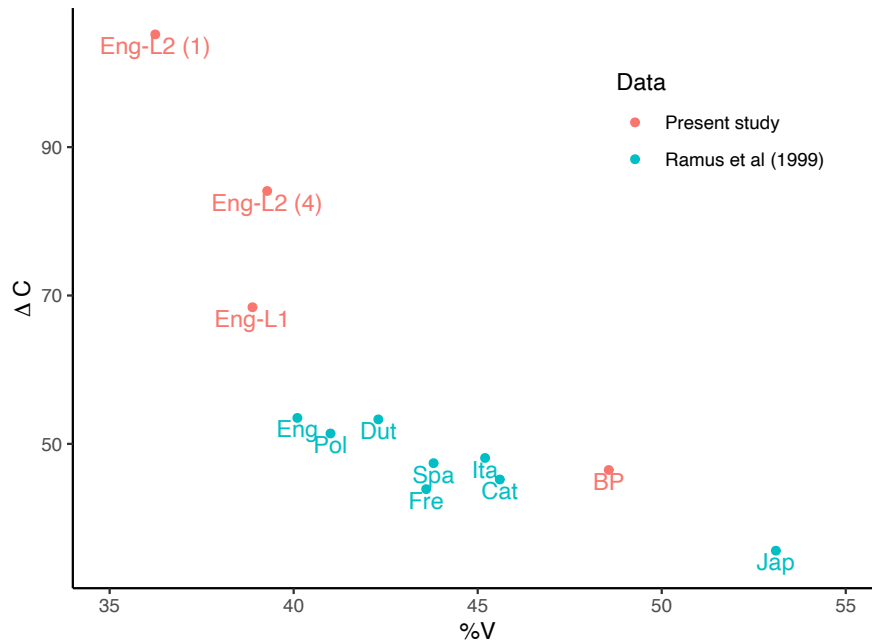
Tabela 2 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para ΔC e %V. Modelos: $\Delta C \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\text{percV} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	ΔC			%V		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	46.48	36.17 – 56.79	<0.001	48.78	46.02 – 51.55	<0.001
Lang [Eng-L1]	21.94	7.35 – 36.52	0.004	-9.66	-12.97 – -6.35	<0.001
Lang [Eng-L2 (1)]	58.71	44.13 – 73.30	<0.001	-11.92	-14.23 – -9.61	<0.001
Lang [Eng-L2 (4)]	37.61	23.02 – 52.19	<0.001	-9.28	-11.47 – -7.09	<0.001

Fonte: Teixeira e Lima Jr. (2021)

Comparando os dados deste estudo aos do estudo seminal de Ramus, Nespore e Mehler (1999), obtemos a distribuição da figura 6:

Figura 6 – Dados do presente estudo e dos de Ramus, Nespore e Mehler (1999) para ΔC e %V, em que Eng = inglês, Pol = polonês, Dut = holandês, Spa = espanhol, Fre = francês, Ita = italiano, Cat = catalão, Jap = japonês.



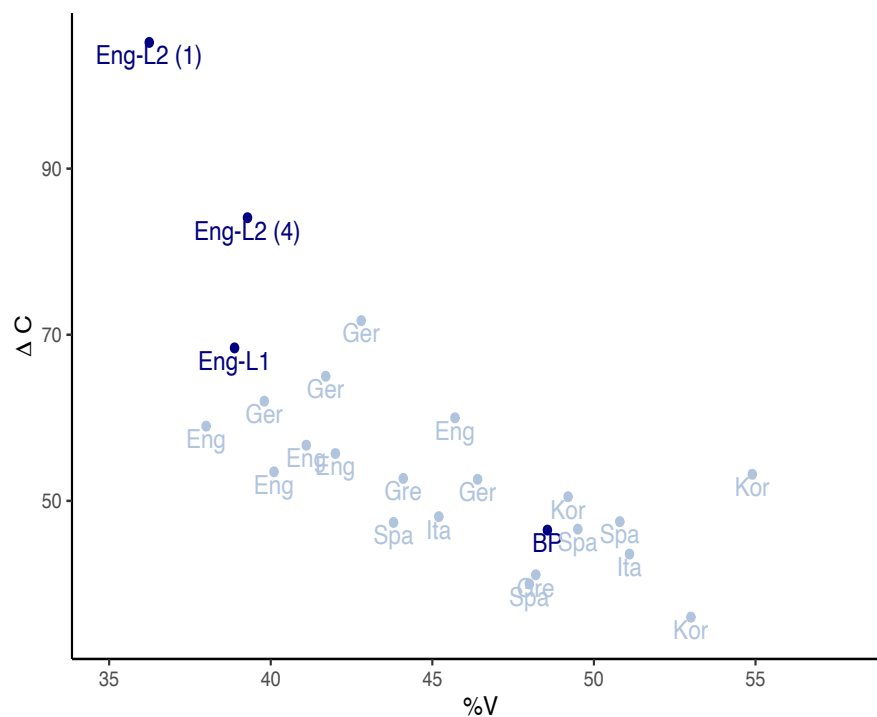
Fonte: Teixeira e Lima Jr. (2021)

Os dados de PB aproximaram-se dos de línguas de ritmo considerado prototipicamente mais silábico (espanhol, italiano, francês e catalão), com valores muito próximos sobretudo no eixo de ΔC . Em relação ao eixo %V, o PB registrou a segunda maior proporção vocálica dentre todas as línguas consideradas no comparativo, apontando para uma maior tendência ao ritmo silábico. Percebe-se, contudo, um distanciamento do PB em relação

ao japonês, de ritmo supostamente moraico. Semelhantemente, os dados de inglês-L1 se aproximaram dos das línguas de ritmo prototipicamente acentual (inglês, polonês e holandês), principalmente no eixo de %V. As diferenças verificadas no eixo ΔC podem estar relacionadas às variedades dialetais utilizadas neste estudo, inglês-norte americano canadense, e em Ramus, Nespore e Mehler (1999), não especificada; ou a diferentes taxas de elocução, conforme assevera Dellwo (2006). As interlínguas se posicionaram distantes das demais línguas, em direção ao ritmo acentual, mas com os dados do inglês-L2 (4) se aproximando mais dos dados de inglês-L1 deste estudo e dos de Ramus, Nespore e Mehler (1999), principalmente em relação ao eixo %V.

Como forma de mais uma comparação, apresentamos (figura 7) os nossos dados em meio aos dados resenhados¹⁴ e obtidos por Arvaniti (2012), cujo estudo abrangente incluiu um total de seis línguas, sendo duas consideradas prototipicamente mais acentuais (inglês e alemão), duas consideradas de ritmo mais silábico (espanhol e italiano), e duas consideradas até então de difícil classificação (grego e coreano).

Figura 7 – Dados do presente estudo (azul escuro) em meio a todos os dados resenhados e obtidos por Arvaniti (2012) (azul claro) para ΔC e %V, em que Eng = inglês, Ger = alemão, Gre = grego, Spa = espanhol, Ita = italiano, Kor = coreano.



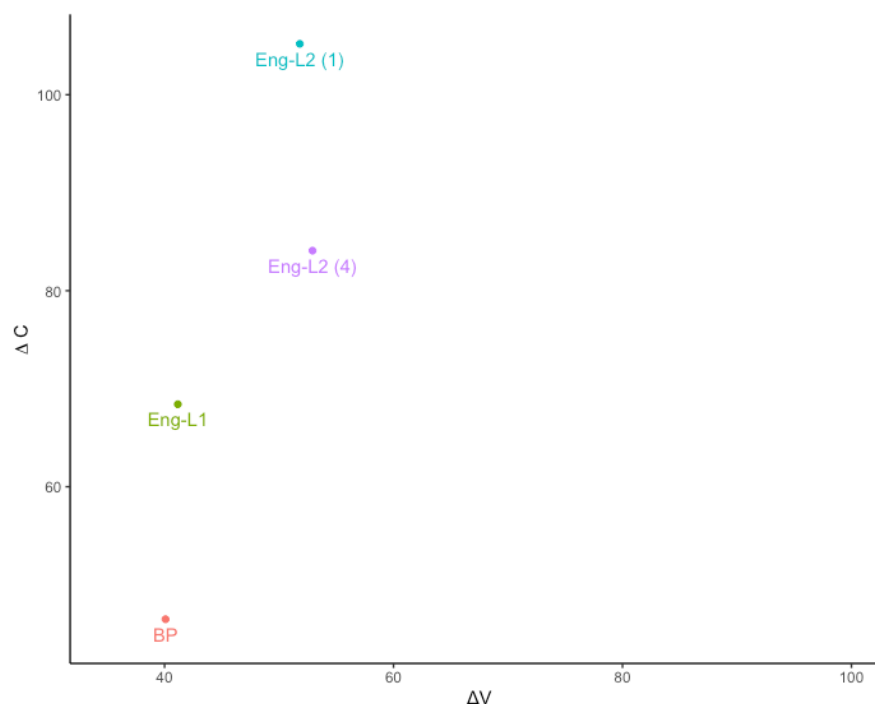
Fonte: Teixeira e Lima Jr. (2021)

¹⁴ Especificamente, os dados de Baltazani (2007); Dellwo e Wagner (2003); Grabe e Low (2002); Jeon (2006); Mok e Lee (2008); Ramus, Nespore e Mehler (1999); Russo e Barry (2008); Tisiartsoni (2003); e White e Mattys (2007).

A análise visual desses agrupamentos no plano ΔC - $\%V$ permite confirmar mais uma vez a peculiaridade temporal do inglês-L2 (1), que se afasta diametralmente do PB e registra os valores mais extremos para ambos os eixos dentre todas as línguas reportadas. Já os dados do inglês-L2 (4) aproximam-se dos dados de línguas consideradas mais acentuais, ocupando uma posição mais consistente e intermediária para o eixo $\%V$ em relação às médias registradas para o inglês e alemão nos estudos de referência. Os dados do inglês-L1 reportados neste estudo parecem consistentes em relação à literatura em ambos os eixos, embora com maior variação no eixo ΔC . Os dados do PB agrupam-se aos de línguas consideradas de ritmo mais silábico, aproximando-se do espanhol e do italiano no eixo ΔC , e do espanhol, do grego e do coreano no plano $\%V$.

Em relação ao desvio-padrão da duração dos intervalos vocálicos (ΔV), os resultados foram menos elucidativos e agruparam as L1s mais próximas uma da outra ($\Delta V_{PB} = 40,08$; $\Delta V_{\text{inglês-L1}} = 41,16$), assim como as interlínguas ($\Delta V_{\text{inglês-L2(1)}} = 51,81$; $\Delta V_{\text{inglês-L2(4)}} = 52,92$). Consoante Ramus, Nespor e Mehler (1999), os resultados obtidos para ΔV também foram plotados bidimensionalmente com ΔC e com $\%V$, como pode ser observado nos gráficos das figuras 8 e 9, respectivamente.

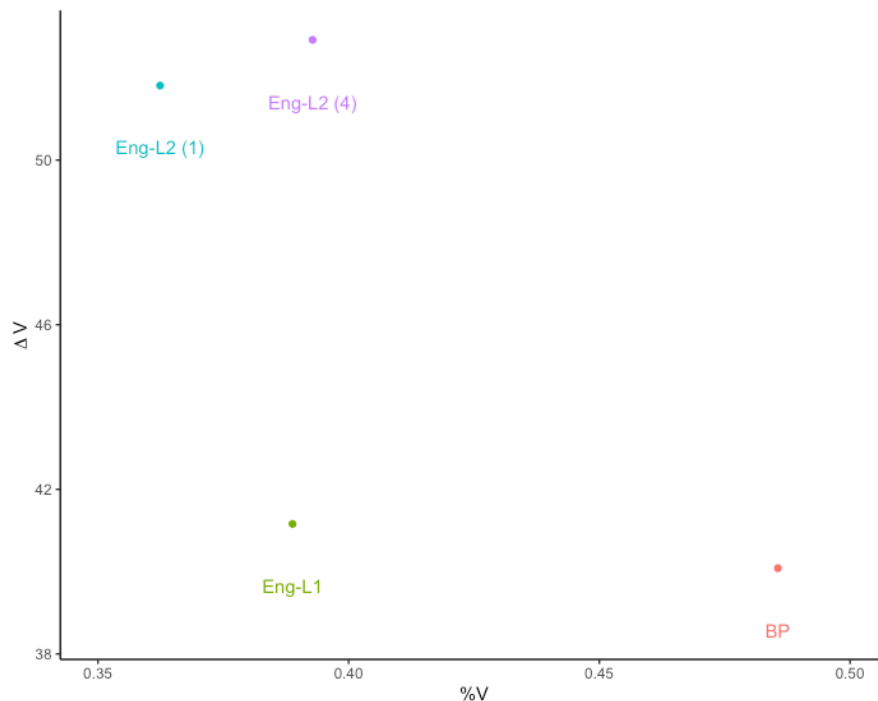
Figura 8 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔC e ΔV



Fonte: elaborado pelo autor.

Os espaços rítmicos gerados pela combinação de ΔC com ΔV , e ΔV com $\%V$ não explicitaram diferenças rítmicas significativas entre as L1s, tampouco flagraram um percurso nítido de desenvolvimento, uma vez que os dados de interlíngua permaneceram praticamente inalterados e muito próximos um do outro, com o inglês-L2(4) ligeiramente mais afastado tanto do PB quanto do inglês-L1.

Figura 9 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos ΔV e $\%V$



Fonte: elaborado pelo autor.

De acordo com Ramus, Nespore e Mehler (1999), a dificuldade em interpretar de forma transparente os resultados para ΔV pode ser atribuída a uma série de variáveis fonológicas que são captados por esse índice e o influencia: redução vocálica, a presença de segmentos vocálicos longos e alongamento contrastivo. É interessante notar que tanto o inglês quanto o PB compartilham as duas primeiras características, embora realizadas de formas distintas, o que pode ter contribuído para a aproximação dessas línguas no eixo ΔV .

Um modelo de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foi ajustado para ΔV e revelou efeitos significativos apenas das L2s em relação ao PB, conforme mostra o resumo dos modelos da tabela 3.

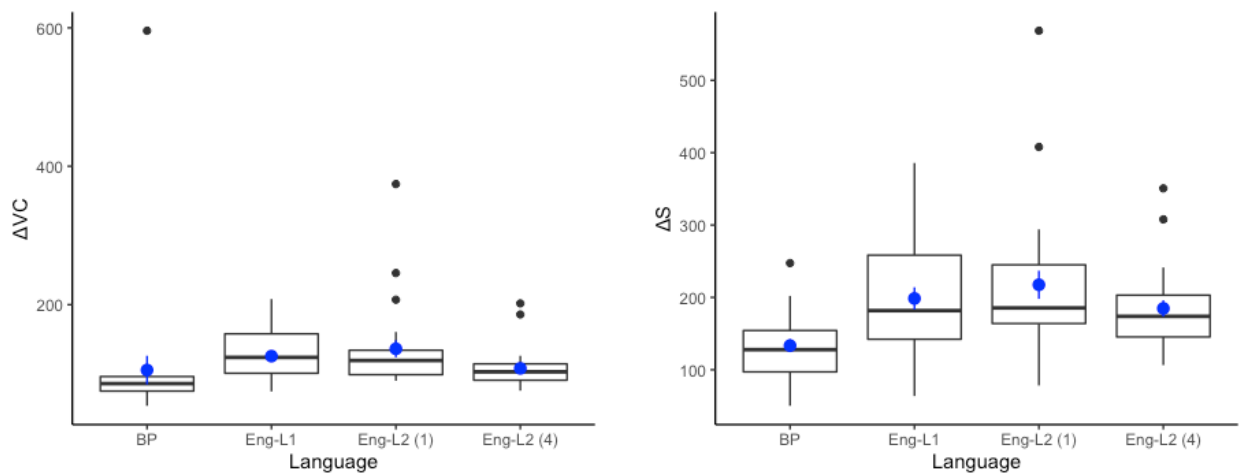
Tabela 3 – Coeficientes, intervalos de confiança (95%) e valores de p um modelo de regressão linear com efeitos mistos ajustados para ΔV . Modelo: $\Delta V \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

Predictors	ΔV		
	Estimates	CI	p
(Intercept)	39.41	31.39 – 47.43	<0.001
Lang [Eng-L1]	1.75	-8.17 – 11.66	0.730
Lang [Eng-L2 (1)]	12.85	5.59 – 20.11	0.001
Lang [Eng-L2 (4)]	12.84	5.91 – 19.77	<0.001

Fonte: elaborada pelo autor.

Finalmente, tanto o desvio-padrão da duração dos intervalos vocálicos e consonantais consecutivos (ΔVC) quanto o desvio-padrão da duração sílabas fonéticas (ΔS) registram uma distribuição de valores com uma hierarquia similar ao padrão de variabilidade detectado para ΔC , como pode ser observado nos *boxplots* da figura 10.

Figura 10 – *Boxplots* de ΔVC (à direita, 10.1) e ΔS (à esquerda, 10.2) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

O inglês-L1 apresentou maior média de desvio-padrão tanto da duração dos intervalos vocálicos e consonantais consecutivos ($\Delta VC_{\text{inglês-L1}} = 125,59$) quanto da duração das sílabas fonéticas ($\Delta S_{\text{inglês-L1}} = 198,53$) em comparação ao PB ($\Delta VC_{\text{PB}} = 105,05$; $\Delta S_{\text{PB}} = 133,40$). Já em relação aos dados das interlínguas, o inglês-L2 (1) registrou valores de ΔVC e ΔS distantes do PB e mais elevados que o inglês-L1 ($\Delta VC_{\text{inglês-L2(1)}} = 136,16$; $\Delta S_{\text{inglês-L2(1)}} = 217,46$). Já o inglês-L2 (4) comportou-se de maneira distinta para cada uma dessas métricas,

com valores intermediários entre o PB e o inglês-L1 para ambas, mas aproximando-se mais do PB para ΔVC ($\Delta VC_{\text{inglês-L2(4)}} = 107,88$) e do inglês-L1 para ΔS ($\Delta S_{\text{inglês-L2(4)}} = 184,75$). Neste sentido, ΔS foi mais consistente na comparação com o percurso de desenvolvimento flagrado por ΔC , que indica o desenvolvimento dos aprendizes rumo a padrões prosódicos do inglês-L1 a partir de valores que se distanciam também do PB.

Como meio de atestar a significância estatística da robustez de ΔS em relação a ΔVC , dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada métrica. Como pode ser observado na tabela 4, o modelo ajustado para ΔS revelou efeitos significativos de cada uma das (inter)línguas sobre essa métrica, o que não foi verificado para ΔVC .

Tabela 4 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para ΔVC e ΔS . Modelos: $\text{deltaVC} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\text{deltaS} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	ΔVC			ΔS		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	105.19	79.10 – 131.29	<0.001	133.41	104.78 – 162.04	<0.001
Lang [Eng-L1]	20.40	-16.51 – 57.31	0.279	65.13	24.64 – 105.61	0.002
Lang [Eng-L2 (1)]	30.99	-4.66 – 66.64	0.088	84.05	43.57 – 124.54	<0.001
Lang [Eng-L2 (4)]	2.38	-32.98 – 37.74	0.895	51.35	10.86 – 91.83	0.013

Fonte: elaborada pelo autor.

4.1.2 Varcos

Nesta subseção são apresentados os resultados para as métricas que computam o coeficiente de variação da duração dos intervalos de referência. Varcos calculam globalmente o desvio-padrão da duração dos intervalos de referência, normalizando-os por meio da divisão pela média e multiplicando-os por 100, como meio de atenuar os efeitos da taxa de elocução. Assim como nas métricas de desvio-padrão não normalizadas (ΔC e ΔV). A tabela 5 apresenta as médias das métricas VarcoV, VarcoC, VarcoVC e VarcoS para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) e respectivos desvios-padrão. As médias obtidas foram multiplicadas por 100 para facilitar o comparativo entre os resultados.

Tabela 5 – Médias absolutas de VarcoV, VarcoC, VarcoVC e VarcoS e desvio-padrão (entre parênteses) para PB, inglês-L1, inglês-L2 (1) e inglês-L2

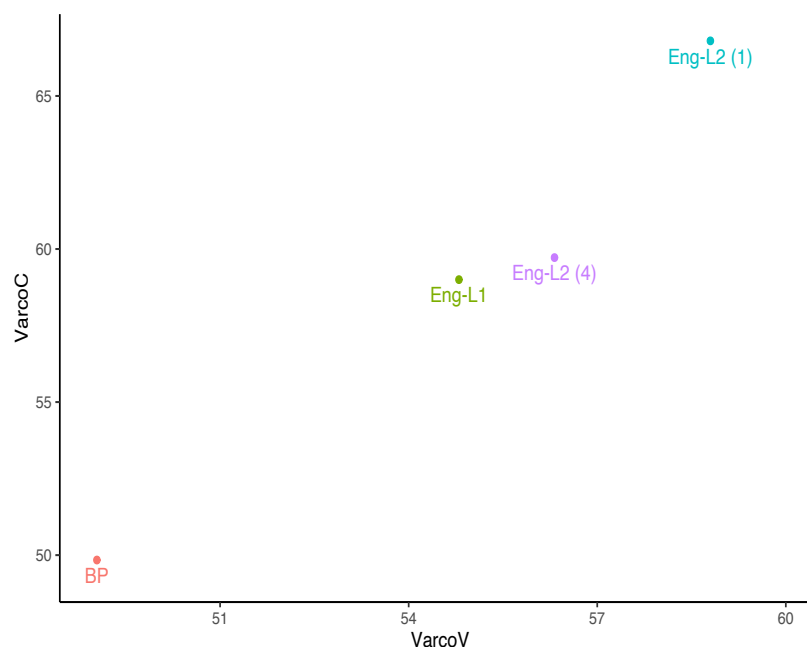
Língua	VarcoV	VarcoC	VarcoVC	VarcoS
BP	49,04 (8,49)	49,84 (7,00)	96,92 (55,93)	67,92 (19,75)
Eng-L1	54,80 (12,52)	59 (11,89)	109,08 (28,98)	82,52 (21,95)
Eng-L2 (1)	58,80 (11,30)	66,80 (13,69)	94,12 (19,53)	73,28 (16,99)
Eng-L2 (4)	56,32 (14,81)	59,72 (22,02)	82,92 (17,27)	67,04 (12,62)

Fonte: elaborada pelo autor.

Primeiramente, os resultados obtidos para neste estudo para o coeficiente de variação da duração dos intervalos consonantais (VarcoC) e o coeficiente de variação da duração dos intervalos vocálicos (VarcoV) são apresentados e comparados aos dados obtidos por Arvaniti (2012). Para tanto, VarcoC e VarcoV são apresentadas de maneira bidimensional. Em seguida, os resultados obtidos para o coeficiente de variação da duração dos intervalos vocálicos e consonantais consecutivos (VarcoVC) e o coeficiente de variação da duração das sílabas fonéticas (VarcoS) são reportados.

A figura 11 apresenta a distribuição dos corpora deste estudo para VarcoC, coeficiente de variação da duração de intervalos consonantais; e VarcoV, coeficiente de variação da duração de intervalos vocálicos.

Figura 11 – Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB no plano formado pelos eixos VarcoC e VarcoV



Fonte: Teixeira e Lima Jr. (2021)

Note-se, primeiramente, que no plano com ΔC e %V nos eixos y e x, respectivamente, as línguas se posicionam do canto superior esquerdo representando as mais acentuais para o canto inferior direito representando as mais silábicas. Num plano com VarcoC e VarcoV nos mesmos eixos, as línguas se posicionam numa diagonal inversa, com as línguas mais acentuais em direção ao canto superior direito, onde há maior variabilidade nas durações tanto de vogais como de consoantes; e as línguas mais silábicas em direção ao canto inferior esquerdo, onde há menor variabilidade (maior estabilidade) nas durações dos intervalos consonantais e vocálicos.

Sendo assim, observamos que, no plano rítmico formado pelos eixos VarcoC e VarcoV, inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB distribuíram-se de maneira análoga ao plano ΔC e %V, com os dados do PB registrando menor índice de variabilidade tanto em relação ao eixo VarcoC ($\text{VarcoC}_{\text{PB}} = 49,84$) como ao de VarcoV ($\text{VarcoV}_{\text{PB}} = 49,04$). Os dados de interlíngua e inglês-L1 também apresentaram consistência em relação aos resultados obtidos para ΔC e %V. O inglês-L2 (1) apresentou as maiores pontuações tanto no eixo VarcoC ($\text{VarcoC}_{\text{inglês-L2(1)}} = 66,8$) quanto no eixo VarcoV ($\text{VarcoV}_{\text{inglês-L2(1)}} = 58,8$). Já o inglês-L2 (4) registrou valores mais baixos para VarcoC ($\text{VarcoC}_{\text{inglês-L2(4)}} = 59,72$) e VarcoV ($\text{VarcoV}_{\text{inglês-L2(4)}} = 56,32$), muito próximos aos valores do inglês-L1, que apresentou as segundas menores médias para VarcoC ($\text{VarcoC}_{\text{inglês-L1}} = 59$) e VarcoV ($\text{VarcoV}_{\text{inglês-L1}} = 54,8$). Apesar de a hierarquia de valores para VarcoV e VarcoC ter sido análoga ao plano ΔC e %V ([+ acentual] inglês-L2 (1) > inglês-L2 (4) > inglês-L1 > PB [+ silábico]), observamos que o inglês-L2 (4) ficou posicionado ainda mais próximo ao inglês-L1 com essas métricas.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada métrica.

Tabela 6 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VarcoC e VarcoV. Modelos: $\text{varcoC} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\text{varcoV} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

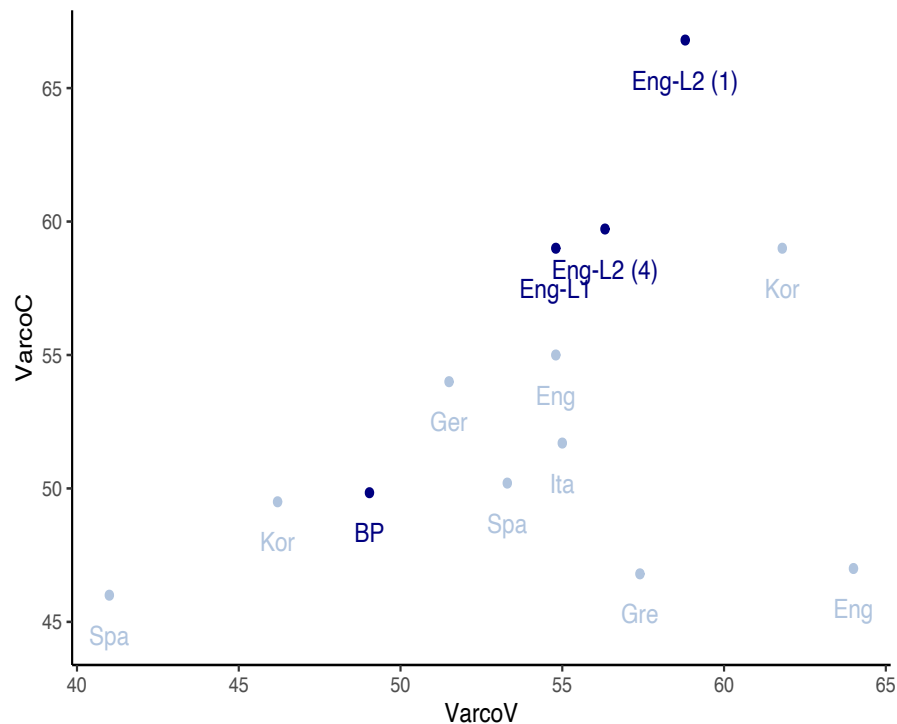
<i>Predictors</i>	varcoC			varcoV		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	49.84	44.01 – 55.67	<0.001	48.20	41.98 – 54.41	<0.001
Lang [Eng-L1]	9.16	0.91 – 17.41	0.030	6.60	-2.20 – 15.40	0.141
Lang [Eng-L2 (1)]	16.96	8.71 – 25.21	<0.001	10.83	4.52 – 17.13	0.001
Lang [Eng-L2 (4)]	9.88	1.63 – 18.13	0.019	7.28	1.27 – 13.29	0.018

Fonte: Teixeira e Lima Jr. (2021)

Ambos revelaram efeitos significativos das (inter)línguas sobre as métricas VarcoC e VarcoV, com exceção de inglês-L1 em comparação ao PB para VarcoV, conforme mostra o resumo dos modelos da tabela 6.

Para efeito de comparação, apresentamos, na figura 12, os dados obtidos neste estudo em meio aos resultados resenhados e obtidos por Arvaniti (2012) para VarcoC e VarcoV:

Figura 12 – Dados do presente estudo em meio a todos os dados resenhados e obtidos por Arvaniti (2012) para VarcoC e VarcoV, em que Eng = inglês, Ger = alemão, Gre = grego, Spa = espanhol, Ita = italiano, Kor = coreano.



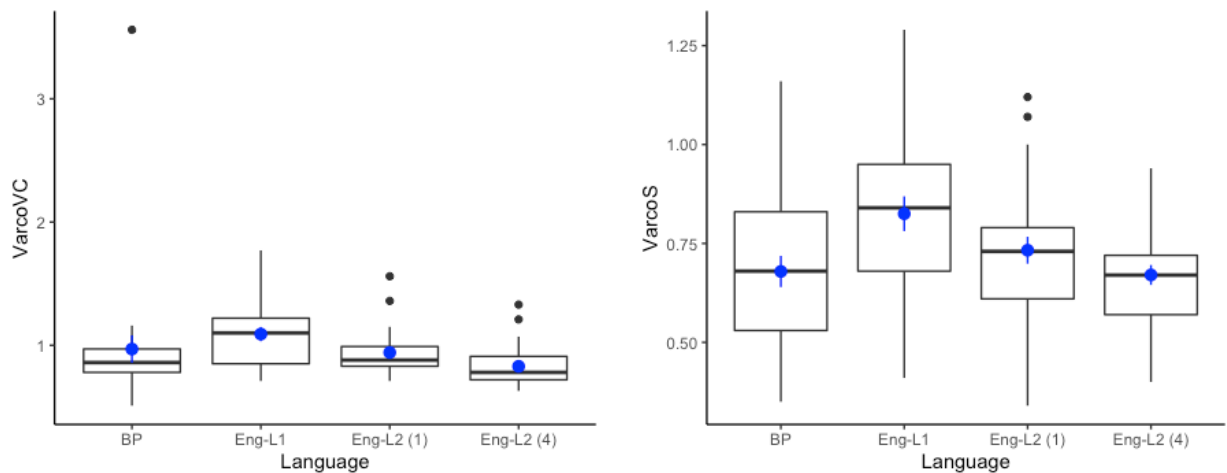
Fonte: Teixeira e Lima Jr. (2021)

Os dados de PB aproximaram-se dos de línguas consideradas prototipicamente mais silábicas (espanhol e italiano). Em relação ao inglês-L1, a média obtida para VarcoC ficou um pouco acima da média dos dados do inglês obtidos por Arvaniti (2012), mas o valor de VarcoV foi exatamente o mesmo nos dados deste estudo e nos dados obtidos por Arvaniti. Há uma ocorrência de inglês com valores bastante baixos para VarcoC, obtidos por White e Mattys (2007). As interlínguas se posicionaram em locais mais extremos de ambos os eixos, apontando para maior variabilidade duracional, mas com os dados da quarta gravação se aproximando bastante dos dados de inglês-L1.

O padrão de desenvolvimento flagrado pelas métricas VarcoC e VarcoV neste estudo não foi detectado na análise do coeficiente de variação da duração dos intervalos

vocálicos e consonantais consecutivos (VarcoVC) e nem do coeficiente de variação da duração das sílabas fonéticas (VarcoS), como visualmente explicitado na figura 13:

Figura 13 – *Boxplots* das médias de VarcoVC (à esquerda, 13.1) e VarcoS (à direita, 13.2) para Inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

O padrão dos resultados flagrados por VarcoVC e VarcoS é menos elucidativo em relação ao percurso de desenvolvimentos dos aprendizes, embora as L1s registrem valores consistentes com a expectativa de maior variabilidade da duração dos intervalos de referência. Para VarcoVC, o inglês-L1 apresentou maior média ($\text{VarcoVC}_{\text{inglês-L1}} = 106,08$) em comparação ao PB ($\text{VarcoVC}_{\text{PB}} = 96,92$). Valores intermediários foram registrados para o inglês-L2(1) ($\text{VarcoVC}_{\text{inglês-L2(1)}} = 94,12$) e inglês-L2(4) ($\text{VarcoVC}_{\text{inglês-L2(4)}} = 82,92$). Já em relação aos valores registrados para VarcoS, o inglês-L1 também apresentou maior variabilidade ($\text{VarcoS}_{\text{inglês-L1}} = 82,52$), seguido do inglês-L2(1) ($\text{VarcoS}_{\text{inglês-L2(1)}} = 73,28$). PB e inglês-L2(4) registraram valores praticamente idênticos (67,92 e 67,94 respectivamente).

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*), que foram ajustados um para cada métrica, também não revelaram efeitos significativos das (inter)línguas sobre essas métricas, com exceção do inglês-L1 em relação ao PB para VarcoS, conforme mostra o resumo dos modelos da tabela 7.

Tabela 7 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VarcoVC e VarcoS. Modelos: $\text{varcoVC} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\text{varcoS} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	varcoVC			varcoS		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.96	0.81 – 1.12	<0.001	0.68	0.61 – 0.75	<0.001
Lang [Eng-L1]	0.13	-0.09 – 0.34	0.251	0.15	0.04 – 0.25	0.005
Lang [Eng-L2 (1)]	-0.02	-0.21 – 0.17	0.840	0.05	-0.05 – 0.16	0.300
Lang [Eng-L2 (4)]	-0.14	-0.32 – 0.04	0.135	-0.01	-0.11 – 0.09	0.864

Fonte: elaborada pelo autor.

4.1.3 PVI_s

Nesta subseção são apresentados os resultados para as métricas que computam o índice de variabilidade pareada (PVI), isto é, a média das diferenças entre os intervalos de referência adjacentes. Estes índices são considerados locais e podem ser brutos (rPVI_s), ou normalizados (nPVI_s), por meio da divisão da média das diferenças pelo somatório da diferença dos intervalos pareados, com posterior multiplicação por 100.

O procedimento de normalização tem como objetivo atenuar ou minimizar os efeitos da taxa de elocução e é normalmente aplicado aos intervalos vocálicos (nPVI_s-V). No entanto, pelo caráter abrangente deste estudo, ambas as versões de PVI foram quantificadas para vogais, consoantes, unidades vogal-consoante consecutivas e sílabas fonéticas (unidades VV). Quanto mais alto o PVI, maior é a diferença entre a duração das unidades de referência consecutivas, sinalizando maior tendência à implementação do ritmo acentual. Em contrapartida, um PVI mais baixo indica que há menos diferença do ponto de vista duracional entre os intervalos de referência adjacentes, o que favorece a implementação de um padrão rítmico com maior tendência silábica.

A tabela 8 apresenta as médias das métricas de PVI brutas (rPVI_s-V, rPVI_s-C, rPVI_s-V e rPVI_s-S) e normalizadas (nPVI_s-V, nPVI_s-C, nPVI_s-VC, nPVI_s-S) para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4). As médias obtidas foram multiplicadas por 100 para facilitar o comparativo entre os resultados.

Tabela 8 – Médias absolutas de PVIs brutos e normalizados e desvio-padrão (entre parênteses) para PB, inglês-L1, inglês-L2 (1) e inglês-L2

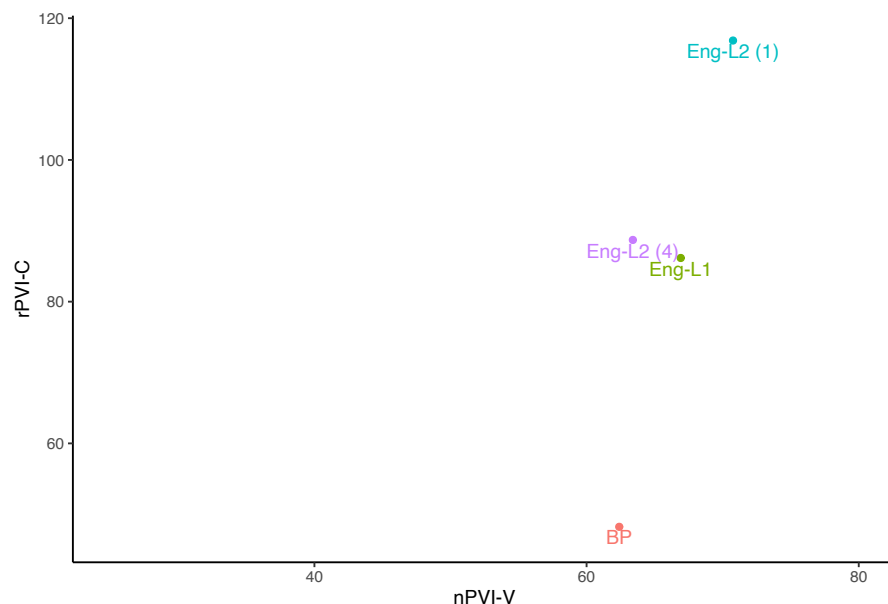
Língua	rPVI-V	rPVI-C	rPVI-VC	rPVI-S	nPVI-V	nPVI-C	nPVI-VC	nPVI-S
BP	65,1 (11,71)	48,22 (8,91)	64,73 (18,72)	102,85 (40,24)	62,4 (9,14)	53,96 (7,21)	59,76 (7,69)	46,6 (9,95)
Eng-L1	70,74 (14,84)	86,16 (20,23)	83,58 (11,12)	130,66 (33,59)	66,92 (12,78)	68,56 (11,72)	68,96 (11,09)	50,56 (7,97)
Eng-L2 (1)	96,98 (19,27)	116,84 (46,64)	114,4 (38,18)	176,27 (90,03)	70,76 (9,64)	72,36 (12,14)	72,6 (9,40)	51,48 (10,97)
Eng-L2 (4)	81,21 (15,75)	88,7 (21,69)	89,1 (14,53)	137,96 (44,62)	63,4 (5,83)	64,84 (11,46)	65,08 (8,12)	45,64 (9,21)

Fonte: elaborada pelo autor.

Primeiramente, os resultados obtidos neste estudo para o índice bruto de variabilidade pareada dos intervalos consonantais consecutivos (rPVI-C) e o índice normalizado de variabilidade pareada para os intervalos vocálicos consecutivos (nPVI-V) são apresentados e comparados aos dados do estudo seminal de Grabe e Low (2002), bem como Arvaniti (2012). Para tanto, rPVI-C e nPVI-V foram plotadas bidimensionalmente. Em seguida, os dados para os demais PVIs serão reportados.

O gráfico da figura 14 apresenta a distribuição dos resultados dos quatro corpora deste estudo obtidos pelas métricas rPVI-C e nPVI-V:

Figura 14 – Inglês-L1, inglês-L2(1), inglês-L2(4) e PB no plano formado pelos eixos rPVI-C e nPVI-V



Fonte: Teixeira e Lima Jr. (2021)

No plano formado pelos eixos rPVI-C e nPVI-V, inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB também assumiram posições distintas, embora as maiores distâncias tenham sido registradas no eixo rPVI-C ($rPVI-C_{PB} = 48,22$; $rPVI-C_{\text{inglês-L1}} = 86,16$; $rPVI-C_{\text{inglês-L2(1)}} = 116,84$; $rPVI-C_{\text{inglês-L2(4)}} = 88,70$), com PB e inglês-L2(1) mais uma vez ocupando posições diametralmente opostas, e inglês-L1 e inglês-L2(4) ocupando posições intermediárias e muito próximas, indicando um percurso de desenvolvimento rumo aos padrões prosódicos da língua não nativa, pelo menos em relação à variabilidade dos intervalos consonantais adjacentes. Já em relação ao eixo nPVI-V, os resultados são menos elucidativos, embora os dados indiquem uma maior tendência ao ritmo silábico para o PB ($nPVI-V = 62,4$) em comparação aos dados do inglês-L1 ($nPVI-V = 66,92$) e aos dados de interlíngua ($nPVI-V_{\text{inglês-L2(1)}} = 70,76$; $nPVI-V_{\text{inglês-L2(4)}} = 63,4$).

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada métrica. O modelo para rPVI-C revelou efeitos significativos de cada uma das (inter)línguas sobre a métrica. O modelo de nPVI-V, no entanto, apresentou efeito apenas para a primeira gravação de inglês-L2 em relação ao PB, conforme mostra o resumo dos modelos da tabela 9.

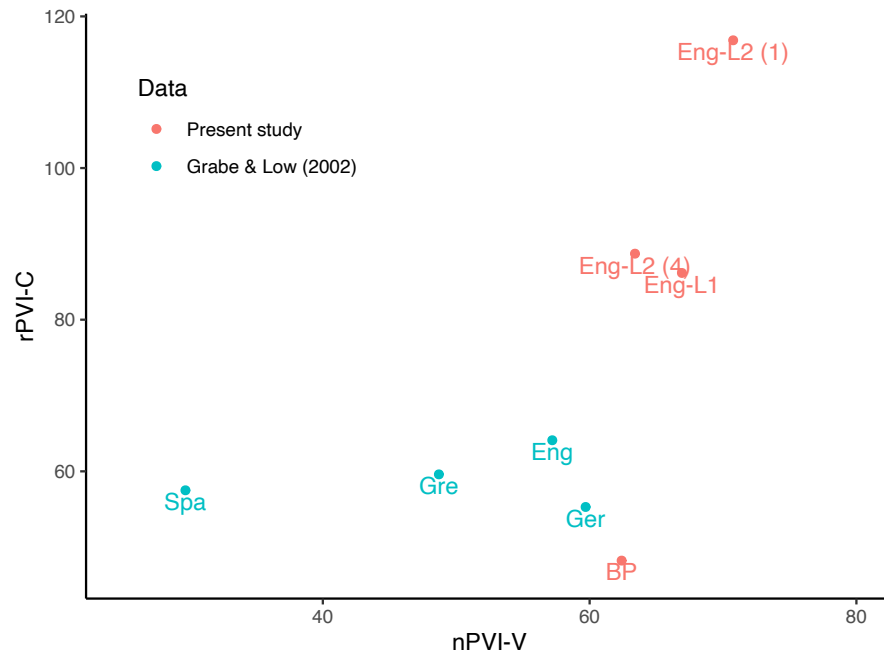
Tabela 9 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para rPVI-C e nPVI-V. Modelos: $rPVI-C \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $nPVI-V \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	rPVI-C			nPVI-V		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	48.09	34.98 – 61.20	<0.001	61.23	55.38 – 67.08	<0.001
Lang [Eng-L1]	38.40	21.81 – 55.00	<0.001	5.27	-2.15 – 12.69	0.164
Lang [Eng-L2 (1)]	69.07	53.98 – 84.16	<0.001	8.99	4.04 – 13.95	<0.001
Lang [Eng-L2 (4)]	40.48	25.68 – 55.28	<0.001	1.00	-3.69 – 5.69	0.676

Fonte: Teixeira e Lima Jr. (2021)

Os dados obtidos neste estudo para rPVI-C e nPVI-V foram comparados, primeiramente, aos resultados reportados por Grabe e Low (2002), quando fizeram a proposta da métrica, conforme figura 15.

Figura 15 – Dados do presente estudo e dos de Grabe e Low (2002) para rPVI-C e nPVI-V

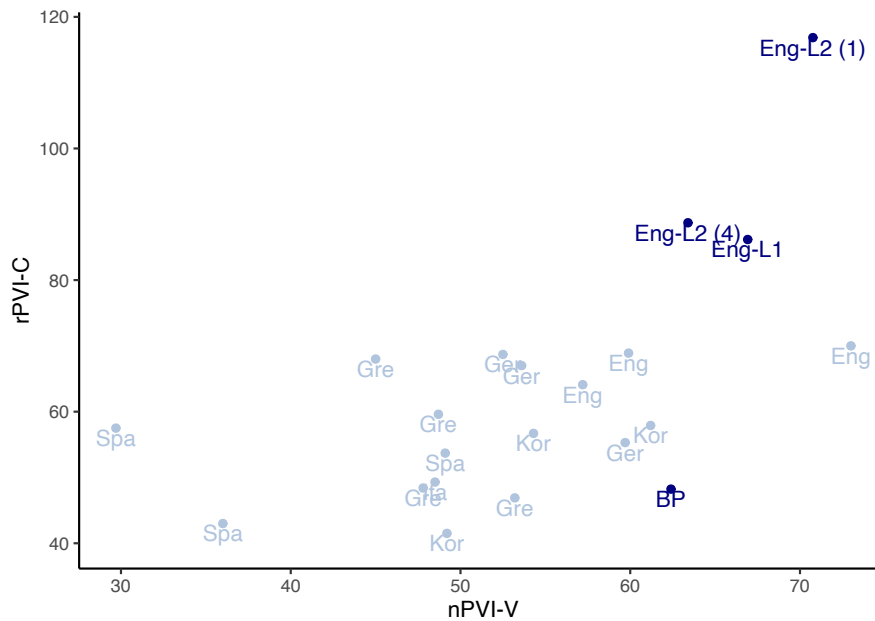


Fonte: Teixeira e Lima Jr. (2021)

Os valores deste estudo para nPVI-V foram maiores do que os encontrados por Grabe e Low (2002), distanciando-se, principalmente, dos dados do grego e do espanhol. Como, em teoria, valores menores de nPVI-V são esperados para línguas de ritmo mais silábico, o comparativo entre os dois estudos mostrou-se inconsistente. Já em relação ao eixo rPVI-C, o BP registrou o menor valor dentre todas as línguas em ambos os estudos, e as interlínguas, seguidas pelo inglês-L1, os maiores valores.

Para uma comparação ainda mais completa, apresentamos os dados deste estudo em meio a todos os dados resenhados e obtidos por Arvaniti (2012) na figura 16. Os dados do inglês-L1 (1) despontam como os mais extremos no eixo nPVI-C e com a segunda maior média para nPVI-V, menor apenas que a média reportada por White e Mattys (2007). Em seguida, tanto os valores do inglês-L1 quanto do inglês-L1 (4), que se aproximam bastante em ambos os eixos, registram valores mais elevados em relação a todas as línguas do estudo, exceto inglês-L2 (1) e o inglês de White e Mattys (2007) no plano nPVI-V. Já o PB agrupa-se com línguas como o grego e o italiano no eixo VarcoC, mas se distancia delas em nPVI-V, apresentando o quinto maior valor nesse eixo.

Figura 16 – Dados do presente estudo (azul escuro) em meio a todos os dados resenhados e obtidos por Arvaniti (2012) (azul claro) para rPVI-C e nPVI-V, em que Eng = inglês, Ger = alemão, Gre = grego, Spa = espanhol, Ita = italiano, Kor = coreano.

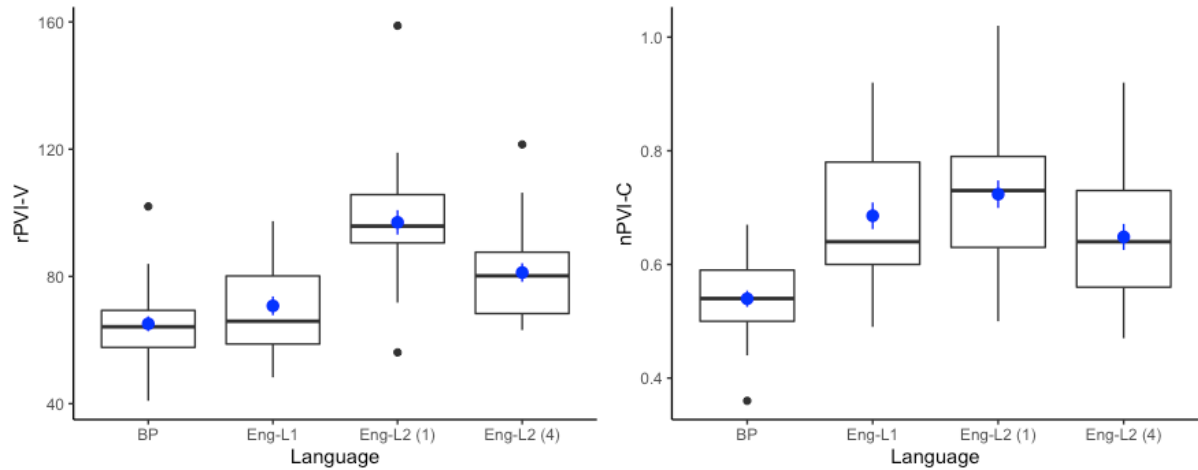


Fonte: Teixeira e Lima Jr. (2021)

Vejamos agora como os demais PVIs que integram este estudo comportaram-se em relação às (inter)línguas. Em relação ao PVI bruto para vogais (rPVI-V) e o PVI normalizado para consoantes (nPVI-C), a distribuição dos resultados apresentou padrão análogo às suas versões normalizada e bruta, respectivamente, como pode ser observado na figura 17.

Repetindo o padrão flagrado para a versão normalizada, inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB assumiram médias distintas para o PVI bruto das vogais, conforme explicitado visualmente na figura 17.1 ($rPVI-V_{PB} = 65,10$; $rPVI-V_{\text{inglês-L1}} = 70,74$; $rPVI-V_{\text{inglês-L2(1)}} = 96,98$; $rPVI-V_{\text{inglês-L2(4)}} = 81,21$). Mais uma vez PB e inglês-L2(1) ocuparam posições opostas, e inglês-L1 e inglês-L2(4) ocuparam posições intermediárias e mais próximas, indicando um percurso de desenvolvimento rumo aos padrões prosódicos da língua não nativa. O PVI normalizado para consoantes (nPVI-C) também repetiu o mesmo padrão de sua versão bruta, como mostra a figura 17.2, com os dados do PB apresentando a menor variabilidade, indicando uma maior tendência ao ritmo silábico ($nPVI-C_{PB} = 53,96$) em comparação aos dados do inglês-L1 ($nPVI-C = 68,56$) e aos dados de interlíngua ($nPVI-C_{\text{inglês-L2(1)}} = 72,36$; $nPVI-C_{\text{inglês-L2(4)}} = 64,84$). Conforme pode ser observado na figura 17, o nPVI-C foi mais eficiente na separação entre as L1s, ao passo que o rPVI-V separou melhor as interlínguas.

Figura 17 – *Boxplots* das médias de rPVI-V (à esquerda, 17.1) e nPVI-C (à direita, 17.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para rPVI-V e nPVI-C. O modelo para nPVI-C revelou efeitos significativos de cada uma das (inter)línguas sobre a métrica. O modelo de rPVI-V não apresentou efeitos significativos apenas para o inglês-L1 em relação ao PB, conforme mostra o resumo dos modelos da tabela 10.

Tabela 10 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para nPVI-C e rPVI-V. Modelos: $nPVIC \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $rPVIV \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

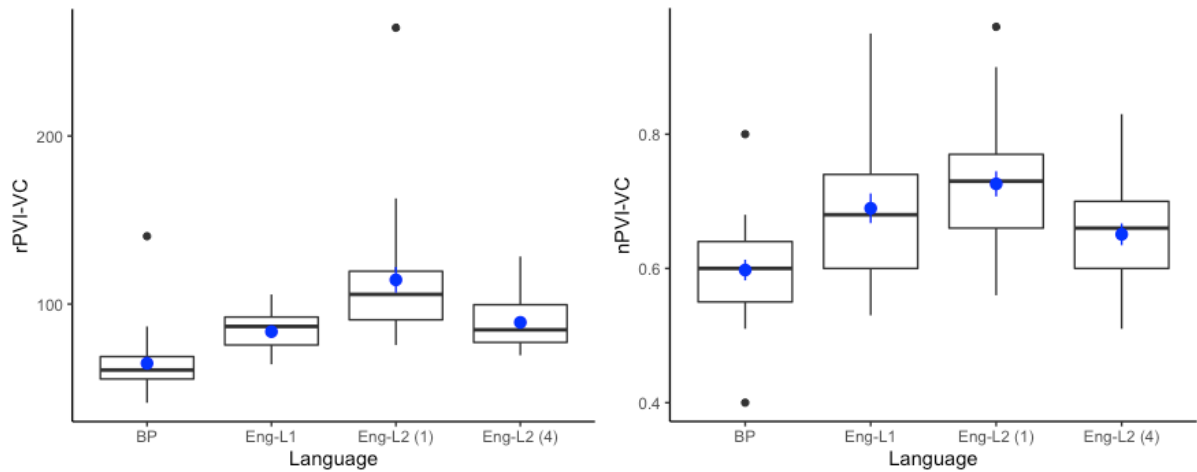
<i>Predictors</i>	nPVIC			rPVIV		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.54	0.47 – 0.60	<0.001	65.11	59.00 – 71.21	<0.001
Lang [Eng-L1]	0.15	0.08 – 0.22	<0.001	5.64	-3.00 – 14.27	0.198
Lang [Eng-L2 (1)]	0.18	0.13 – 0.24	<0.001	31.88	23.25 – 40.52	<0.001
Lang [Eng-L2 (4)]	0.11	0.06 – 0.16	<0.001	16.11	7.47 – 24.75	<0.001

Fonte: elaborada pelo autor.

A hierarquia de valores [+ acentual] inglês-L2 (1) > inglês-L2 (4) > inglês-L1 > PB [+ silábico] também foi detectada para os PVIs brutos e normalizados dos segmentos vocálicos e consonantais consecutivas (rPVI-VC e nPVI-VC, respectivamente) e para o PVI bruto das

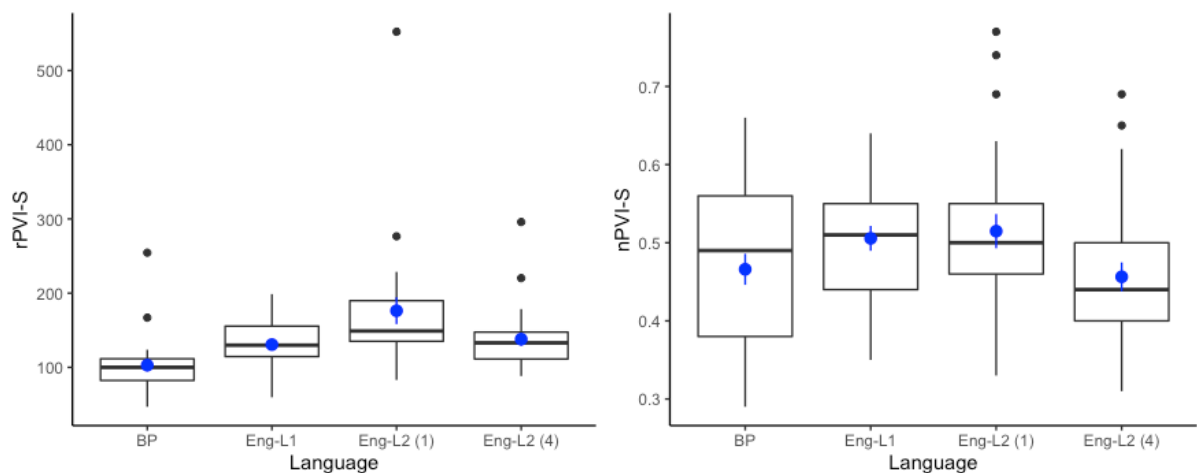
sílabas fonéticas (rPVI-S), mas não foi detectada para o PVI normalizada das sílabas fonéticas (nPVI-S) conforme evidência visual dos gráficos das figuras 18 e 19.

Figura 18 – *Boxplots* das médias de rPVI-VC (à esquerda, 18.1) e nPVI-S (à direita, 18.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Figura 19 – *Boxplots* das médias de rPVI-S (à esquerda, 19.1) e nPVI-S (à direita, 19.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Os resultados para rPVI-VC, nPVI-VC e ePVI-S (figuras 18.1, 18.2 e 19.1, respectivamente) posicionaram o PB com maior tendência ao ritmo silábico, com os menores

valores médios ($rPVI-VC_{PB} = 64,73$; $nPVI-VC_{PB} = 59,76$; $rPVI-S_{PB} = 102,85$). O inglês-L1 registrou maior variabilidade para esses índices, ocupando uma posição intermediária entre o PB e as interlínguas ($rPVI-VC_{\text{inglês-L1}} = 83,58$; $nPVI-VC_{\text{inglês-L1}} = 68,96$; $rPVI-S_{\text{inglês-L1}} = 130,66$) e o inglês-L2(4) apresentou valores muito próximos ao inglês-L1 ($rPVI-VC_{\text{inglês-L2(4)}} = 89,10$; $nPVI-VC_{\text{inglês-L2(4)}} = 65,08$; $rPVI-S_{\text{inglês-L2(4)}} = 137,96$), deflagrando um percurso de desenvolvimento pelos aprendizes rumo aos padrões prosódicos da L2. Já o inglês-L2(1) apresentou os valores mais extremos para essas métricas, registrando médias mais distantes do PB e superando também o inglês-L1 ($rPVI-VC_{\text{inglês-L2(1)}} = 114,40$; $nPVI-VC_{\text{inglês-L2(1)}} = 72,60$; $rPVI-S_{\text{inglês-L2(1)}} = 176,27$).

Já os valores médios registrados para o PVI normalizado das sílabas fonéticas mostraram-se menos conclusivos, com grande sobreposição de dados, conforme pode ser observado no gráfico na figura 19.2. Mesmo assim, houve consistência em relação à menor variabilidade do PB ($nPVI-S_{PB} = 46,60$) em comparação ao inglês-L1 ($nPVI-S_{\text{inglês-L1}} = 50,56$). Em relação às médias dessa métrica para as interlínguas, o inglês-L2(1) registrou uma média bastante próxima ao inglês-L1 ($nPVI-S_{\text{inglês-L2(1)}} = 51,48$) e o inglês-L2(4) aproximou-se mais do PB ($nPVI-S_{\text{inglês-L2(4)}} = 45,64$), com a menor média de $nPVI-S$ dentre os corpora analisados.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para $rPVI-VC$ e $nPVI-VC$. Ambos os modelos revelaram efeitos significativos de cada uma das (inter)línguas sobre essas métricas, como pode ser observado no resumo dos modelos da tabela 11.

Tabela 11 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para $rPVI-VC$ e $nPVI-VC$. Modelos: $rPVIVC \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $nPVIVC \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	rPVIVC			nPVIVC		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	64.11	53.81 – 74.41	<0.001	0.60	0.56 – 0.63	<0.001
Lang [Eng-L1]	19.48	5.31 – 33.64	0.007	0.09	0.04 – 0.14	0.001
Lang [Eng-L2 (1)]	50.37	37.67 – 63.07	<0.001	0.13	0.08 – 0.18	<0.001
Lang [Eng-L2 (4)]	24.37	11.95 – 36.79	<0.001	0.05	0.00 – 0.10	0.043

Fonte: elaborada pelo autor.

Também foram ajustados dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para $rPVI-S$ e $nPVI-S$ (tabela 12). Como esperado, apenas $rPVI-S$ apresentou efeitos significativos e apenas para as interlínguas.

Tabela 12 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para rPVI-S e nPVI-S. Modelos: rPVIVC ~ $\text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e nPVIVC ~ $\text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	rPVIS			nPVIS		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	101.55	74.26 – 128.84	<0.001	0.47	0.43 – 0.50	<0.001
Lang [Eng-L1]	29.12	-7.13 – 65.37	0.115	0.04	-0.01 – 0.09	0.148
Lang [Eng-L2 (1)]	75.55	45.02 – 106.08	<0.001	0.05	-0.01 – 0.10	0.075
Lang [Eng-L2 (4)]	35.11	5.52 – 64.69	0.020	-0.01	-0.06 – 0.04	0.724

Fonte: elaborada pelo autor.

4.1.4 RRs

Nesta subseção, são apresentados os resultados para a Taxa de Ritmo (*rhythm ratio* – RR), que é uma variação do índice de variabilidade pareada (PVI), diferindo apenas quanto à técnica de normalização, que consiste na divisão entre a duração de dois intervalos adjacentes, em que o mais longo sempre será dividido pelo mais curto, cumulando em um numerador sempre positivo que será dividido por sua vez pelo número total de sílabas menos 1 e multiplicado por 100. Valores mais baixos de RR são esperados para línguas de ritmo mais acentual, sendo 100 o valor de um inventário fonético perfeitamente isócrono.

A tabela 13 apresenta as médias da taxa de ritmo (RR) aplicadas às vogais (RR-V), às consoantes (RR-C), às unidades vocálicas e consonantais consecutivas (RR-VC) e às sílabas fonéticas (RR-S) para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4).

Tabela 13 – Médias absolutas e desvio-padrão (entre parênteses) de RR-V, RR-C, RR-VC e RR-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2

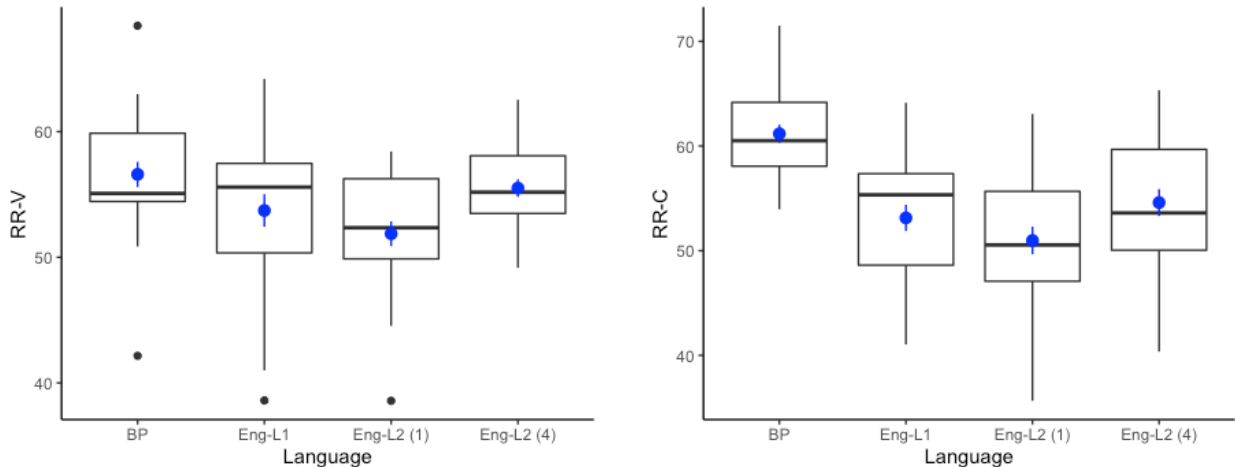
Língua	RR-V	RR-C	RR-VC	RR-S
BP	56,59 (5,01)	61,17 (4,36)	58,07 (4,35)	65,9 (5,89)
Eng-L1	53,72 (6,49)	53,13 (6,29)	52,8 (5,65)	63,4 (4,96)
Eng-L2 (1)	51,87 (4,96)	50,97 (6,59)	50,91 (4,97)	62,78 (6,07)
Eng-L2 (4)	55,5 (3,51)	54,59 (6,42)	54,55 (4,59)	66,21 (5,46)

Fonte: elaborada pelo autor.

No comparativo entre as L1s, os resultados para RR aplicados a todos os segmentos sugerem que o PB apresenta maior tendência ao ritmo silábico ($RR-V_{PB} = 56,59$; $RR-C_{PB} =$

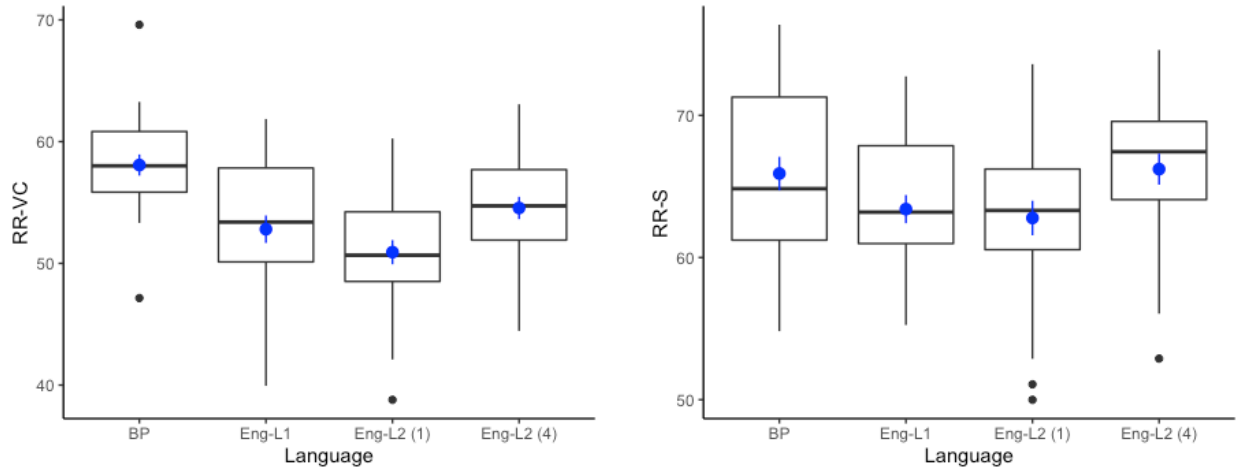
61,17; $RR-VC_{PB} = 58,07$; $RR-S_{PB} = 65,9$) em comparação ao inglês-L1 ($RR-V_{\text{inglês-L1}} = 53,72$; $RR-C_{\text{inglês-L1}} = 53,13$; $RR-VC_{\text{inglês-L1}} = 52,8$; $RR-S_{\text{inglês-L1}} = 63,4$), já que apresentou médias mais elevadas para essa métrica. Já os dados das interlínguas figuram como menos elucidativos, com o inglês-L2(1) despontando como a interlíngua com maior variabilidade duracional em relação à taxa de ritmo dentre os corpora analisados ($RR-V_{\text{inglês-L2(1)}} = 51,87$; $RR-C_{\text{inglês-L2(1)}} = 50,97$; $RR-VC_{\text{inglês-L2(1)}} = 50,91$; $RR-S_{\text{inglês-L2(1)}} = 62,78$), e o inglês-L2(4) ocupando uma posição intermediária entre o PB e o inglês-L1 ($RR-V_{\text{inglês-L2(4)}} = 55,5$; $RR-C_{\text{inglês-L2(4)}} = 54,59$; $RR-VC_{\text{inglês-L2(4)}} = 54,55$), exceto pela taxa de ritmo das sílabas fonéticas (RRS), para a qual a média obtida para o inglês-L2(4) supera a média do PB ($RR-S_{\text{inglês-L2(4)}} = 66,21$). Como pode ser observado nos *boxplots* das figuras 20 e 21, os valores registrados para RR apresentam alto grau de sobreposição de dados, sobretudo o RR das sílabas fonéticas (RR-S). Já o RR-C emergiu como a métrica mais eficiente na separação das L1s.

Figura 20 – *Boxplots* das médias de RR-V (à esquerda, 20.1) e RR-C (à direita, 20.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Figura 21 – *Boxplots* das médias de RR-VC (à esquerda, 21.1) e RR-S (à direita, 21.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Apesar de as médias de RR não terem diferido radicalmente nas médias absolutas, o padrão observado para o inglês-L2(1) como (inter)língua com valores mais extremos também foi replicado para essa métrica, embora o percurso de desenvolvimento flagrado não seja tão nítido. A fim de verificar a significância desses resultados, dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para RR-V e RR-C. Ambos revelaram efeitos significativos das (inter)línguas sobre as métricas, com exceção de inglês-L1 e do inglês-L2(4) em comparação ao PB para RR-V, conforme mostra o resumo dos modelos da tabela 14.

Tabela 14 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para RR-V e RR-C. Modelos: $RRV \sim Lang + (1|Chunk) + (1|Speaker)$ e $RRC \sim Lang + (1|Chunk) + (1|Speaker)$

<i>Predictors</i>	RR-V			RR-C		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	57.01	54.05 – 59.96	<0.001	61.17	58.80 – 63.55	<0.001
Lang [Eng-L1]	-3.29	-7.04 – 0.46	0.086	-8.04	-11.40 – -4.68	<0.001
Lang [Eng-L2 (1)]	-5.07	-7.72 – -2.43	<0.001	-10.20	-13.56 – -6.84	<0.001
Lang [Eng-L2 (4)]	-1.09	-3.61 – 1.42	0.394	-6.58	-9.94 – -3.22	<0.001

Fonte: elaborada pelo autor.

Outros dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para RR-VC e RR-S. RR-VC revelou efeitos significativos de todas as (inter)línguas sobre essa métrica, o que não foi atestado para RR-S, conforme mostra o resumo dos modelos da tabela 15.

Tabela 15 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para RR-VC e RR-S. Modelos: RRVC ~ Lang + (1|Chunk) + (1|Speaker) e RRS ~ Lang + (1|Chunk) + (1|Speaker)

<i>Predictors</i>	RR-VC			RR-S		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	58.36	55.43 – 61.30	<0.001	65.91	63.68 – 68.14	<0.001
Lang [Eng-L1]	-5.56	-9.04 – -2.08	0.002	-2.50	-5.66 – -0.65	0.118
Lang [Eng-L2 (1)]	-7.40	-9.90 – -4.89	<0.001	-3.13	-6.28 – -0.03	0.052
Lang [Eng-L2 (4)]	-3.53	-5.91 – -1.14	0.004	0.31	-2.85 – 3.46	0.846

Fonte: elaborada pelo autor.

4.1.5 VIs

Nesta subseção, são apresentados os resultados para VI (*variability index – VI*), que também se assemelha ao índice de variabilidade pareada (PVI), mas com uma técnica de normalização diferente: a duração de cada intervalo de referência é dividida pela duração média de todos os intervalos do mesmo tipo no enunciado. Em teoria, valores mais altos para VI representam maior variabilidade na duração de intervalos de referência consecutivos e, portanto, são indicativos de tendência ao ritmo acentual. A tabela 16 apresenta as médias e os desvios-padrão de VI para os intervalos vocálicos (VI-V), consonantais (VI-C), vocálico-consonantais consecutivos (VI-VC) e silábico-fonéticos (VI-S) para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4).

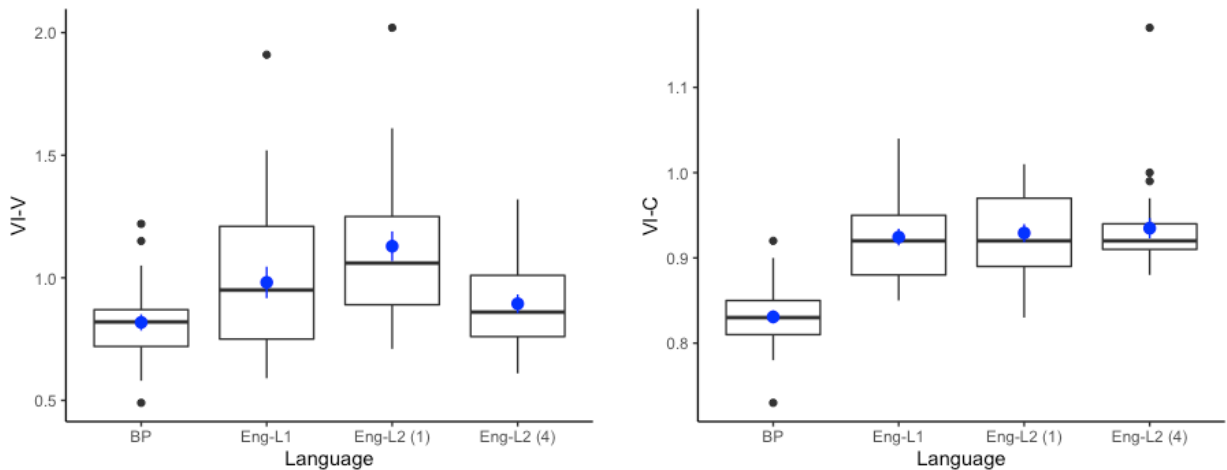
Tabela 16 – Médias absolutas e desvio-padrão (entre parênteses) de, VI-C, VI-VC e VI-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2

Língua	VI-V	VI-C	VI-VC	VI-S
BP	0,818 (0,166)	0,830 (0,037)	0,684 (0,101)	0,516 (0,120)
Eng-L1	0,981 (0,322)	0,924 (0,049)	0,834 (0,157)	0,606 (0,136)
Eng-L2 (1)	1,128 (0,302)	0,929 (0,052)	0,859 (0,120)	0,615 (0,160)
Eng-L2 (4)	0,894 (0,188)	0,934 (0,059)	0,746 (0,116)	0,538 (0,126)

Fonte: elaborada pelo autor.

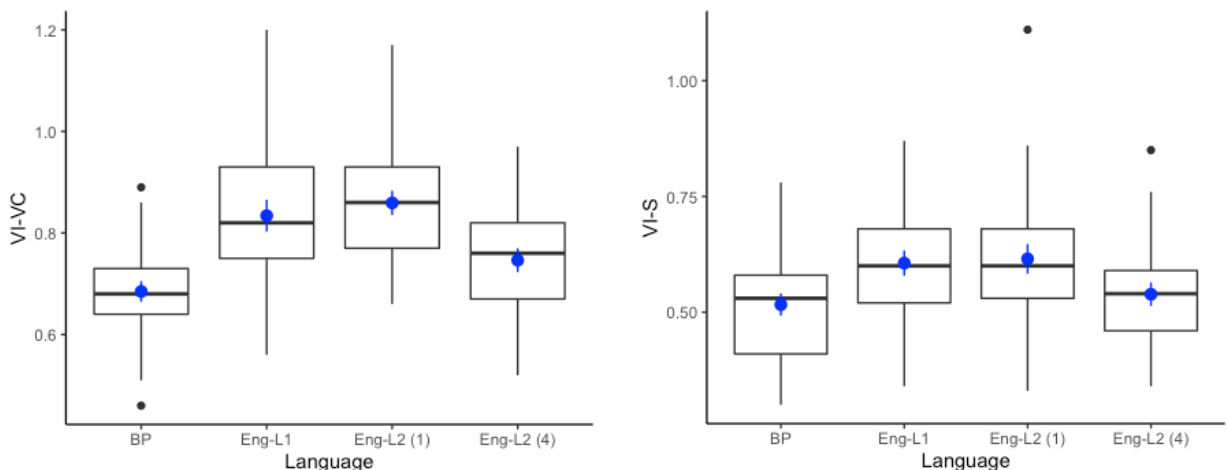
A análise dos resultados obtidos para VI posicionou o PB como o sistema com tendência mais silábica dentre todos os corpora analisados, conforme visualmente explicitado nos *boxplots* das figuras 22 e 23 ($VI-V_{PB} = 0,81$; $VI-C_{PB} = 0,83$; $VI-VC_{PB} = 0,68$; $VI-S_{PB} = 0,51$). Quando se compara essas médias com as do inglês-L1 ($VI-V_{\text{inglês-L1}} = 0,98$; $VI-C_{\text{inglês-L1}} = 0,92$; $VI-VC_{\text{inglês-L1}} = 0,83$; $VI-S_{\text{inglês-L1}} = 0,6$), percebe-se que as métricas de VI cumpriram seu papel em separar as L1s, especialmente VI-C e VI-VC. As maiores médias foram registradas por inglês-L2(1) ($VI-V_{\text{inglês-L2(1)}} = 1,12$; $VI-VC_{\text{inglês-L2(1)}} = 0,85$; $VI-S_{\text{inglês-L2(1)}} = 0,61$) e inglês-L2 (4) ($VI-C_{\text{inglês-L2(4)}} = 0,93$).

Figura 22 – *Boxplots* das médias de VI-V (à esquerda, 22.1) e VI-C (à direita, 22.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Figura 23 – *Boxplots* das médias de VI-VC (à esquerda, 23.1) e VI-S (à direita, 23.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para VI-V e VI-C. Ambos os modelos revelaram efeitos significativos de todas as (inter)línguas sobre as métricas, exceto inglês-L2(4) para VI-V, conforme mostra o resumo dos modelos da tabela 17.

Tabela 17 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VI-V e VI-C. Modelos: VI-V ~ Lang + (1|Chunk) + (1|Speaker) e VI-C ~ Lang + (1|Chunk) + (1|Speaker)

<i>Predictors</i>	VI-V			VI-C		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.80	0.66 – 0.95	<0.001	0.83	0.81 – 0.85	<0.001
Lang [Eng-L1]	0.18	0.00 – 0.35	0.048	0.09	0.07 – 0.12	<0.001
Lang [Eng-L2 (1)]	0.32	0.19 – 0.45	<0.001	0.10	0.07 – 0.13	<0.001
Lang [Eng-L2 (4)]	0.08	-0.05 – 0.20	0.226	0.10	0.08 – 0.13	<0.001

Fonte: elaborada pelo autor.

Outros dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para VI-VC e VI-S. Ambos os modelos revelaram efeitos significativos de todas as (inter)línguas sobre as métricas, exceto inglês-L2(4) para VI-S, conforme mostra o resumo dos modelos da tabela 18.

Tabela 18 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para VI-VC e VI-S. Modelos: VI-V ~ Lang + (1|Chunk) + (1|Speaker) e VI-C ~ Lang + (1|Chunk) + (1|Speaker)

<i>Predictors</i>	VI-VC			VI-S		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.67	0.59 – 0.76	<0.001	0.52	0.46 – 0.57	<0.001
Lang [Eng-L1]	0.16	0.05 – 0.27	0.003	0.09	0.01 – 0.17	0.022
Lang [Eng-L2 (1)]	0.19	0.13 – 0.25	<0.001	0.10	0.02 – 0.18	0.012
Lang [Eng-L2 (4)]	0.06	0.00 – 0.12	0.042	0.02	-0.05 – 0.10	0.563

Fonte: elaborado pelo autor.

Os resultados obtidos para as métricas com base em VI reforçam a tendência de ritmo mais silábica do PB em comparação com o inglês-L1, e a desorganização temporal das interlínguas.

4.1.6 YARDs

Nesta subseção, são apresentados os resultados para YARD (*Yet Another Rhythm Determination*), que, assim como o RR e o VI, é uma métrica local que pode ser considerada uma variação do PVI. O método de normalização aplicado para YARD consiste na utilização da transformada Z (*z-score*) para posterior cálculo do índice YARD. Valores mais elevados para YARD são indicativos de tendência ao ritmo acentual porque implica em maior variabilidade duracional entre os intervalos referência consecutivos.

A tabela 19 apresenta as médias e os desvios-padrão de YARD aplicadas às vogais (YARD-V), às consoantes (YARD-C), às unidades vocálicas e consonantais consecutivas (YARD-VC) e às sílabas fonéticas (YARD-S) para os dados de PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4).

Tabela 19 – Médias absolutas de YARD-V, YARD-C, YARD-VC e YARD-S para PB, inglês-L1, inglês-L2 (1) e inglês-L2

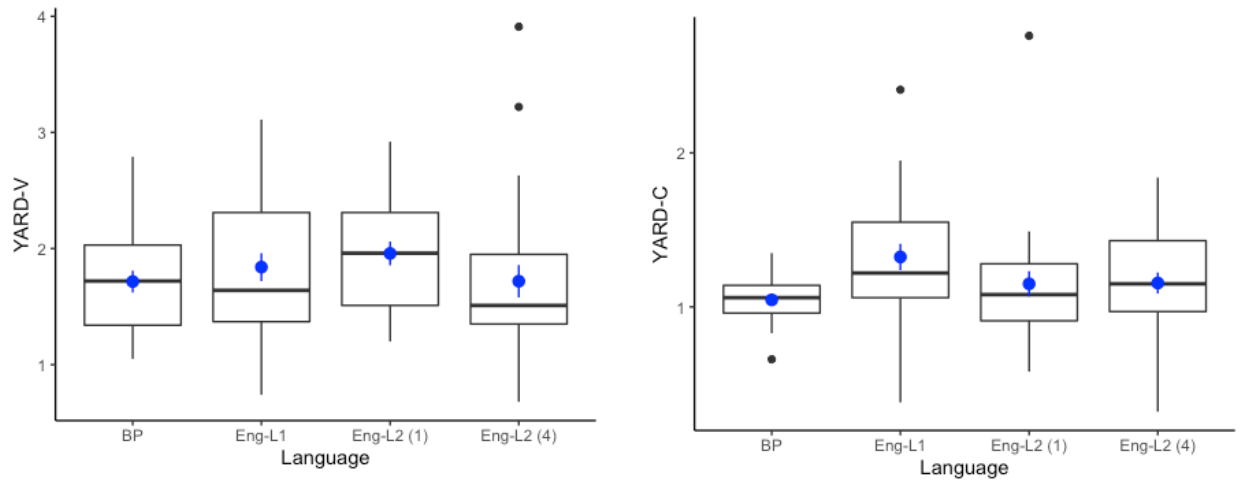
Língua	YARD-V	YARD-C	YARD-VC	YARD-S
BP	1,715 (0,472)	1.045 (0,155)	0.717 (0,150)	0.757 (0,163)
Eng-L1	1,840 (0,603)	1.324 (0,428)	0.695 (0,133)	0.709 (0,158)
Eng-L2 (1)	1,957 (0,511)	1.150 (0,412)	0.869 (0,113)	0.824 (0,170)
Eng-L2 (4)	1,718 (0,696)	1.155 (0,342)	0.848 (0,123)	0.759 (0,134)

Fonte: elaborada pelo autor.

Os resultados para YARD não atingiram significância, com grande sobreposição de valores, como pode ser observado nas figuras 24 e 25. Para YARD-V, o PB, juntamente com o inglês-L2(4) registraram as menores médias ($YARD-V_{PB} = 1,715$; $YARD-V_{\text{inglês-L2(4)}} = 1,718$), seguidos do inglês-L1 ($YARD-V_{\text{inglês-L1}} = 1,84$) e do inglês-L2(1) ($YARD-V_{\text{inglês-L2(1)}} = 1,957$), o que não evidencia um percurso de desenvolvimento nítido. Já YARD-C foi capaz de separar as L1s ($YARD-C_{PB} = 1,045$; $YARD-C_{\text{inglês-L1}} = 1,324$), mas registrou valores muito próximos para inglês-L2(1) e inglês-L2(4) ($YARD-C_{\text{inglês-L2(1)}} = 1,150$; $YARD-C_{\text{inglês-L2(4)}} = 1,115$), posicionando-as entre as L1s, com tendência mais acentual que o PB. Já as médias obtidas para YARD-VC e YARD-S agruparam as L1s e as interlínguas em espaços distintos de valores aproximados. O PB e o inglês-L1 registraram as menores médias para essas métricas ($YARD-VC_{PB} = 0,717$; $YARD-VC_{\text{inglês-L1}} = 0,695$; $YARD-S_{PB} = 0,757$; $YARD-S_{\text{inglês-L1}} = 0,709$), com o inglês-L1 apresentando uma tendência ligeiramente mais silábica que o PB. O inglês-L2(1) apresentou os valores mais elevados para YARD-VC (0,869) e YARD-S (0,824), seguido do

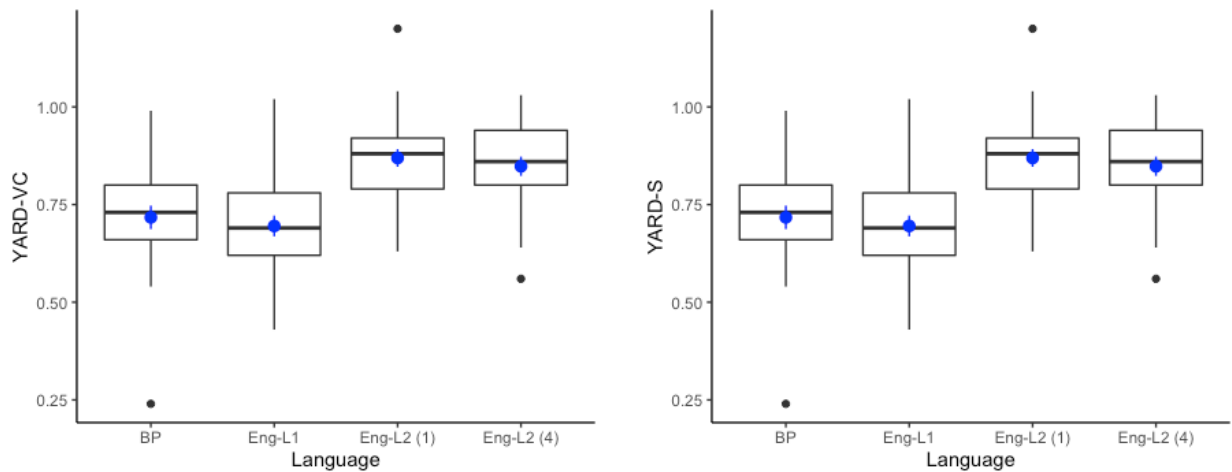
inglês-L2(4) ($\text{YARD-VC}_{\text{inglês-L2(4)}} = 0,848$; $\text{YARD-S}_{\text{inglês-L2(4)}} = 0,759$), cuja média para YARD-S praticamente idêntica à média do PB.

Figura 24 – *Boxplots* das médias de YARD-V (à esquerda, 24.1) e YARD-C (à direita, 24.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Figura 25 – *Boxplots* das médias de YARD-VC (à esquerda, 25.1) e YARD-S (à direita, 25.2) para inglês-L1, inglês-L2(1), inglês-L2(4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para YARD-V e YARD-C. Em ambos os modelos, não houve efeitos significativos de todas das (inter)línguas sobre as métricas, exceto inglês-L1 para YARD-C, conforme mostra o resumo dos modelos da tabela 20.

Tabela 20 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para YARD-V e YARD-C. Modelos: $YARD-V \sim Lang + (1|Chunk) + (1|Speaker)$ e $YARD-C \sim Lang + (1|Chunk) + (1|Speaker)$

<i>Predictors</i>	YARD-V			YARD-C		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.72	1.40 – 2.03	<0.001	1.05	0.91 – 1.19	<0.001
Lang [Eng-L1]	0.12	-0.26 – 0.51	0.536	0.28	0.08 – 0.48	0.006
Lang [Eng-L2 (1)]	0.22	-0.08 – 0.52	0.145	0.10	-0.09 – 0.30	0.297
Lang [Eng-L2 (4)]	0.00	-0.28 – 0.29	0.985	0.11	-0.09 – 0.31	0.273

Fonte: elaborada pelo autor.

Outros dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados para YARD-VC e YARD-S. Apenas o modelo implementado para YARD-VC revelou efeitos significativos das (inter)línguas, exceto para inglês-L1. Já o modelo para YARD-S não revelou efeitos significativos sobre essa métrica, como pode ser observado no resumo desses modelos na tabela 21.

Tabela 21 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para YARD-VC e YARD-S. Modelos: $YARD-VC \sim Lang + (1|Chunk) + (1|Speaker)$ e $YARD-S \sim Lang + (1|Chunk) + (1|Speaker)$

<i>Predictors</i>	YARD-VC			YARD-S		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.72	0.65 – 0.78	<0.001	0.76	0.70 – 0.82	<0.001
Lang [Eng-L1]	-0.02	-0.10 – 0.06	0.588	-0.05	-0.14 – 0.04	0.283
Lang [Eng-L2 (1)]	0.15	0.08 – 0.22	<0.001	0.07	-0.02 – 0.16	0.134
Lang [Eng-L2 (4)]	0.13	0.06 – 0.20	<0.001	0.00	-0.09 – 0.09	0.964

Fonte: elaborada pelo autor.

4.1.7 Distâncias euclidianas

Um aspecto de interesse para os objetivos deste estudo era verificar o posicionamento do PB, do inglês-L1, do inglês-L2 (1) e do inglês-L2 (4), observando se as métricas rítmicas flagrariam algum percurso de desenvolvimento entre a primeira e a quarta gravação dos aprendizes. Observou-se que os pares clássicos $\Delta C\text{-}\%V$, $VarcoC\text{-}VarcoV$ e $rPVI\text{-}C\text{-}nPVI\text{-}V$ captaram um percurso de desenvolvimento análogo, com os dados do inglês-L2 (1)

distantes tanto do PB quanto do inglês-L1, e os dados do inglês-L2(4) aproximando-se do inglês-L1, indicando uma tendência de (re) organização temporal rumo aos padrões prosódicos da língua não nativa. A fim de se contabilizar a distância entre as interlínguas e o inglês-L1 em cada plano rítmico formado pelos pares ΔC - $\%V$ (RAMUS; NESPOR; MEHLER, 1999), VarcoC-VarcoV (DELLWO, 2006; WHITE; MATTYS, 2007) e rPVI-C-nPVI-V (LOW; GRABE; NOLAN, 2000), calculamos, em cada um deles, as distâncias euclidianas entre inglês-L1 e as interlínguas, bem como entre as interlínguas e o PB para fins de comparação. A distância euclidiana é um cálculo de dissimilaridade, que é basicamente a soma das diferenças, e que pode ser utilizado para medir a distância entre dois pontos em um plano cartesiano, como os formados pelos pares de métricas. Neste caso, foram calculadas com:

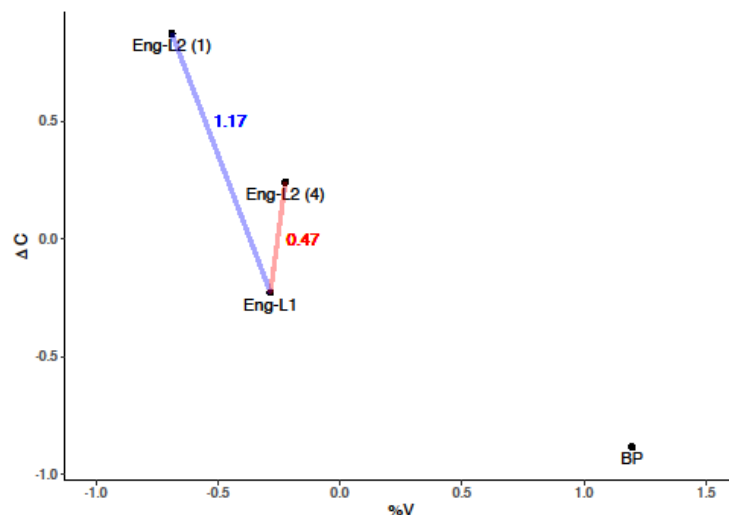
$$\sqrt{(y_{Ling1} - y_{Ling2})^2 + (x_{Ling1} - x_{Ling2})^2} \quad (5)$$

onde x y são duas métricas

e Ling as duas línguas comparadas

Para evitar que diferenças entre as escalas dos eixos x e y influenciassem os resultados, os cálculos foram realizados com valores normalizados das métricas (*z-score*), que colocam os dois eixos na mesma escala (de desvios-padrão). Para que o leitor tenha uma noção da escala resultante, a figura 26 ilustra as distâncias euclidianas de 1,17 entre o inglês-L1 e o inglês-L2 da primeira gravação, e de 0,47 entre o inglês-L1 e o inglês-L2 da quarta gravação no plano formado por ΔC e $\%V$.

Figura 26 – Distâncias euclidianas entre inglês-L1-L2(1) e inglês-L1-L2(4) no plano cartesiano das métricas ΔC e $\%V$.



Fonte: Teixeira e Lima Jr. (2021)

A tabela 22 apresenta as distâncias euclidianas nos planos rítmicos analisados. Para todas as métricas, a maior distância está entre o PB e o inglês-L1, seguida da distância entre o inglês-L1 e o inglês-L2 da primeira gravação, e com o inglês-L2 da quarta gravação sempre mais próximo ao inglês-L1:

Tabela 22 – Distâncias euclidianas entre inglês-L1 e interlínguas, e entre as L1s

Métrica	Línguas	Distância Euclidiana
$\Delta C-\%V$	PB-L1 – inglês-L1	1,62
	Inglês-L1 – L2(1)	1,17
	Inglês -L1 – L2(4)	0,47
Varcos	PB-L1 – inglês-L1	0,75
	Inglês -L1 – L2(1)	0,59
	Inglês -L1 – L2(4)	0,13
PVIIs	PB-L1 – inglês-L1	1,12
	Inglês -L1 – L2(1)	0,91
	Inglês -L1 – L2(4)	0,36

Fonte: Teixeira e Lima Jr. (2021)

De modo geral, a métrica que posicionou as línguas de maneira mais distante foi $\Delta C-\%V$, e a métrica que deixou as duas línguas menos separadas foi a que utilizou Varcos. Esses resultados complementam a observação visual captada pelos planos rítmicos bidimensionais e atestam quantitativamente o percurso de desenvolvimento dos aprendizes na dimensão duracional, que pode representar um efeito do ensino explícito de pronúncia.

4.2 Parâmetros acústicos

Os parâmetros acústicos tomam por base correlatos outros que não a duração com vistas a investigar o ritmo em suas múltiplas dimensões. Neste trabalho são analisados parâmetros acústicos com base em f_0 , intensidade e sonoridade.

De acordo com Cumbers (2013), alguns autores sugerem o termo frequência fundamental da fala para referir-se aos padrões de f_0 manifestados em situações de conversação ou fala em cadeia, tendo em vista que o termo f_0 pode ser aplicado a qualquer tipo de realização fonética (por exemplo, uma vogal prolongada). Neste sentido, o termo SFF representa a taxa média de vibração das pregas vocais durante a fala conectada, cuja característica distintiva em relação à produção de segmentos isolados é a gama de frequências e a mudança contínua de padrões vibratórios. Essa distinção é importante porque medidas de SFF não refletem, por exemplo, máximas ou mínimas de f_0 fisicamente possíveis de serem produzidas por um indivíduo, mas antes a gama de variação de f_0 manifestada na fala. Neste trabalho, f_0 equivale a

SFF. Uma outra distinção importante em estudos envolvendo medidas de f_0 é a diferença entre medidas distribucionais de longo termo (*long-term distributional – LTD measures*) e índices “linguísticos”, em que picos, mínimos e extensão espacial de f_0 se conectam a marcações de contorno entonacional, que, por sua vez, se relacionam a aspectos fonológicos e linguísticos propriamente ditos (MENNEN; SCHAEFFLER; DOCHERTY, 2012). Os parâmetros acústicos com base em f_0 investigados neste trabalho encaixam-se na primeira categoria, embora o modelo de segmentação das unidades de referência tenha adotado, de forma não estrita, critérios prosódico-sintáticos que permitem, no futuro, a análise desses índices “linguísticos”.

Com base em f_0 , foram analisados os parâmetros acústicos de mediana de f_0 ($f_0\text{med}$), pico de f_0 ($f_0\text{peak}$), f_0 mínimo ($f_0\text{min}$), desvio-padrão de f_0 (σf_0), assimetria de f_0 ($f_0\text{sk}$), média da primeira derivada de f_0 ($\mu\Delta 1-f_0$), primeira derivada do desvio-padrão de f_0 ($\sigma\Delta 1-f_0$), primeira derivada da assimetria de f_0 ($\text{sk}\Delta 1-f_0$) e taxa de f_0 ($f_0\text{R}$). Outros parâmetros acústicos analisados nesta seção são a ênfase espectral, cujo correlato de proeminência subjacente é a intensidade, isto é, a quantidade de energia da onda sonora advinda do esforço respiratório/vocal empreendido (GONÇALVES, 2018); a média normalizada da duração do pico silábico, que se relaciona à sonoridade e à percepção de altura; e a taxa de elocução. A abrangência de múltiplas dimensões fonético-acústicas na análise permite observar de que maneira a alternância de unidades mais e menos proeminentes correlacionam-se nas (inter)línguas, contribuindo para uma caracterização multidimensional do ritmo nos corpora analisados.

A tabela 23 apresenta as médias dos parâmetros acústicos de mediana de f_0 ($f_0\text{med}$), pico de f_0 ($f_0\text{peak}$), f_0 mínimo ($f_0\text{min}$), desvio-padrão de f_0 (σf_0), assimetria de f_0 ($f_0\text{sk}$), média da primeira derivada de f_0 ($\mu\Delta 1-f_0$), primeira derivada do desvio-padrão de f_0 ($\sigma\Delta 1-f_0$), primeira derivada da assimetria de f_0 ($\text{sk}\Delta 1-f_0$), taxa de f_0 ($f_0\text{R}$), ênfase espectral (emph), média normalizada da duração do pico silábico ($\mu\text{dur-Sil}$) e taxa de elocução (SR):

Tabela 23 – Médias absolutas dos parâmetros acústicos para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4)

Língua	$f_0\text{med}$	$f_0\text{peak}$	$f_0\text{min}$	σf_0	$f_0\text{sk}$	$\mu\Delta 1-f_0$	$\sigma\Delta 1-f_0$	$\text{sk}\Delta 1-f_0$	$f_0\text{R}$	emph	$\mu\text{dur-Sil}$	SR
BP	85,31	91,83	78,23	2,62	0,06	-0,317	3,61	-0,39	1,195	2,73	526,39	5,22
Eng-L1	88,74	96,06	78,88	3,79	0,05	-0,381	5,51	-0,20	1,164	4,34	653,47	4,43
Eng-L2 (1)	87,18	93,72	79,23	2,71	0,04	-0,315	3,73	-0,26	1,160	3,23	576,36	3,59
Eng-L2 (4)	85,70	93,10	77,80	3,34	0,02	-0,158	4,61	0,04	1,178	2,56	477,55	3,74

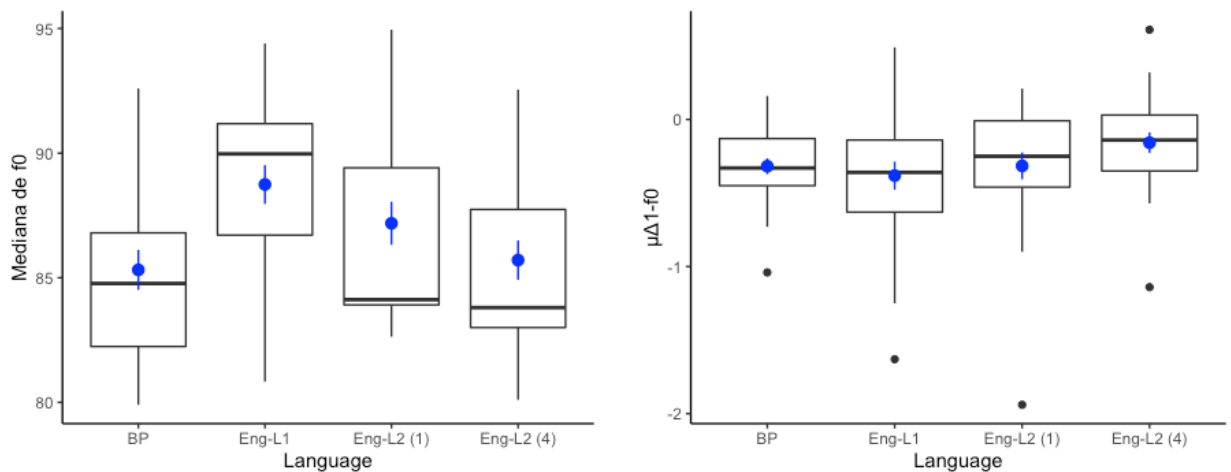
Fonte: elaborada pelo autor.

A maioria dos parâmetros acústicos com base em f_0 registrou valores mais elevados para o inglês-L1, sinalizando a importância desse correlato de proeminência nessa língua. O inglês-L1 também registrou a maior média de ênfase espectral, cujo correlato de proeminência é a intensidade. Já a taxa de elocução (SR) registrou os menores valores de média para ambas as interlínguas, o que pode estar relacionado a um maior nível de atenção ao nível segmental. Nas subseções a seguir, esses parâmetros serão reportados de forma pormenorizada.

4.2.1 Mediana de f_0 e média da primeira derivada de f_0 ($\mu\Delta 1-f_0$)

A mediana é uma medida de tendência central que separa a metade dos valores mais elevados e a metade dos valores mais baixos correspondentes a um conjunto de dados ordenados. Portanto, a mediana de f_0 corresponde ao ponto em que metade dos valores da amostra estão localizados acima dele e a outra metade abaixo. Ao contrário da média, a mediana não sofre influência de valores extremos e, por esse motivo, é considerada mais robusta em relação à primeira. A fim de verificar a dispersão dos dados de média em relação à mediana, também analisamos a média da primeira derivada de f_0 . Os *boxplots* da figura 27 apresentam a distribuição dos resultados dos quatro corpora deste estudo obtidos para a mediana de f_0 e média da primeira derivada de f_0 :

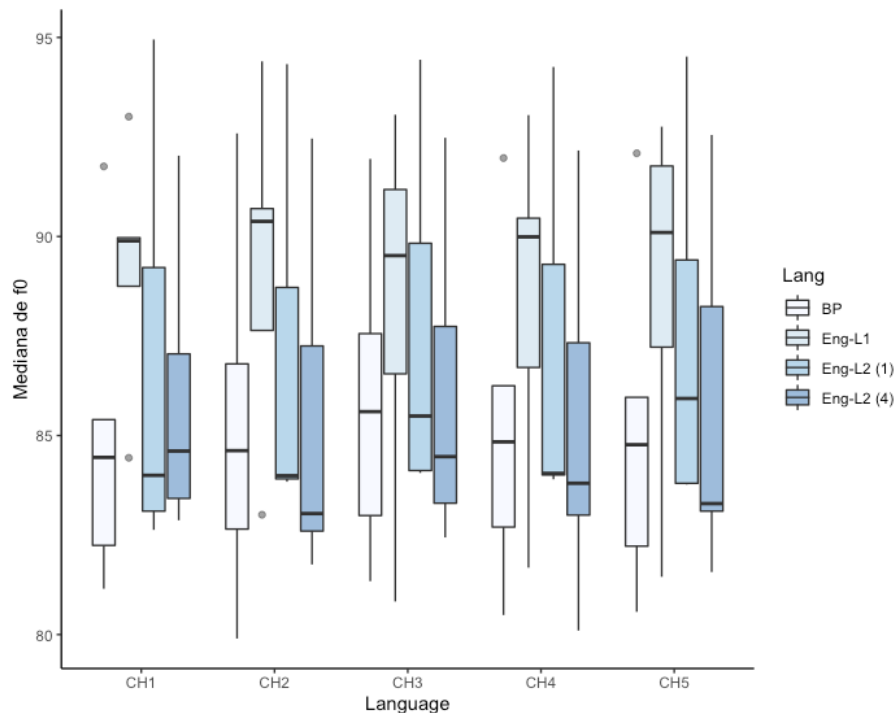
Figura 27 – *Boxplots* da mediana de f_0 para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB (à esquerda, 27.1) e média da primeira derivada de f_0 (à direita, 27.2). Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

O inglês-L1 apresentou o maior valor para mediana de f_0 ($f_{0med_{\text{inglês-L1}}} = 88,74$) e o PB, o menor valor ($f_{0med_{\text{PB}}} = 85,31$). Já o inglês-L2(1) apresentou um valor de mediana de f_0 muito próximo ao inglês-L1 ($f_{0med_{\text{inglês-L2(1)}}} = 87,18$), ao passo que o inglês-L2(4) apresentou um valor intermediário para esse parâmetro ($f_{0med_{\text{inglês-L2(4)}}} = 85,70$), mais próximo do PB. Também analisamos o comportamento desse parâmetro de acordo com as unidades sintático-prosódicas - *chunks*, tendo em vista que valores de f_0 podem correlacionar-se aos limites das estruturas sintáticas e podem estar associados a determinadas informações linguísticas, por exemplo, se estamos lidando com uma oração interrogativa ou declarativa. Na figura 28, a distribuição dos valores para as medianas dos quatro corpora analisados é apresentada por *chunks*.

Figura 28 – Medianas de f_0 para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por *chunk*



Fonte: elaborado pelo autor.

Como pode ser observado, PB e inglês-L1 assumiram posições distintas em todos os *chunks*, o que indica que essas línguas mobilizam diferentes padrões de f_0 . Nota-se também que os valores de mediana para o inglês-L1 foram mais altos para os dois primeiros *chunks*, justamente os que contêm frases interrogativas, o que pode estar associado à tendência dos falantes nativos da língua inglesa em elevar a frequência fundamental e, conseqüentemente, o *pitch* para orações interrogativas fechadas (e.g. “*Is English your native language?*”). Em

relação às interlínguas, as medianas do inglês-L2(1) aproximam-se mais do inglês-L1, exceto para o primeiro *chunk*.

Em relação à média da primeira derivada de f_0 , conforme pode ser observado no *boxplot* da figura 27 (*boxplot* 27.2), o inglês-L1 registrou a maior média negativa de coeficiente angular ($\mu\Delta 1 - f_{0\text{inglês-L1}} = -0,381$), seguido pelo PB ($\mu\Delta 1 - f_{0\text{PB}} = -0,317$), o inglês-L2(1) ($\mu\Delta 1 - f_{0\text{inglês-L2(1)}} = -0,315$), e o inglês-L2(4) ($\mu\Delta 1 - f_{0\text{inglês-L2(4)}} = -0,158$). O fato de as derivadas para a média terem registrado valores negativos evidencia uma tendência decrescente em relação aos valores a partir da média, o que indica que os valores registrados para a média sofreram a influência de valores mais extremos. Essa constatação, combinada à convergência dos resultados para as (inter)línguas, corrobora a robustez da mediana de f_0 em relação à média de f_0 como medida de registro ou altura habitual desse correlato de proeminência.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada parâmetro acústico, conforme resumo apresentado na tabela 24. O modelo para $f_{0\text{med}}$ revelou efeitos significativos apenas para a primeira gravação do inglês-L2 em relação ao PB, e o modelo de $\mu\Delta 1 - f_0$ não apresentou efeitos significativos.

Tabela 24 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para $f_{0\text{med}}$ e $\mu\Delta 1 - f_0$. Modelos: $f_{0\text{med}} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\mu\Delta 1 - f_0 \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	$f_{0\text{med}}$			$\mu\Delta 1 - f_0$		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	85.69	82.42 – 88.96	<0.001	-0.32	-0.48 – -0.16	<0.001
Lang [Eng-L1]	3.05	-1.80 – 7.90	0.217	-0.06	-0.29 – 0.16	0.571
Lang [Eng-L2 (1)]	1.71	1.12 – 2.29	<0.001	0.00	-0.22 – 0.23	0.985
Lang [Eng-L2 (4)]	0.39	-0.15 – 0.93	0.153	0.16	-0.06 – 0.38	0.159

Fonte: elaborada pelo autor.

4.2.2 Pico de f_0 e f_0 mínimo

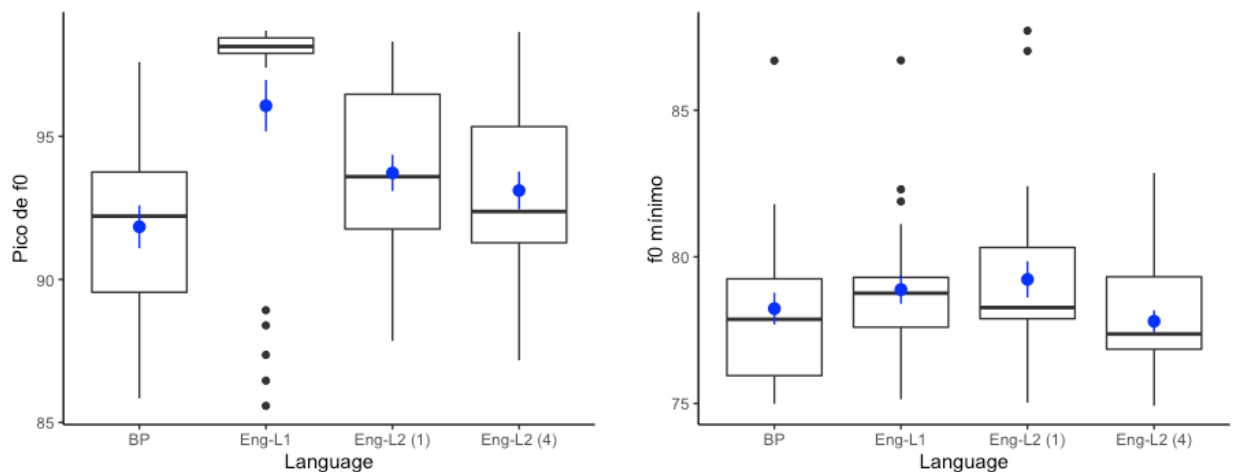
Média e mediana são medidas de tendência central que mensuram a altura relativa de f_0 . Já o cálculo de pico de f_0 (ou f_0 máximo) e f_0 mínimo podem fornecer informações acerca do escopo de variação de f_0 , evidenciando o quão complexo é a utilização desse parâmetro pelo falante. Em outras palavras, espera-se que falantes que possuem um pleno domínio dos aspectos

prosódicos de uma L1 ou L2 apresentem um interespaço mais amplo entre máximos e mínimos de f_0 , embora esse escopo de variação também seja linguisticamente determinado.

A frequência fundamental – f_0 é o correlato acústico da frequência de vibração das pregas vocais, cujo correlato perceptual é denominado *pitch*. Quando enunciamos, alternamos entre frequências mais altas e mais baixas de f_0 . Padrões de f_0 incidem sobre porções do enunciado, formando frases entonacionais. Geralmente o pico de f_0 incide sobre a sílaba tônica da palavra que recebe mais ênfase no enunciado, que se alterna com elementos menos proeminentes. Essa alternância pode contribuir para a percepção do ritmo, na medida em que as diferentes línguas do mundo mobilizam diferentes padrões melódicos a partir da variação de f_0 na realização fonética de enunciados.

Os *boxplots* da figura 29 apresentam a distribuição dos resultados dos quatro corpora deste estudo obtidos para o pico de f_0 e f_0 mínimo.

Figura 29 – *Boxplots* para médias de pico de f_0 (à esquerda, 29.1) e f_0 mínimo (à direita, 29.2) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



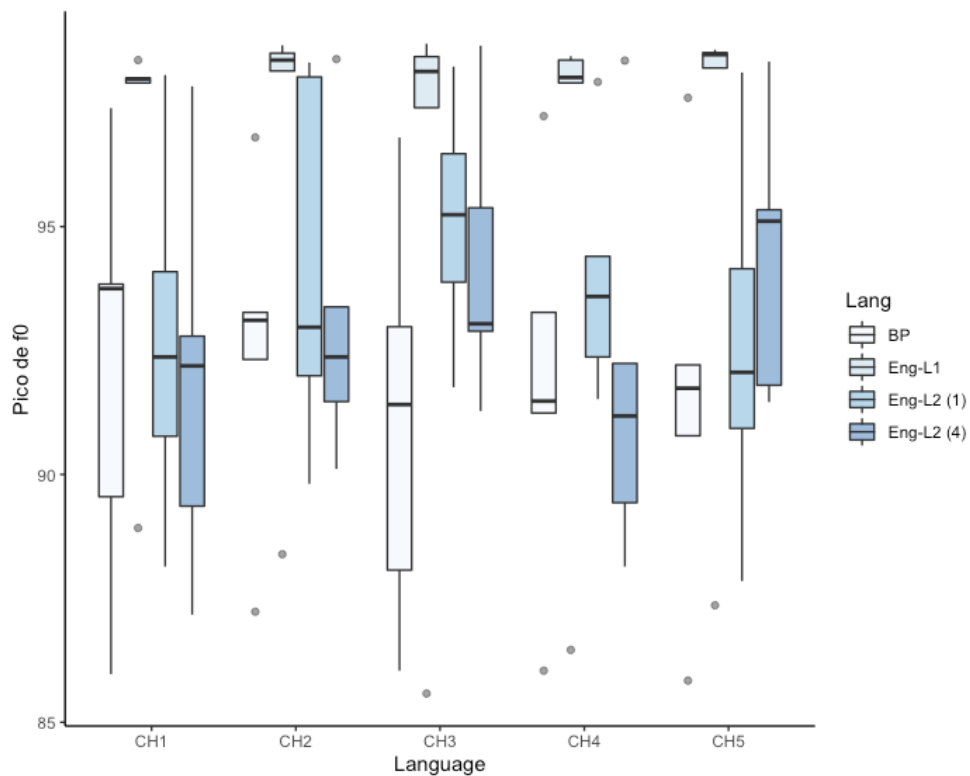
Fonte: elaborados pelo autor.

Em relação ao pico de f_0 , o inglês-L1 apresentou a maior média ($f_{0peak_{inglês-L1}} = 96,06$) e o PB a menor média ($f_{0peak_{PB}} = 91,83$). No entanto, essa tendência não foi corroborada pelo parâmetro de f_0 mínimo, uma vez que, conforme Silva Jr. e Barbosa (2019b), há uma relação direta entre f_0 mínimo mais alto e menor variabilidade de f_0 . Embora muito próximas, a média de f_0 mínimo do PB ($f_{0min_{PB}} = 78,23$) foi ligeiramente menor que a do inglês-L1 ($f_{0min_{inglês-L1}} = 78,88$). Já os resultados para as interlínguas sugerem um percurso de desenvolvimento compatível com a literatura, pois é esperado que os aprendizes apresentem

menor variabilidade e complexidade nos contornos de f_0 de suas produções orais em L2. Embora muito próximas, a média de f_{0min} do inglês-L2(1) superou a do inglês-L1(4) ($f_{0min_{inglês-L2(1)}} = 79,23 > f_{0min_{inglês-L2(4)}} = 77,80$). Neste sentido, considerando isoladamente as interlínguas e suas médias de f_0 mínimo, há indícios de maior variabilidade de f_0 entre a primeira e a quarta gravação. Para a média de f_{0peak} , os resultados foram menos elucidativos, e as médias do inglês-L2(1) e inglês-L2(4) foram praticamente idênticas e intermediários entre o PB e o inglês-L1 ($f_{0peak_{inglês-L2(1)}} = 93,72$); $f_{0peak_{inglês-L2(4)}} = 93,10$).

A fim de verificar a correlação entre picos de f_0 e f_0 mínimos, e estruturas sintáticas, os dados para essas duas métricas foram analisados por *chunks*, como mostram os *boxplots* da figura 30.

Figura 30 – *Boxplot* de pico de f_0 para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por *chunk*



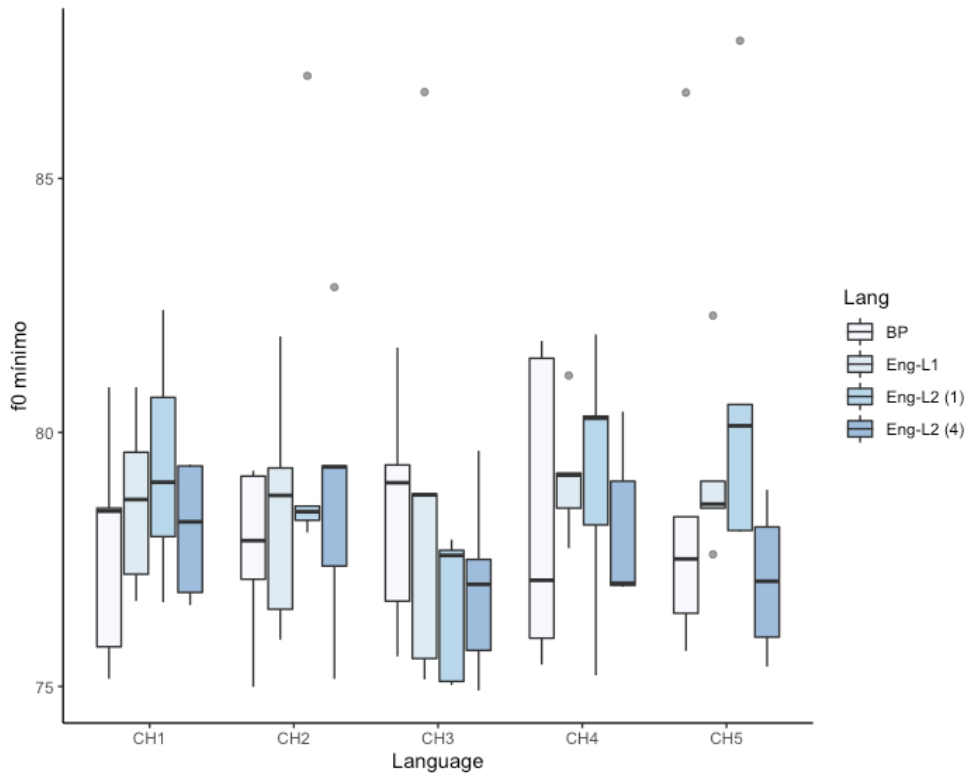
Fonte: elaborado pelo autor.

Em relação ao pico de f_0 , nota-se que o inglês-L1 atingiu as maiores médias em todos os *chunks*, tendo registrado as maiores médias de pico de f_0 nos dois primeiros (96,23 e 96,41 respectivamente), o que pode estar associado a um aumento da frequência fundamental na realização fonética das orações interrogativas fechadas. Esse parâmetro foi bem-sucedido na discriminação entre inglês-L1 e PB em todos os *chunks*. Quanto às interlínguas, houve grande sobreposição de resultados para o pico de f_0 , com o inglês-L2(1) atingindo as maiores médias,

exceto para o último *chunk*. Já os resultados para f_0 mínimo por *chunk* demonstram pouco efeito de língua, com grande convergência de valores em todos os *chunks*.

Em relação ao f_0 mínimo, cada *chunk* registrou uma hierarquia de valores distinta, com as (inter) línguas alternando-se em suas posições relativamente às demais. No entanto, percebe-se que o PB registrou as menores médias em quatro dos cinco *chunks*, embora a elevada sobreposição de valores não permita falar em significância quanto a esses resultados.

Figura 31 – *Boxplot de f_0 mínimo para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por chunk*



Fonte: elaborado pelo autor.

Tomados isoladamente, picos de f_0 e f_0 mínimos são pouco informativos em relação ao escopo de variação de f_0 do falante. Pesa também o fato de que, neste estudo, a distribuição da variável sexo é assimétrica entre os grupos controle e experimental, sendo 1 homem e 4 mulheres para o primeiro, e o contrário, 4 homens e 1 mulher para o último. Como há evidências de que o escopo de variação de f_0 é mais amplo para as mulheres do que para os homens (CUMBERS, 2013), a apresentação dos resultados apenas por (inter)língua poderia não ser tão realista. Dessa forma, decidimos calcular o escopo de variação de f_0 por falante a fim de verificar o comportamento individual dos participantes e a possível influência da variável sexo nos resultados. Para tanto, subtraímos f_0 mínimo de pico de f_0 , o que constitui também, segundo

Mennen, Schaeffler e Docherty (2012) uma medida distribucional de longo prazo (*LTD*) do tipo escopo.

Pellegrino (2006) assevera ainda que valores baixos de escopo de f_0 (ou intervalo tonal, utilizando-se da metalinguagem da autora) sinalizam uma fala monotônica, ao passo que um amplo escopo de f_0 revela um discurso mais variado e dinâmico. A tabela 25 apresenta os resultados para pico de f_0 (f_{0peak}), f_0 (f_{0min}) e escopo de f_0 ($f_{0peak} - f_{0min} = f_{orange}$) por falante. Os participantes do grupo experimental foram identificados por uma letra maiúscula e foi atribuída uma combinação de três letras aos que compõem o grupo controle. A participante N é a única do sexo feminino no grupo experimental e o participante Ros é o único do sexo masculino no grupo controle.

Tabela 25 – Médias de pico de f_0 e f_0 mínimo, e escopo de f_0 em semitons para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.

Língua	participante	f_{0peak}	f_{0min}	f_{orange}
PB	A	92,49	79,33	13,16
	F	90,63	75,52	15,11
	K	92,67	78,64	14,02
	L	86,22	77,69	8,52
	N	97,16	79,97	17,18
Eng-L1	Car	98,12	78,56	19,55
	Rho	98	77,19	20,8
	Ros	87,34	79,06	8,27
	Roz	98,54	78,77	19,76
	Van	98,33	80,82	17,5
Eng-L2 (1)	A	91,12	78,84	12,28
	F	92,88	77,19	15,68
	K	94,8	79,29	15,5
	L	91,72	78,34	13,37
	N	98,06	82,47	15,59
Eng-L2 (4)	A	91,76	78,95	12,81
	F	90,95	75,86	15,09
	K	92,39	77,96	14,43
	L	92,1	77	15,1
	N	98,3	79,24	19,06

Fonte: elaborada pelo autor.

De fato, é evidente a correlação entre f_0 e sexo. No grupo experimental, observa-se que a participante N, única do sexo feminino, é a que registra os maiores picos de f_0 , bem como os escopos mais amplos de f_0 para PB e inglês-L2(4). Já no grupo controle, há uma diferença considerável entre o único participante do sexo masculino e as demais participantes do sexo feminino em relação ao escopo de variação de f_0 , tendo o participante Ros apresentado um interespaço bem mais estreito de f_0 ($f_0\text{range}_{\text{Ros}} = 8,27$) em comparação ao menor valor de escopo de f_0 registrado entre as participantes do sexo feminino do grupo controle ($f_0\text{range}_{\text{Van}} = 17,5$) que representa pouco mais que o dobro do valor de Ros.

Na análise dos resultados do grupo experimental, houve variação quanto ao percurso individual dos aprendizes. Para A e N, o escopo de variação no PB ($f_0\text{range}_{\text{A-PB}} = 13,16$; $f_0\text{range}_{\text{N-PB}} = 17,18$) foi maior que os valores registrados para o mesmo índice no inglês-L2(1) ($f_0\text{range}_{\text{A-inglês-L2(1)}} = 12,28$; $f_0\text{range}_{\text{N-inglês-L2(1)}} = 15,59$). Os resultados para esses aprendizes dialogam com evidências de que falantes não nativos tendem a mobilizar um escopo menor de variação de f_0 na L2 em comparação à L1 (ALTENBERG; FERRAND, 2006; URBANI, 2012). Além disso, A e N, assim como L, apresentaram maior escopo de variação para o inglês-L2(4) em comparação ao inglês-L2(1), o que demonstra uma tendência de desenvolvimento rumo aos padrões prosódicos da L2. Para a participante N, o escopo de variação de f_0 para o inglês-L2(4) superou inclusive o valor verificado para o PB ($f_0\text{range}_{\text{N-PB}} = 17,18 < f_0\text{range}_{\text{N-inglês-L2(4)}} = 19,06$), agrupando-se com os valores aferidos para as participantes do sexo feminino do grupo controle.

Essa última constatação corrobora estudos que sugerem que influências culturais podem exercer maior pressão sobre o falar das mulheres, o que implica para o contexto de desenvolvimento de L2 que as mulheres em geral são mais propensas a replicar os padrões da L2 (URBANI, 2012). Por fim, F e K registram um maior escopo de variação para o inglês-L2(1) em comparação ao PB e ao inglês-L2(4), com diferenças pouco significativas entre as três gravações.

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada parâmetro acústico. O modelo para $f_0\text{peak}$ revelou efeitos significativos para inglês-L2 (1) e inglês-L2(4) e o modelo de $f_0\text{min}$ não apresentou efeitos significativos das (inter)línguas sobre os parâmetros acústicos.

Tabela 26 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para f_{0peak} e f_{0min} . Modelos: $f_{0peak} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $f_{0min} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	<i>f_{0peak}</i>			<i>f_{0min}</i>		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	92.05	88.94 – 95.16	<0.001	78.23	77.21 – 79.25	<0.001
Lang [Eng-L1]	4.02	-0.57 – 8.61	0.086	0.65	-0.79 – 2.10	0.372
Lang [Eng-L2 (1)]	1.64	0.51 – 2.77	0.004	1.00	-0.45 – 2.44	0.174
Lang [Eng-L2 (4)]	1.27	0.23 – 2.31	0.017	-0.43	-1.87 – 1.01	0.555

Fonte: elaborada pelo autor.

4.2.3 Desvio-padrão de f_0 (σf_0), assimetria (skewness) de f_0 e respectivas derivadas ($\sigma \Delta 1 - f_0$ e $sk \Delta 1 - f_0$)

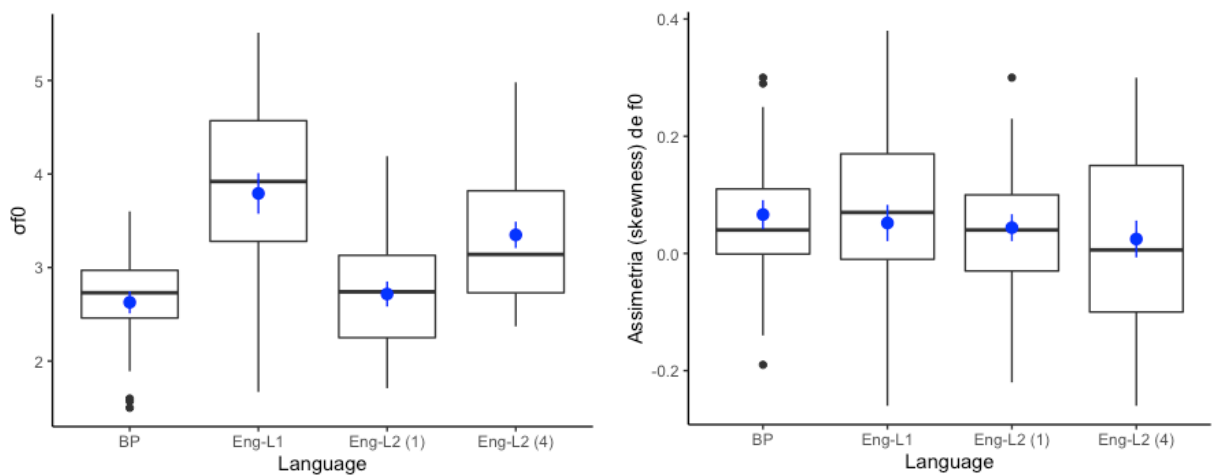
Como o desvio-padrão é uma medida de variabilidade que computa o nível de dispersão de um conjunto de dados, o cálculo do desvio-padrão de f_0 pode nos oferecer pistas do quão variável e complexa é a utilização desse parâmetro para os usuários de uma língua. Ainda segundo Cumbers (2013), o cálculo do desvio-padrão constitui uma medida mais robusta da frequência fundamental de um falante pelo fato de o desvio-padrão medir a distância média dos valores em relação à média, evitando que valores extremos distorçam a medida de variabilidade de f_0 . Já a assimetria mede o grau de afastamento da simetria de uma distribuição de frequência. Curvas perfeitamente simétricas ocorrem quando a média, a mediana e a moda coincidem. Quando essas medidas de tendência não coincidem, a distribuição será assimétrica, podendo ser positiva (*right skewed*), quando há concentração de valores na extremidade inferior da escala, com distribuição gradativa em direção à extremidade superior; ou negativa (*left skewed*), quando há concentração de valores na extremidade superior da escala, com distribuição gradativa em direção à extremidade inferior.

Conhecer o tipo e intensidade da assimetria de f_0 pode trazer informações úteis para a compreensão do ritmo, pois assimetria positiva acentuada significa que apesar da alta concentração de dados registrando valores menores, a média é influenciada pela cauda à direita e descola-se em sua direção. Desse modo, a quantidade de valores abaixo da média supera a quantidade de valores que a superam. Caso a assimetria seja negativa, o contrário ocorre. Como os valores de f_0 tendem a não apresentar distribuição normal ao redor da média, o cálculo de assimetria sinaliza o pico de distribuição de valores (URBANI, 2012).

Os *boxplots* da figura 32 apresentam a distribuição dos resultados dos quatro corpora deste estudo obtidos para o desvio-padrão (σ_{f_0}) e assimetria de f_0 . Como pode ser observado, o inglês-L1 apresentou o mais alto desvio-padrão dentre os corpora analisados ($\sigma_{f_0\text{inglês-L1}} = 3,79$), seguido pelo inglês-L2(4) ($\sigma_{f_0\text{inglês-L2(4)}} = 3,34$), inglês-L2(1) ($\sigma_{f_0\text{inglês-L2(1)}} = 2,71$) e PB ($\sigma_{f_0\text{PB}} = 2,62$). Os resultados para esse parâmetro sugerem um gradativo desenvolvimento prosódico dos aprendizes rumo aos padrões de variação de f_0 verificados na língua alvo.

No que concerne à assimetria de f_0 todos os corpora apresentaram assimetria positiva. Isso significa que a média de f_0 para todos os corpora sofreu a influência de valores mais extremos. O PB apresentou tendência de assimetria mais forte à direita ($f_0\text{sk}_{\text{PB}} = 0,06$), seguido do inglês-L1 ($f_0\text{sk}_{\text{inglês-L1}} = 0,05$) e das interlínguas ($f_0\text{sk}_{\text{inglês-L2(4)}} = 0,04$; $f_0\text{sk}_{\text{inglês-L2(1)}} = 0,02$). Os valores positivos indicam ainda a presença de picos acentuados e longas caudas na distribuição de frequência (MENNEN; SCHAEFFLER; DOCHERTY, 2012).

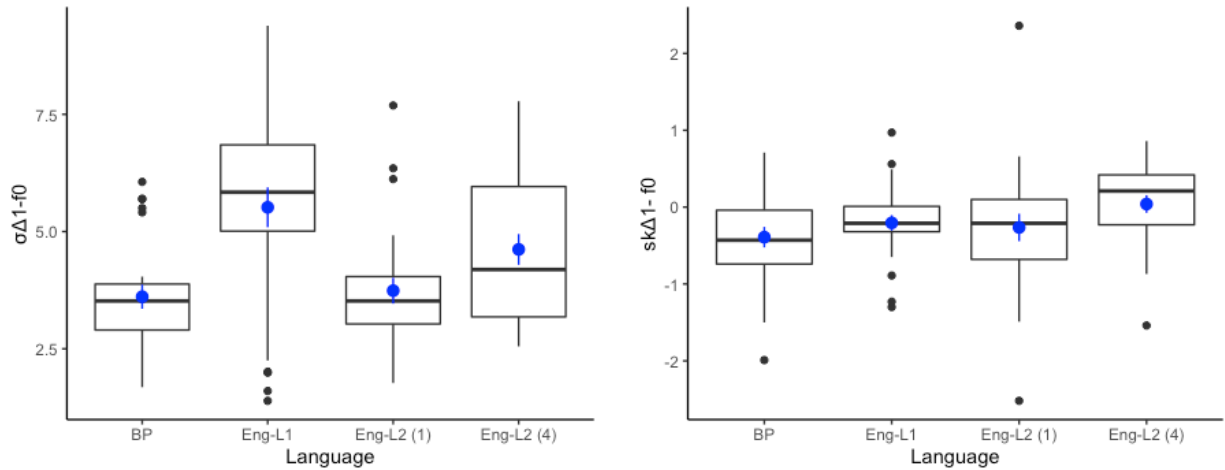
Figura 32 – *Boxplots* para médias de desvio-padrão de f_0 (à esquerda, 32.1) e assimetria de f_0 (à direita, 32.2) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

A primeira derivada das médias de desvio-padrão de f_0 ($\sigma\Delta 1-f_0$) e assimetria de f_0 ($\text{sk}\Delta 1-f_0$) também foram calculadas e a representação visual da distribuição desses resultados pode ser observada nos *boxplots* da figura 33.

Figura 33 – *Boxplots* para $\sigma\Delta 1-f_0$ (à esquerda, 33.1) e $sk\Delta 1-f_0$ (à direita, 33.2) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborados pelo autor.

Os padrões detectados para os parâmetros acústicos de desvio-padrão de f_0 e assimetria de f_0 foram reforçados pelos valores registrados para suas respectivas primeiras derivadas. A primeira derivada das médias de desvio-padrão de f_0 ($\sigma\Delta 1-f_0$) foi bem-sucedida na separação entre as L1s e também captou um percurso de desenvolvimento similar ao que foi flagrado pelo cálculo de desvio-padrão de f_0 , sendo mais elevada para o inglês-L1 ($\sigma\Delta 1-f_{0\text{inglês-L1}} = 5,51$), mais baixa para o PB ($\sigma\Delta 1-f_{0\text{PB}} = 3,61$) e apresentando valores intermediários para as interlínguas, com valores mais próximos ao inglês-L1 para o corpus da última gravação ($\sigma\Delta 1-f_{0\text{inglês-L2(1)}} = 3,73 < \sigma\Delta 1-f_{0\text{inglês-L2(4)}} = 4,61$). Quanto à primeira derivada da assimetria de f_0 ($sk\Delta 1-f_0$), com exceção do parâmetro do inglês-L2(4) ($sk\Delta 1-f_{0\text{inglês-L2(4)}} = 0,04$), as (inter) línguas registraram valores negativos, o que evidencia um coeficiente angular de tendência decrescente ($sk\Delta 1-f_{0\text{PB}} = -0,39$; $sk\Delta 1-f_{0\text{inglês-L1}} = -0,20$; $sk\Delta 1-f_{0\text{inglês-L2(1)}} = -0,26$).

Dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) foram ajustados, um para cada parâmetro acústico, analisando-se conjuntamente o desvio-padrão de f_0 (σf_0) e sua respectiva derivada ($\sigma\Delta 1-f_0$). Em ambos os modelos, houve efeitos significativos do inglês-L1 e do inglês-L2(4) sobre os parâmetros acústicos, conforme os dados apresentados na tabela 27.

Tabela 27 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para σf_0 e $\sigma \Delta 1-f_0$. Modelos: $\sigma f_0 \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $\sigma \Delta 1-f_0 \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	σf_0			$\sigma \Delta 1-f_0$		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	2.62	1.98 – 3.26	<0.001	3.61	2.96 – 4.26	<0.001
Lang [Eng-L1]	1.17	0.23 – 2.11	0.015	1.91	0.98 – 2.83	<0.001
Lang [Eng-L2 (1)]	0.09	-0.23 – 0.41	0.580	0.13	-0.79 – 1.05	0.782
Lang [Eng-L2 (4)]	0.72	0.42 – 1.02	<0.001	1.01	0.09 – 1.93	0.032

Fonte: elaborada pelo autor.

Também foram ajustados dois modelos de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para assimetria de f_0 (f_{0sk}) e sua respectiva derivada ($sk\Delta 1-f_0$), como mostra o resumo da tabela 28. Em nenhum dos modelos houve efeitos significativos das (inter)línguas sobre os parâmetros acústicos, à exceção do inglês-L2(4) em comparação ao PB para $sk\Delta 1-f_0$, o que corrobora a sobreposição visual verificada nos *bloxpots* e nos valores absolutos registrados para esses dois parâmetros acústicos.

Tabela 28 – Coeficientes, intervalos de confiança (95%) e valores de p para os dois modelos de regressão linear com efeitos mistos ajustados para f_{0sk} e $sk\Delta 1-f_0$. Modelos: $f_{0sk} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$ e $sk\Delta 1-f_0 \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	f_{0sk}			$sk\Delta 1-f_0$		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.06	-0.04 – 0.15	0.239	-0.39	-0.66 – -0.12	0.005
Lang [Eng-L1]	-0.00	-0.14 – 0.13	0.957	0.19	-0.19 – 0.57	0.335
Lang [Eng-L2 (1)]	-0.03	-0.09 – 0.04	0.416	0.13	-0.25 – 0.51	0.510
Lang [Eng-L2 (4)]	-0.04	-0.10 – 0.02	0.163	0.43	0.05 – 0.81	0.027

Fonte: elaborada pelo autor.

Considerando os efeitos verificados para a maioria das (inter)línguas sobre o desvio-padrão de f_0 (σf_0) e a primeira derivada do desvio-padrão de f_0 ($\sigma \Delta 1-f_0$), decidimos analisar os dados individuais dos participantes para esses dois parâmetros acústicos. A tabela 29 apresenta os resultados para esses parâmetros por indivíduo e corpora:

Tabela 29 – Médias de desvio-padrão de f_0 (σf_0) e primeira derivada do desvio-padrão de f_0 ($\sigma \Delta f_0$) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.

Língua	participante	σf_0	$\sigma \Delta f_0$
PB	A	2,82	3,52
	F	3,02	3,47
	K	2,97	3,58
	L	1,7	1,8
	N	2,62	5,67
Eng-L1	Car	3,83	5,7
	Rho	4,67	6,8
	Ros	2,01	1,84
	Roz	4,77	7,34
	Van	3,66	5,88
Eng-L2 (1)	A	2,83	3,77
	F	3,31	4,09
	K	2,68	3,41
	L	2,05	2,57
	N	2,69	4,84
Eng-L2 (4)	A	3,12	3,77
	F	3,46	4,19
	K	2,96	3,55
	L	2,92	4,52
	N	4,27	7,05

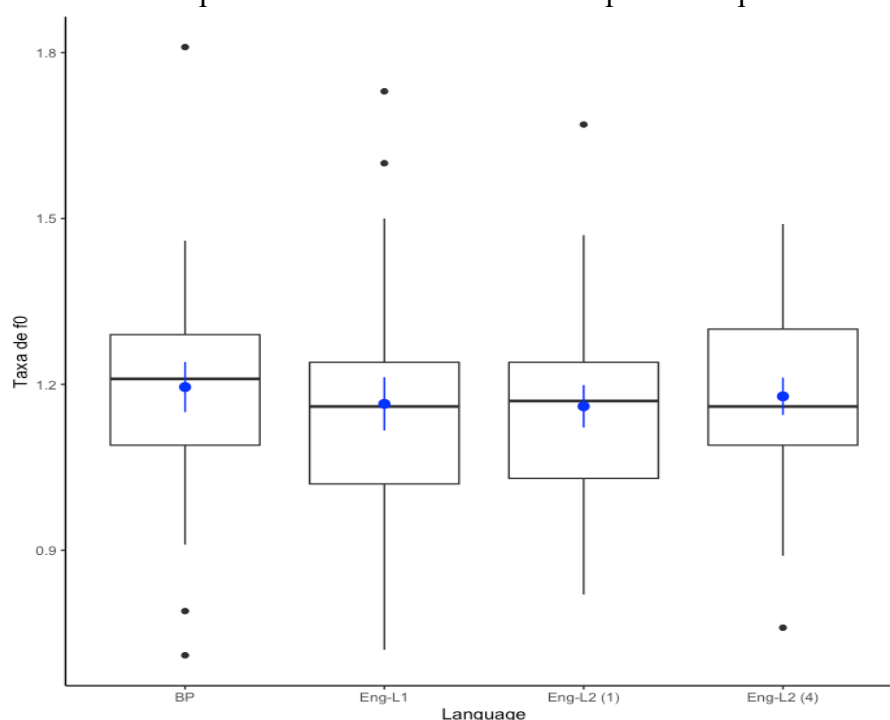
Fonte: elaborada pelo autor.

Como pode ser verificado, na análise do percurso individual dos aprendizes, todos registram médias de desvio-padrão de f_0 na primeira gravação muito próximas às médias do PB. Já na última gravação, o desvio-padrão de f_0 foi mais elevado para todos os aprendizes, o que sugere maior variabilidade e complexidade dos contornos de f_0 e um possível efeito do ensino explícito de pronúncia. Um padrão parecido pode ser verificado para a primeira derivada do desvio-padrão de f_0 ($\sigma \Delta f_0$), ainda que as médias tenham permanecido inalteradas para o participante A e discretamente mais elevadas para os falantes F e K entre a primeira e a última gravação. Esses resultados também corroboraram a influência da variável sexo já atestada para o escopo de f_0 . Conforme a tabela 15, dentre os falantes do grupo controle, o único do sexo masculino, Ros, apresentou médias de desvio-padrão de f_0 e primeira derivada de f_0 muito mais baixas que as demais participantes do sexo feminino. No grupo experimental, por outro lado, a única falante do sexo feminino foi a que registrou valores mais altos para esses parâmetros na última gravação, o que fortalece a hipótese que as mulheres são mais propensas a incorporar os padrões prosódicos da L2.

4.2.4 Taxa de f_0 (f_0 -R)

Valores mais elevados para a taxa de f_0 podem ser indicativos de maior variação de f_0 na linha do tempo, e, conseqüentemente, de padrões de uso mais complexos deste correlato de proeminência pelos usuários do sistema linguístico. O gráfico da figura 34 apresenta a distribuição dos resultados dos quatro corpora deste estudo para a taxa de f_0 .

Figura 34 – *Bloxpot* da taxa de f_0 para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborado pelo autor.

Os resultados para esse parâmetro foram pouco elucidativos e as médias registradas ocuparam posições muito próximas. O PB e o inglês-L2(4) apresentaram a maior média (f_0 -R_{PB} = 1,195; f_0 -R_{inglês-L2(4)}} = 1,195), seguidos pelo inglês-L1 (f_0 -R_{inglês-L1}} = 1,164), que registrou uma média praticamente idêntica ao inglês-L2(1) (f_0 -R_{inglês-L2(1)}} = 1,160). Esses resultados tomados isoladamente sugerem que não há diferenças significativas quanto à mobilização de padrões de f_0 entre os usuários dessas (inter)línguas e são pouco informativos quanto ao percurso de desenvolvimento dos aprendizes. A análise dos resultados individuais também não revelou diferenças interindividuais significativas, assim como o modelo de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para assimetria de taxa de f_0 , que não

apresentou efeito significativo das (inter)línguas sobre o parâmetro acústico analisado, conforme mostra o resumo da tabela 30.

Tabela 30 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para f_0 -R. Modelo: f_0 -R \sim Lang + (1|Chunk) + (1|Speaker)

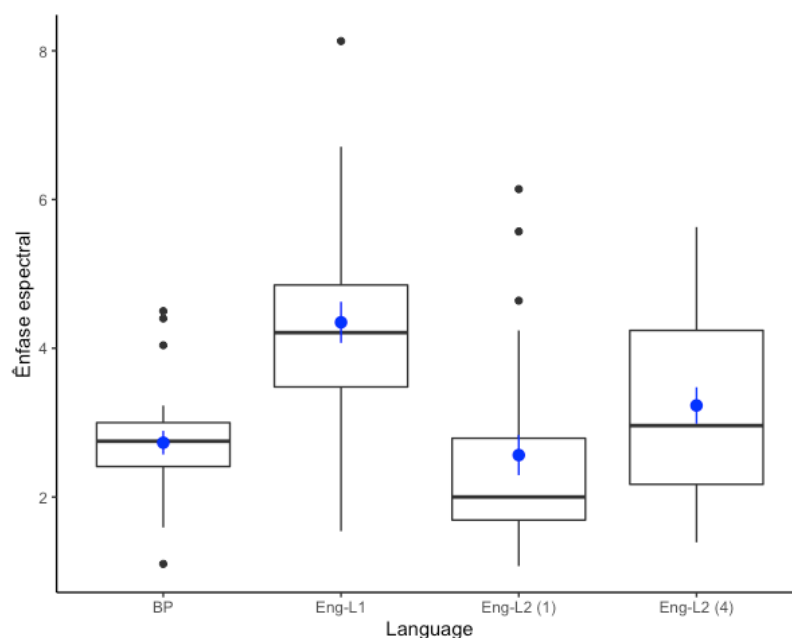
<i>Predictors</i>	<i>f₀-R</i>		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.19	1.08 – 1.29	<0.001
Lang [Eng-L1]	-0.02	-0.16 – 0.11	0.726
Lang [Eng-L2 (1)]	-0.04	-0.15 – 0.07	0.513
Lang [Eng-L2 (4)]	-0.02	-0.13 – 0.09	0.762

Fonte: elaborada pelo autor.

4.2.5 Ênfase espectral

A ênfase espectral está relacionada ao correlato de proeminência da intensidade e, portanto, à quantidade de energia da onda sonora e ao esforço vocal empreendido na realização fonética dos enunciados (GONÇALVES, 2018). O gráfico da figura 35 apresenta a distribuição dos resultados dos quatro corpora deste estudo obtidos para ênfase espectral:

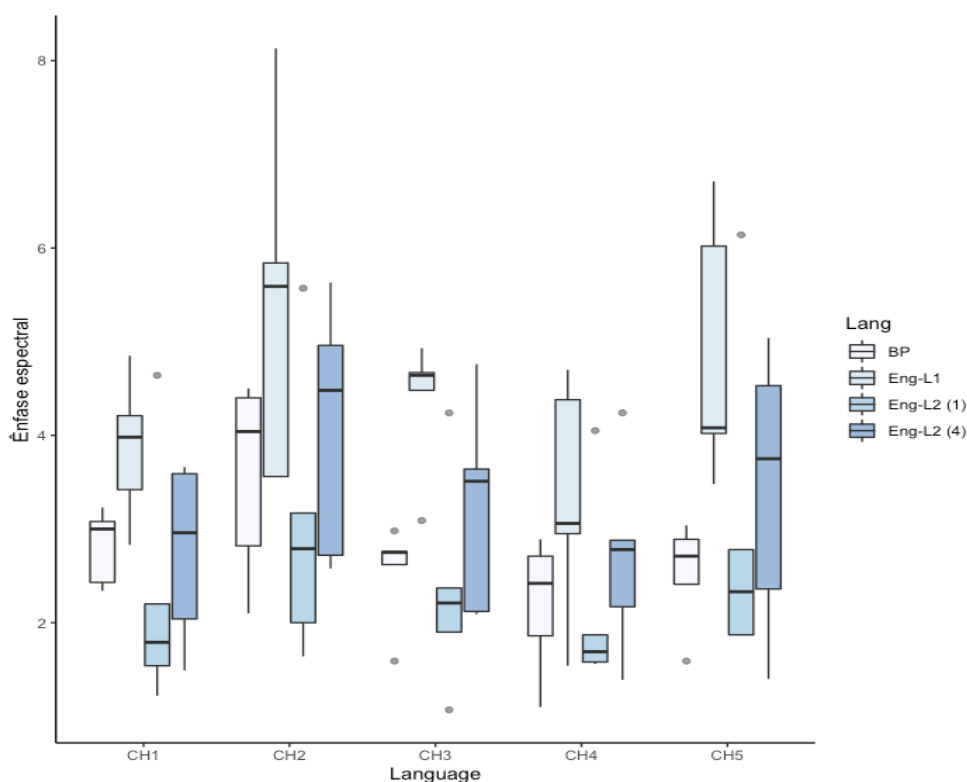
Figura 35 – *Bloxpot* da ênfase espectral (emph) para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborado pelo autor.

Como visualmente explicitado na figura 35 a ênfase espectral foi capaz de separar as L1s, com o inglês-L1 registrando a maior média dentre os corpora analisados ($emph_{\text{inglês-L1}} = 4,34$), superior ao PB ($emph_{\text{PB}} = 2,73$). Se considerarmos trabalhos que evidenciam a correlação entre ênfase espectral e acento frasal (HELDNER, 2001), esse resultado sugere que os falantes nativos do inglês realizam mais esforço como pista acústica na marcação do acento do que os falantes do português. Relativamente às interlínguas, o inglês-L2(1) obteve a menor média de ênfase espectral, muito próxima ao PB, ($emph_{\text{inglês-L2(1)}} = 2,56$), e o inglês-L2(4) aproximou-se do inglês-L1 ($emph_{\text{inglês-L2(4)}} = 3,23$). Isto significa que o esforço vocal empreendido pelos aprendizes na primeira gravação do texto em inglês foi mais similar ao PB, indicando transferência dos padrões de ênfase espectral da L1, e aproximou-se dos padrões prosódicos do inglês-L1 na última gravação. Esses resultados também sugerem uma correlação entre ensino explícito de pronúncia e desenvolvimento do ritmo, pois pode estar relacionado a uma melhora global dos aprendizes na marcação de sílabas acentuadas. Para fins de verificação da possível correlação entre intensidade e fronteiras sintáticas, as médias de ênfase espectral foram analisadas por *chunks*, conforme *boxplot* da figura 36:

Figura 36 – Ênfase espectral para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB por *chunk*



Fonte: elaborado pelo autor.

Como pode ser observado, o padrão detectado para a média geral de ênfase espectral foi replicado para todos os chunks ($\text{emph}_{\text{inglês-L1}} > \text{emph}_{\text{inglês-L2(4)}} > \text{emph}_{\text{inglês-PB}} > \text{emph}_{\text{inglês-L2(1)}}$), o que indica que essas línguas mobilizam diferentes padrões de intensidade. Nota-se também que o pico de ênfase espectral para o inglês-L1 coincide com o pico de mediana para o f_0 , o que sugere uma possível correlação entre esses dois correlatos de proeminência para o inglês.

Também foi ajustado um modelo de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para ênfase espectral. Houve efeitos significativos das (inter)línguas sobre o parâmetro acústico da ênfase espectral, à exceção do inglês-L2(1).

Tabela 31 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para ênfase espectral. Modelo: $\text{emph} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	emph		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	2.72	1.80 – 3.65	<0.001
Lang [Eng-L1]	1.62	0.44 – 2.81	0.007
Lang [Eng-L2 (1)]	-0.37	-0.82 – 0.09	0.115
Lang [Eng-L2 (4)]	0.50	0.08 – 0.92	0.020

Fonte: elaborada pelo autor.

Considerando os efeitos verificados para a maioria das (inter)línguas sobre a ênfase espectral, decidimos analisar os dados individuais dos participantes para esse parâmetro acústico. A tabela 32 apresenta os resultados para esses parâmetros por indivíduo e corpora. Como pode ser verificado, na análise do percurso individual dos aprendizes, houve grande variabilidade de trajetórias entre os aprendizes. Para F, L e N, as médias de ênfase espectral do inglês-L2(1) foram mais baixas que as registradas para o PB e foram superadas pelas médias do inglês-L2(4), tendo a média dos participantes L e N superado também a do PB na última gravação. Para A e K, no entanto, a média de ênfase espectral do inglês-L2(1) superou tanto a média do PB quanto a média do inglês-L2(4). Em geral, à exceção do participante F, os aprendizes apresentaram níveis de ênfase espectral mais elevados que o PB na última gravação.

Tabela 32 – Médias de ênfase espectral (emph) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.

Língua	participante	emph
PB	A	3,29
	F	2,88
	K	1,76
	L	2,54
	N	3,16
Eng-L1	Car	4,27
	Rho	3,72
	Ros	5,05
	Roz	2,9
	Van	5,8
Eng-L2 (1)	A	4,92
	F	1,83
	K	2,32
	L	1,47
	N	2,26
Eng-L2 (4)	A	4,08
	F	2,09
	K	1,98
	L	3,98
	N	4,01

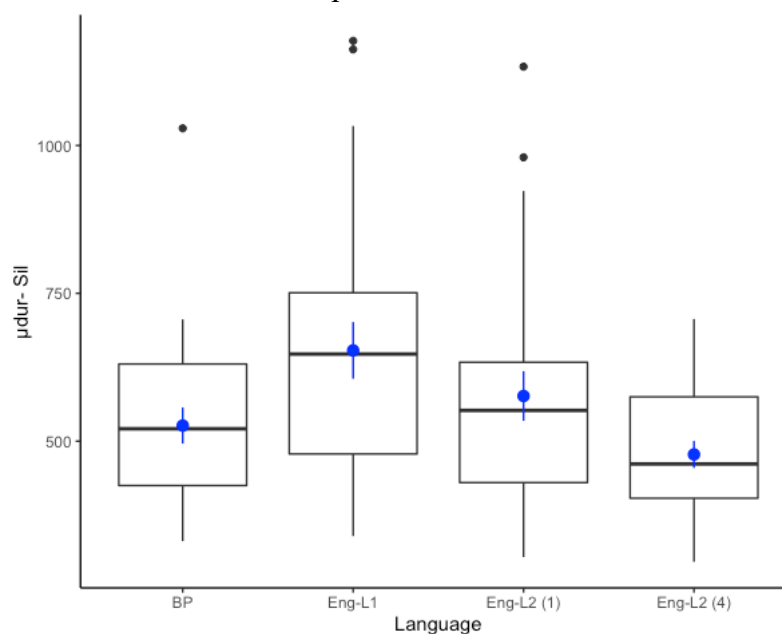
Fonte: elaborada pelo autor.

4.2.6 Duração média dos picos silábicos (μ dur- Sil)

A sonoridade é uma propriedade gradiente e corresponde à percepção de altura de um determinado segmento em comparação a outros produzidos com o mesmo grau de duração, tonicidade e *pitch*. Como as vogais possuem mais sonoridade, o pico silábico normalmente coincide com a sonoridade máxima dos segmentos vocálicos ocupando o núcleo silábico. Medidas de sonoridade podem contribuir para a compreensão do ritmo porque os usuários de uma língua natural podem utilizar-se dessa pista acústica para segmentar a cadeia da fala, auxiliando-os, por exemplo, no reconhecimento de sílabas (LADEFOGED; JOHNSON, 2011).

O *bloxpot* da figura 37 apresenta a distribuição das durações médias de picos silábicos (μ dur- Sil) para os quatro corpora analisados neste estudo:

Figura 37 – Duração média dos picos silábicos ($\mu\text{dur-Sil}$) para Inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborado pelo autor.

O inglês-L1 registrou maior duração média de picos silábicos ($\mu\text{dur-Sil}_{\text{inglês-L1}} = 653,47$) em comparação ao PB ($\mu\text{dur-Sil}_{\text{PB}} = 526,36$). Quanto às interlínguas, houve maior aproximação da média do inglês-L2(1) com o inglês-L1 ($\mu\text{dur-Sil}_{\text{inglês-L2(1)}} = 576,36$), e o inglês-L2(4) registrou a menor duração média de picos silábicos dentre os corpora analisados ($\mu\text{dur-Sil}_{\text{inglês-L2(4)}} = 477,55$). Esses resultados são pouco elucidativos em relação ao desenvolvimento do ritmo pelos aprendizes. Também foi ajustado um modelo de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para a duração média de picos silábicos (μdur). Não houve efeitos significativos das (inter) línguas sobre esse parâmetro acústico.

Tabela 33 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para ênfase espectral. Modelo: $\text{emph} \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

<i>Predictors</i>	$\mu\text{dur-Sil}$		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	512.30	391.97 – 632.62	<0.001
Lang [Eng-L1]	141.17	-3.20 – 285.55	0.055
Lang [Eng-L2 (1)]	70.20	-20.50 – 160.90	0.129
Lang [Eng-L2 (4)]	-48.84	-134.38 – 36.69	0.263

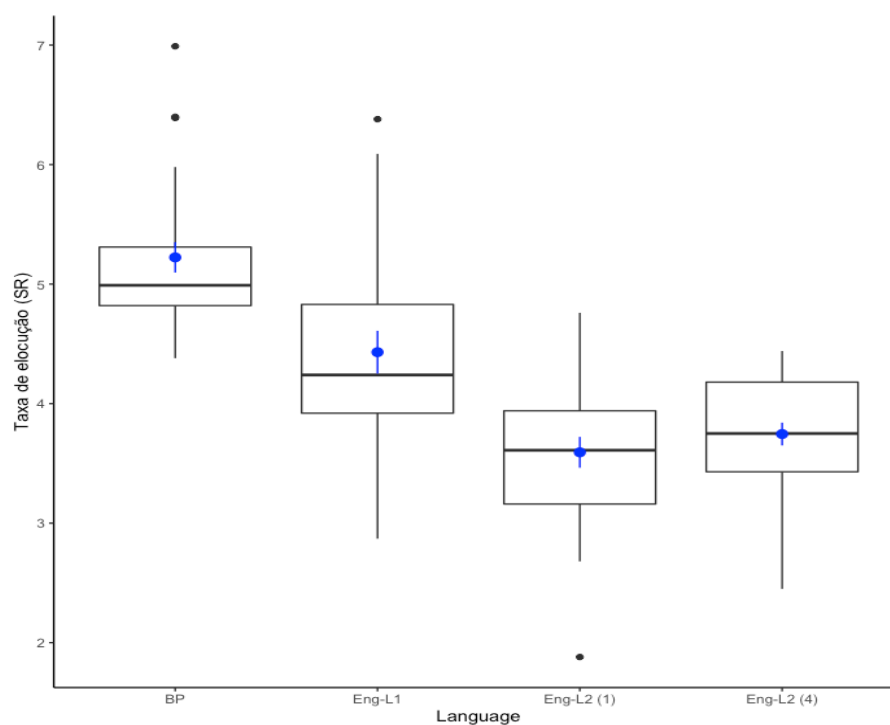
Fonte: elaborada pelo autor.

4.2.7 Taxa de elocução (SR)

A velocidade articulatória pode ser medida por meio da taxa de elocução e pode ser obtida pelo número de unidades de referência produzidas pela soma total de suas durações. Um fator que pode estar correlacionado à taxa de elocução é a fluência, e é esperado que aprendizes em estágios iniciais de desenvolvimento realizem foneticamente a língua alvo a uma taxa de elocução mais baixa devido às dificuldades articulatórias inerentes ao processo de aprendizagem de uma L2.

A figura 38 apresenta a distribuição das médias da taxa de elocução para os quatro corpora analisados neste trabalho:

Figura 38 – Taxa de elocução (SR) para inglês-L1, inglês-L2 (1), inglês-L2 (4) e PB. Os pontos e linhas azuis representam as médias e os erros-padrão respectivamente.



Fonte: elaborado pelo autor.

Como esperado, as L1s apresentaram taxas de elocução mais elevadas, com o PB registrando uma taxa de elocução mais alta em comparação ao inglês-L1 ($SR_{PB} = 5,22 > SR_{\text{inglês-L1}} = 4,43$). Além disso, o inglês-L2 (1) apresentou a menor taxa de elocução dentre os corpora analisados ($SR_{\text{inglês-L2(1)}} = 3,59$) e o inglês-L2(4) aproximou-se do inglês-L1 ($SR_{\text{inglês-L2(4)}} = 3,74$). O aumento verificado na taxa de elocução das interlínguas entre a primeira e a segunda gravação

pode estar relacionado aos efeitos da instrução explícita, especialmente do curso de Fonologia Suprasegmental da Língua Inglesa.

A fim de verificar a significância estatística desses resultados, foi ajustado um modelo de regressão de efeitos mistos (com *intercepts* aleatórios para participantes e *chunks*) para a taxa de elocução (SR). O modelo para a taxa de elocução revelou efeitos significativos de cada uma das (inter)línguas sobre as métricas, como pode ser observado no resumo da tabela 34, o que evidencia a robustez desse parâmetro acústico na distinção entre (inter)línguas e dialoga com os resultados reportados por Silva Jr. e Barbosa (2019b) para esse parâmetro.

Tabela 34 – Coeficientes, intervalos de confiança (95%) e valores de p para um modelo de regressão linear com efeitos mistos ajustados para taxa de elocução. Modelo: $SR \sim \text{Lang} + (1|\text{Chunk}) + (1|\text{Speaker})$

Taxa de Elocução (SR)			
<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	5.22	4.68 – 5.77	<0.001
Lang [Eng-L1]	-0.79	-1.46 – -0.12	0.020
Lang [Eng-L2 (1)]	-1.62	-1.90 – -1.34	<0.001
Lang [Eng-L2 (4)]	-1.48	-1.74 – -1.22	<0.001

Fonte: elaborada pelo autor.

Considerando que todas as (inter)línguas revelaram efeitos sobre a taxa de elocução, decidimos analisar os dados individuais dos participantes para esse parâmetro acústico. A tabela 35 apresenta os resultados para esses parâmetros por indivíduo e corpora:

Tabela 35 – Médias da taxa de elocução (SR) para PB, inglês-L1, inglês-L2 (1) e inglês-L2 (4) por falante.

Língua	participante	SR
PB	A	5,65
	F	5,45
	K	5,2
	L	5
	N	4,8
Eng-L1	Car	3,93
	Rho	4,11
	Ros	5,86
	Roz	3,88
	Van	4,35

	A	3,62
	F	3,68
Eng-L2 (1)	K	3,60
	L	2,83
	N	4,21
<hr/>		
	A	3,63
	F	3,95
Eng-L2 (4)	K	3,85
	L	3,63
	N	3,65
<hr/> <hr/>		

Fonte: elaborada pelo autor.

Como pode ser verificado, há uma consistência em relação aos valores mais elevados de taxa de elocução entre as L1s, com o PB apresentando as médias mais elevadas e com valores bem agrupados. Para os aprendizes, as médias de taxa de elocução do inglês-L2(1) foram mais baixas que as registradas para o PB e houve um aumento discreto dessas médias no inglês-L2(4), para quatro dos cinco aprendizes.

5 DISCUSSÃO

Como anteriormente explicitado, três questões norteadoras guiaram a condução deste trabalho. A primeira delas diz respeito ao posicionamento do inglês-L1 norte-americano, do inglês-L2 de brasileiros, e do PB-L1 a partir dos resultados para as métricas rítmicas e parâmetros acústicos. Vejamos como os resultados para métricas rítmicas e parâmetros acústicos responderam a essa questão.

Em relação às métricas rítmicas, que neste trabalho se referem aos índices estatísticos cujo correlato de proeminência computado é a duração, as métricas posicionaram o PB, o inglês-L1, o inglês-L2 (1) e o inglês-L2 (4) como ritmicamente diferentes. Isso ficou evidente nos planos bidimensionais formados pelos pares clássicos %V e ΔC , VarcoV e VarcoC, e nPVI-V e rPVI-C. Além disso os *bloxpots* e a distribuição dos resultados para as demais métricas que atingiram significância estatística dão suporte à hipótese de que as (inter)línguas analisadas são sistemas ritmicamente distintos. Nesse sentido, também foram detectadas diferenças rítmicas entre o inglês-L2 dos falantes nas duas diferentes etapas do desenvolvimento, com valores muito mais aproximados entre inglês-L2 (4) e inglês-L1, conforme as distribuições nos planos rítmicos e as distâncias euclidianas, o que pode ser atribuído ao efeito da instrução. Consequentemente, as métricas rítmicas da primeira gravação dos aprendizes foram mais dissimilares ao inglês-L1, mas também não foram mais próximas ao PB, como esperado devido à influência da prosódia da L1 sobre a L2. As médias do inglês-L2 (1) registraram valores ainda mais dissimilares em relação ao PB, perfazendo um percurso de desenvolvimento diferente do esperado. Quando analisamos as 19 métricas duracionais que atingiram significância estatística, 14 agruparam as (inter)línguas com esse padrão de posicionamento na distribuição dos resultados.

No que concerne aos parâmetros acústicos, que neste trabalho tomam por base correlatos de proeminência outros que não a duração (f_0 , intensidade e sonoridade), ao considerarmos os resultados dos cinco índices que atingiram significância estatística, a saber, pico de f_0 (f_{0peak}), desvio-padrão de f_0 (σf_0), primeira derivada do desvio-padrão de f_0 ($\sigma \Delta 1-f_0$), ênfase espectral (*emph*) e taxa de elocução (*SR*), as (inter)línguas assumiram posições diferentes do padrão de posicionamento captado pelas métricas rítmicas. As L1s foram bem separadas pelos parâmetros acústicos σf_0 , $\sigma \Delta 1-f_0$, *emph* e *SR*, com o PB registrando sempre os valores mais baixos em comparação ao inglês-L1, que registrou os valores mais altos dentre todos os corpora, exceto para a taxa de elocução. Em relação às L2s, à exceção de f_{0peak} , cujos valores para ambas as interlínguas permaneceram praticamente inalterados, com o inglês-L2(4)

registrando valores marginalmente menores que o inglês-L2(1), os parâmetros acústicos flagraram um percurso de desenvolvimento rumo aos padrões prosódicos do inglês-L1, com o inglês-L2(1) mais próximo ao PB (considerando σf_0 , $\sigma \Delta 1-f_0$ e ênfase espectral), e o inglês-L2(4) mais próximo ao inglês-L1. Isso corrobora uma hipótese já antecipada em etapas preliminares deste estudo de que coexistem padrões ortogonais de desenvolvimento rítmico em função das diferentes dimensões de proeminência. Em outras palavras, embora os aprendizes possam ter mobilizado em uma primeira etapa de desenvolvimento padrões rítmicos mais extremos na dimensão da duração, esse comportamento não foi replicado para outras dimensões de proeminência, o que é compatível com a abordagem multidimensional, gradiente e não categórica para o ritmo adotada neste trabalho. Neste sentido, tanto as métricas rítmicas quanto os parâmetros acústicos situaram o PB, o inglês-L1, o inglês-L2(1) e o inglês-L2(4) como sistemas ritmicamente diferentes, mas o posicionamento das interlínguas variou em função das dimensões de proeminência consideradas.

Isso nos leva à segunda questão norteadora desta pesquisa, que concerne a influência do PB no desenvolvimento do inglês-L2 dos aprendizes. Na dimensão duracional, houve um percurso inesperado, com a interlíngua em estágio mais inicial apresentando valores mais extremos em direção ao ritmo mais acentual tanto no eixo das métricas vocálicas quanto das métricas consonantais, destoando de estudos que argumentam em favor da existência de um percurso universal de desenvolvimento do ritmo, que deveria desenvolver-se do mais silábico ao mais acentual, independentemente da L1 (BUNTA; INGRAM, 2007; LLÉO, RAKOW, KEHOE, 2007; ORDIN; POLYANSKAIA, 2015). No entanto, também não é possível falar de forma objetiva em transferência de padrões prosódicos da L1 para a L2 na dimensão da duração, porque os valores registrados para o inglês-L2 (1) apresentam valores diametralmente opostos ao PB. Esse percurso de desenvolvimento é consistente com a definição de interlíngua que se apresenta como sistema relativamente independente da L1 e da L2 (LI; POST, 2014), e com a não linearidade do processo de desenvolvimento linguístico como processo dinâmico (DE BOT; LOWIE; VERSPOOR, 2007; DE BOT, 2008).

Entretanto, isso também não quer dizer que as produções de aprendizes brasileiros iniciantes serão sempre mais acentuais em termos duracionais. Vale ressaltar que o parâmetro utilizado neste estudo para caracterizar o estágio de desenvolvimento dos aprendizes foi semestre na graduação em Letras-Inglês, o que não necessariamente indica um nível de proficiência mais inicial. É possível que esses aprendizes tenham de fato atravessado um estágio em que a realização rítmica do inglês-L2 apresentou padrões duracionais com menor variabilidade vocálica e consonantal (i.e., com maior tendência ao ritmo silábico), pela

influência dos padrões prosódicos da L1. Além disso, pelo fato de estarem cursando, à época, uma graduação em Letras, e no esforço de realizarem foneticamente o inglês-L2 de maneira marcadamente distinta em relação ao PB, os aprendizes podem ter mobilizado um processo de dissimilação de categorias fonéticas, imprimindo valores duracionais exagerados para manter a distinção entre L1 e L2, de maneira semelhante ao que é previsto no Modelo de Aprendizagem de Fala para o nível segmental (FLEGE, 1995, FLEGE; BOHN, 2021). À medida que os aprendizes avançaram no desenvolvimento da L2, e após terem cursado as disciplinas de Fonologia Segmental e Fonologia Suprasegmental da Língua Inglesa, os padrões temporais da realização fonética da interlíngua podem ter sido reorganizados de maneira mais próxima à produção do inglês pelos falantes nativos.

A influência dos padrões prosódicos do PB manifestou-se de forma mais evidente na dimensão de f_0 , tendo em vista que quase todos os parâmetros acústicos que atingiram significância estatística nesse nível revelaram uma proximidade maior entre PB e inglês-L2(1), cujos valores médios foram os mais baixos dentre os corpora analisados. Por exemplo, a análise individual das médias de desvio-padrão de f_0 e primeira derivada do desvio-padrão de f_0 revelou níveis de variabilidade e complexidade dos contornos de f_0 bastante similares entre PB e inglês-L2(1). Esses resultados dialogam com estudos que sugerem a influência da L1 sobre os padrões de f_0 da L2. Silva Jr. e Barbosa (2019b) detectaram menor variabilidade na trajetória de variação melódica de falantes brasileiros de inglês-L2, com contornos de f_0 mais baixos e menos complexos na comparação com falantes nativos de AmE. Segundo esses autores, isso se deve a uma maior tendência por parte dos falantes não nativos em atentar mais à informação segmental que aos aspectos prosódicos, gerando padrões rítmicos e entonacionais destoantes dos padrões nativos. Já Urbani (2012), em sua pesquisa com aprendizes italianos de inglês-L2, não conseguiu chegar a uma conclusão sobre a possível transferência dos padrões de f_0 da L1 para a L2, devido a diferença verificada entre homens e mulheres do grupo experimental: embora os dados dos indivíduos do sexo masculino tenham mostrado que a L1 tinha uma influência na gama de contornos de f_0 na L2, os dados das mulheres não o fizeram. Segunda a autora, a hipótese de transferência deveria ser confirmada caso os falantes não nativos apresentassem um escopo de variação de f_0 similar à L1. Como a autora realizou um estudo transversal, diferentes estados de língua não puderam ser comparados.

Nossos dados vão na direção de confirmação da hipótese de transferência prosódica na dimensão de f_0 , tendo em vista que o escopo de variação dos aprendizes na realização fonética do PB e do inglês-L2(1) ficaram muito próximos, à exceção do falante L, cujo escopo de f_0 registrado para o PB ficou bem abaixo do registrado para o inglês-L2(1). Além disso, como os

dados dos indivíduos do sexo masculino sofreram poucas alterações na última gravação, e a única participante do sexo feminino registrou um aumento do escopo de variação de f_0 , atingindo uma média análoga às registradas pelas participantes do sexo feminino do grupo controle, arriscamos sugerir provisoriamente que a interferência dos padrões de f_0 da L1 sobre a L2 existe e tende a ser mais persistente e duradoura nos indivíduos do sexo masculino. Isso explicaria as diferenças com base na variável sexo encontradas por Urbani (2012). No entanto, devido à natureza limitada deste estudo no tocante ao número de participantes por grupo (apenas uma mulher e quatro homens no grupo experimental), mais estudos longitudinais com um maior número de participantes são necessários para que se verifique se esse comportamento se manifesta de forma consistente em grupos mais numerosos.

A terceira e última questão norteadora deste estudo diz respeito aos efeitos do ensino explícito de pronúncia. Os percursos de desenvolvimento do ritmo da L2 flagrados tanto pelas métricas quanto pelos parâmetros acústicos nesta análise sugerem que a instrução explícita pode ter impactado positivamente o desenvolvimento prosódico dos aprendizes, uma vez que os dados do inglês-L2 (4) foram coletados após os falantes terem cursado as disciplinas Fonologia Segmental da Língua Inglesa e Fonologia Suprasegmental da Língua Inglesa.

Na dimensão duracional, foi detectada uma maior organização temporal pelos aprendizes registrada pelas métricas dos dados do inglês-L2 (4) e uma aproximação das distâncias euclidianas entre inglês-L1 e inglês-L2(4) nos planos cartesianos formados pelos pares %V- Δ C, VarcoV-VarcoC e nPVI-V-rPVI-C.

Na dimensão de f_0 , os parâmetros acústicos flagraram maior complexidade nas curvas melódicas manifestadas na realização fonética da L2 pelos aprendizes. Além disso, na análise individual de escopo de f_0 , para três dos cinco aprendizes, foi detectada uma ampliação do interespaço de f_0 na última gravação em comparação à primeira, embora mais discreta para os homens. Na mesma direção, a análise individual do desvio-padrão de f_0 mostrou que todos os aprendizes manifestaram maior variabilidade nos contornos de f_0 na última gravação no comparativo com a primeira gravação e com o PB.

Na dimensão da intensidade, o aumento nas médias de ênfase espectral na última gravação demonstra um maior nível de esforço vocal/respiratório médio pelos aprendizes, o que pode estar relacionado a uma melhoria na habilidade de estabelecer contraste entre sílabas acentuadas e não acentuadas e/o ou na marcação do acento frasal, já que, segundo Heldner (2001), a ênfase espectral opera como pista acústica para a marcação do acento frasal em língua inglesa.

Finalmente, a taxa de elocução mais elevada para os dados do inglês-L2(4) sugere maior fluência dos aprendizes, que pode ter sido possível a partir do deslocamento da atenção de aspectos segmentais (supostamente mais automatizados nesta etapa de desenvolvimento) para aspectos prosódicos. O aumento verificado na taxa de elocução das interlínguas entre a primeira e a segunda gravação pode estar relacionado aos efeitos da instrução explícita, especialmente do curso de Fonologia Suprasegmental da Língua Inglesa. Portanto, é possível que a instrução explícita da pronúncia na aula de língua não nativa, com benefícios para o desenvolvimento de aspectos segmentais da pronúncia já reportados na literatura (LIMA JR., 2010; ALVES; MAGRO, 2011; LIMA JR.; ALVES, 2019), também exerça um efeito positivo no desenvolvimento de aspectos prosódicos em L2.

Importa discutir também os resultados obtidos pelas métricas rítmicas e parâmetros acústicos deste estudo e suas capacidades previsoras para padrões rítmicos similares e dissimilares entre (inter)línguas. Das 30 métricas rítmicas empregadas neste estudo, 19 revelaram efeitos de pelo menos 3 das (inter)línguas, e destas, 13 revelaram efeitos de todas as (inter)línguas: %V, %C, ΔC , ΔS , VarcoC, rPVI-C, rPVI-VC, nPVI-C, nPVI-VC, RR-C, RR-VC, VI-C e VI-VC. Partindo-se do pressuposto de que os quatro corpora analisados aqui têm ritmos distintos, a maioria das métricas rítmicas cumpriu o papel de posicionar cada (inter)língua em um local diferente do espaço gradiente de classificação rítmica. Dentre as treze métricas que revelaram efeitos significativos para todas as (inter)línguas, o padrão preponderante de posicionamento flagrado foi aquele em que o inglês-L2(1) apresentou os valores mais extremos na direção acentual, o PB os valores mais extremos na direção silábica, e o inglês -L1 e inglês-L2 (4) em maior proximidade, o que era previsto dado o desenvolvimento da interlíngua na quarta gravação (como em Lin e Wang [2005]).

Em geral, as métricas consonantais clássicas (ΔC , VarcoC e rPVI-C) apresentaram maior variabilidade e dispersão em comparação às suas contrapartes para outras unidades de referência. A maior dispersão e variabilidade dessas métricas consonantais pode estar associada a diferentes taxas de elocução, uma vez que duas das métricas consonantais empregadas não são normalizadas para a taxa de elocução, embora haja indícios na literatura de que esse fator não exerça tanta influência sobre o escopo duracional dos intervalos consonantais (LOW; GRABE; NOLAN, 2000). Além disso, não é possível afirmar que as métricas normalizadas sempre serão mais robustas em relação às métricas não normalizadas. Por exemplo, no estudo de Arvaniti (2012), %V, ΔC e rPVI-C foram as mais consistentes e apresentaram maior efeito de língua do que Varcos e nPVI-V. Outra possível explicação para a maior variabilidade temporal das métricas consonantais, sobretudo para os dados do inglês-L2 (1), é o fato de

termos analisado o parágrafo inicial do texto lido, com maior probabilidade de os participantes terem produzido os enunciados sob tensão e, conseqüentemente, com maior monitoramento da fala. Apesar dessa diferença em relação a estudos prévios, é evidente a coerência interna entre os dados do estudo, com o inglês-L1 e o PB ocupando posições claramente distintas nos planos rítmicos, como previsto pelas propostas, os dados do inglês-L2 (1) registrando valores mais extremos em ambos os eixos, e o inglês-L2 (4) posicionando-se mais próximo ao inglês-L1, o que foi constatado pela distribuição dos resultados, e visualmente, pelos planos rítmicos bidimensionais, pelos *bloxpots* e pelo cálculo das distâncias euclidianas.

A despeito da maior variabilidade, as consoantes representam neste estudo o intervalo de referência com maior número de métricas que revelaram efeitos das (inter)línguas, o que foi atestado pelos modelos de regressão de efeitos mistos ajustados para cada uma dessas métricas: apenas YARD-C não atingiu significância estatística. Neste sentido, este estudo dialoga com os resultados reportados por Arvaniti (2012), cujos resultados mais estáveis resultaram dos intervalos consonantais, e destoa de White e Mattys (2007) e Wiget et. al. (2010), para os quais diferenças interlinguísticas foram flagradas mais eficientemente pelas métricas vocálicas. Argumentamos consoante Arvaniti (2012) que uma possível explicação para a eficiência das métricas consonantais reside no fato de que a variabilidade na estrutura silábica, ou, em outras palavras, os tipos de agrupamentos consonantais permitidos por uma língua, reflete mais diretamente na duração, o que pode ser captado de forma relativamente eficiente pelas métricas consonantais. Outra unidade de referência que apresentou resultados robustos e capacidade distintiva entre as línguas foram as unidades vocálico-consonantais consecutivas, que só não apresentaram efeitos significativos de (inter)língua para VarcoVC. Em geral, a grande quantidade de métricas duracionais que revelaram resultados significativos destoa de estudos como Silva Jr. e Barbosa (2019b), para o qual poucas dessas métricas passaram nos pressupostos estatísticos para a realização de uma ANOVA paramétrica. Uma outra diferença em relação a Silva Jr. e Barbosa (2019b) é o fato de o segmento das unidades VV, as sílabas fonéticas, terem revelado, neste estudo, resultados significativos para apenas 3 das 5 métricas reportadas pelos autores como significativas: somente ΔS revelou efeitos para cada uma das (inter)línguas, ao passo que rPVI-S e VI-S revelaram efeitos apenas parciais. No entanto, esses resultados reforçam o potencial de utilização das unidades VV no processo de segmentação para investigação do ritmo.

A variabilidade dos resultados também foi verificada em relação aos estudos de referência. No geral, o PB apresentou valores mais próximos de línguas consideradas mais silábicas em relação ao ritmo. No entanto, houve inconsistência em relação a métricas como

nPVI-V, já que valores menores para esse índice são esperados para línguas de ritmo mais silábico, refletindo menor variabilidade entre os intervalos vocálicos consecutivos. Já o inglês-L1 apresentou médias consistentes em relação aos estudos de referência, embora com valores mais elevados para o eixo dos intervalos consonantais. É importante destacar que variabilidade de resultados entre o presente estudo e estudos de referência pode estar relacionada também às escolhas metodológicas empregadas na coleta de dados. Por exemplo, Ramus, Nespor e Mehler (1999) utilizaram-se apenas da leitura de frases. Grabe e Low (2002) coletaram os dados a partir da leitura de um texto. Silva Jr. e Barbosa (2019b) trabalharam com um número maior de informantes (20) e dois corpora. Finalmente, Arvaniti (2012), utilizou-se de três modos de enunciação distintos em sua análise: leitura de frases, leitura de texto e fala espontânea.

Em relação aos 12 parâmetros acústicos analisados, 5 revelaram efeitos para pelo menos três das interlínguas: pico de f_0 , desvio-padrão de f_0 , primeira derivada do desvio-padrão de f_0 , ênfase espectral e taxa de elocução. Dentre esses parâmetros acústicos, apenas a taxa de elocução revelou efeitos de todas as (inter)línguas. O fato de a taxa de elocução ter sido o parâmetro acústico que revelou mais efeitos significativos dialoga com os resultados reportados por Silva Jr. e Barbosa (2019b), para os quais a taxa de elocução se mostrou o parâmetro acústico mais consistente, revelando menor fluência por parte dos aprendizes, mesmo com a realização acurada no nível segmental. De acordo com Pellegrino (2012), no nível suprasegmental, a fluência, juntamente com um menor número de pausas silenciosas e um escopo amplo de f_0 contribuem para uma menor percepção do grau de acento estrangeiro. De fato, se compararmos a taxa de elocução e o escopo de f_0 obtidos neste estudo, parece haver uma correlação entre esses dois indicadores. Investigações no nível perceptual são necessárias para verificar se este é o caso.

Considerando que os quatro corpora ora analisados constituem sistemas rítmicamente distintos, os parâmetros acústicos também cumpriram o papel de posicionar cada (inter)língua em um local diferente do espaço rítmico gradiente. Dentre esses parâmetros que revelaram efeitos significativos para todas as (inter)línguas, o padrão preponderante de posicionamento flagrado foi aquele em que o inglês-L2(1) aproximou-se mais do PB (à exceção da taxa de elocução), dando suporte à hipótese de transferência da L1 para a L2, e o inglês-L2(4) aproximou-se do inglês-L1. Ainda que partindo de uma direção oposta e mais condizente com o que havia sido hipotetizado, os resultados dos parâmetros acústicos e das métricas rítmicas convergem na direção de flagrar um percurso de desenvolvimento. No entanto, a influência da L1 sobre a L2 mostrou-se mais evidente na dimensão de f_0 do que na dimensão duracional. Isso reforça a tese levantada neste estudo de que o desenvolvimento de uma

dimensão da prosódia em L2 é ortogonal em relação às demais dimensões. Se considerarmos a definição de trabalho adotada neste estudo para o ritmo, que o situa como função da distribuição de elementos de proeminência no sinal acústico, nossos dados dão suporte a uma perspectiva multidimensional de investigação do ritmo, como meio de se chegar a descrições mais refinadas e abrangentes acerca desse aspecto da linguagem.

Por fim, algumas breves considerações podem ser tecidas em relação ao ensino do componente fonético-fonológico no que diz respeito ao ritmo. Considerando as diferenças fonológicas entre o PB e o inglês, o ensino do ritmo em sala de aula poderia: (i) conscientizar os aprendizes acerca das diferenças duracionais entre sílabas acentuadas e não-acentuadas na realização fonética do inglês em relação ao português; (ii) enfatizar o processo de redução silábica no inglês, realizando o treinamento fonético para a percepção e produção de vogais reduzidas, sobretudo o *schwa*; (iii) chamar atenção para a inadequação da inserção de vogais empetéticas ao início de palavras iniciadas pela letras ‘s’ ou ao final de palavras terminadas em consoantes oclusivas ou ‘e’ silencioso; (iv) enfatizar a função distintiva do acento em língua inglesa, nos contextos semântico e pragmático; (v) promover a percepção de *pitch* em sua correlação com a estrutura sintático-discursiva e a produção de contornos entonacionais mais variados.

Essas propostas podem ser implementadas em sala de aula de acordo com os princípios da abordagem comunicativa aplicados ao ensino de pronúncia. De acordo com Celce-Murcia, Brinton e Goodwin (2010), o ensino do componente fonético-fonológico de base comunicativa possui cinco estágios: (i) descrição e análise; (ii) discriminação auditiva; (iii) prática controlada e feedback; (iv) prática guiada e feedback; e (v) prática comunicativa. No primeiro estágio, descrição e análise, os aprendizes são conduzidos à análise da linguagem, inferindo, analisando e discutindo sobre os padrões que sistematizam o sistema da L2. No segundo estágio, discriminação auditiva, os aprendizes são expostos aos sons da L2. Conforme assevera Lima Júnior e Alves (2019, p. 41), “a percepção desempenha um papel fundamental no desenvolvimento do sistema fonético-fonológico da L2” e “esta etapa é fundamental para o estabelecimento de categorias funcionais de distinções sonoras em L2”. Na terceira etapa, prática controlada e *feedback*, exercícios e práticas mecânicas são implementados com o objetivo de ajudar os aprendizes a compreenderem as características articulatórias dos sons de maneira automática. Na quarta etapa, prática guiada e feedback, os aprendizes devem resolver uma tarefa comunicativa de completar lacunas, que requer uma produção oral inteligível como pressuposto à adequada comunicação. Finalmente, na prática comunicativa, os aprendizes utilizam a língua livremente, sem vocabulário pré-selecionado ou aspecto linguístico

previamente definidos pelo professor. O uso do aspecto linguístico alvo da instrução explícita será um dos muitos aspectos que os alunos deverão mobilizar na resolução da tarefa comunicativa proposta.

Outrossim, salientamos, consoante Lima Jr. e Alves (2019) que o ensino do componente fonético-fonológico, por ser parte da língua em uso, deve ser implementado de forma integrada aos demais componentes linguísticos, tais como o sintático, o morfológico, o semântico e o pragmático, e convergir para tópicos e objetivos comunicativos previamente estabelecidos.

6 CONCLUSÃO

Respondendo à primeira questão que orientou este estudo, “*Como as métricas e parâmetros posicionam o inglês-L1 norte-americano, o inglês-L2 de brasileiros, e o PB-L1?*”, todas as métricas analisadas e parâmetros acústicos que atingiram significância estatística posicionaram o inglês-L1 norte-americano, o inglês-L2 de brasileiros, e o PB-L1 em posições distintas do espaço rítmico gradiente, confirmando a primeira hipótese, de são sistemas rítmicamente diferentes.

Já o percurso do desenvolvimento assumiu trajetórias diferentes em função das dimensões de proeminência consideradas. Na dimensão duracional, esse percurso partiu de uma direção diametralmente oposta à esperada, já que os valores do inglês-L2 (1) se distanciaram ainda mais daqueles do PB. Isso significa que em termos duracionais, nossos dados não apontaram para a transferência de padrões prosódicos do PB para o inglês-L2(1) dos aprendizes brasileiros na dimensão duracional. Esse percurso de desenvolvimento atípico corrobora a ideia de interlíngua (e língua) enquanto sistema dinâmico e não linear, pois os dados do inglês-L2 (1) apresentaram uma desorganização temporal, com altos níveis de variabilidade e dispersão, mas tendendo aos padrões rítmicos do inglês-L1, o que fica mais evidente na constatação de maior proximidade dos dados do inglês-L2 (4) com o inglês-L1 em todas as métricas verificadas. Em contrapartida, sobretudo na dimensão de f_0 , os valores do inglês-L2(1) foram mais próximos ao PB. Isso nos leva à segunda pergunta norteadora deste estudo – “*Qual é a influência do ritmo do PB-L1 no desenvolvimento do inglês-L2 dos aprendizes?*”. Nossos dados não revelaram transferência direta de padrões rítmicos do PB na dimensão duracional, como inicialmente hipotetizado. No entanto, mais investigações são necessárias, pois, neste estudo, os aprendizes do primeiro semestre não necessariamente apresentavam um perfil de proficiência iniciante, podendo ter mobilizado um processo de dissimilação fonética com o objetivo de realizar L1 e L2 de maneira marcadamente distinta. Por outro lado, na dimensão de f_0 , nossos dados apontam para a transferência dos padrões prosódicos da L1 para a L2, pelo menos quando se consideram dois dos três parâmetros acústicos com base nesse correlato de proeminência que atingiram significância estatística (σf_0 e $\sigma \Delta 1-f_0$).

Para a terceira e última pergunta, “*Qual é o efeito do ensino explícito de pronúncia no desenvolvimento do ritmo do inglês-L2 dos aprendizes?*”, há indícios de que o ensino explícito de pronúncia pode ter exercido um efeito positivo no processo de desenvolvimento do ritmo da L2, pois foi confirmada, tanto pelas métricas rítmicas quanto pelos parâmetros acústicos, a hipótese de que os dados do inglês-L2 (1) seriam mais dissimilares em relação ao

inglês-L1 em comparação aos dados do inglês-L2 (4). A menor diferença entre inglês-L1 e inglês-L2 (4), após os aprendizes terem cursado as disciplinas de Fonologia Segmental da Língua Inglesa e Fonologia Suprasegmental Língua Inglesa sugere também que o desenvolvimento da prosódia em aprendizes de L2 pode beneficiar-se do ensino explícito de pronúncia. É importante destacar que, devido à ausência do emprego de um grupo controle para esta variável (i.e., um grupo que não recebeu instrução explícita), a instrução explícita emerge neste estudo apenas como um dos possíveis fatores que podem ter interferido no percurso de desenvolvimento dos aprendizes.

Uma outra limitação deste estudo diz respeito à quantidade de sentenças analisadas, apenas sete. É possível que a análise de um número maior de sentenças no futuro resulte em mais estabilidade nos índices métricos (ARVANITI, 2012) e nos parâmetros acústicos. Assim, as análises futuras deverão incluir também os corpora das gravações 2 e 3, para uma abordagem mais longitudinal, além de incluir os outros dois parágrafos do texto lido. Além disso, o aspecto perceptual associado ao ritmo deverá ser explorado, por meio do emprego de juízes nativos para avaliação das gravações modificadas através de filtros passa-baixa, como meio de neutralizar a informação segmental. Dessa forma, será possível refinar a descrição rítmica das línguas e interlínguas de maneira multidimensional, e investigar de maneira mais pormenorizada o processo de desenvolvimento da prosódia do inglês-L2 por aprendizes brasileiros, verificando os efeitos da L1 e do ensino explícito de pronúncia nesse percurso em correlação com aspectos como inteligibilidade, compreensibilidade e grau de acento.

Os resultados obtidos neste estudo demonstram a utilidade das métricas rítmicas e parâmetros acústicos para a descrição do desenvolvimento do ritmo em L2, e a necessidade de se considerar múltiplas dimensões em investigações dessa natureza. Acreditamos que nas futuras etapas desta pesquisa, a expansão do material analisado e a inclusão da dimensão perceptual, com o emprego de juízes para avaliar o grau de inteligibilidade, compreensibilidade e acento associados às produções dos aprendizes, permitirá uma caracterização ainda mais precisa do desenvolvimento do ritmo por aprendizes brasileiros de L2, a partir de uma perspectiva multidimensional.

Finalmente, vale ressaltar que futuras investigações são necessárias para dar suporte aos achados preliminares a que essa pesquisa chegou. Com isso, será possível generalizar a coexistência de padrões ortogonais de desenvolvimento rítmico em função das diferentes dimensões de proeminência consideradas e a validade das métricas e parâmetros acústicos utilizados neste estudo.

REFERÊNCIAS

- ABAURRE-GNERRE, M. B. M. Processos fonológicos segmentais como índices de padrões prosódicos diversos os estilos formal e casual do Português do Brasil. **Cadernos de Estudos Lingüísticos**, Campinas, v. 2, p. 23-44, 1981. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8636610>. Acesso em: 12 jan. 2021.
- ABERCROMBIE, D. **Elements of general phonetics**. Chicago: Aldine, 1967.
- ACKERMANN, H; HERTRICH, I. Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllabic timing. **Folia Phoniatria et Logopaedica**, Basel, v. 46, n. 2, p. 70–78, jun. 2009.
- ADAMS, C. **English speech rhythm and the foreign learner**. Haia: Mouton, 1979.
- ALTENBERG, E. P.; FERRAND, C. T. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. **Journal of Voice**, Amsterdã, v. 20, n. 1, p. 89–96, mar. 2006. Disponível em: [doi:10.1016/j.jvoice.2005.01.005](https://doi.org/10.1016/j.jvoice.2005.01.005). Acesso em: 30. jul. 2021.
- ALVES, U. K.; MAGRO, V. Raising awareness of L2 phonology: explicit instruction and the acquisition of aspirated /p/ by Brazilian Portuguese speakers. **Letras de Hoje**, Porto Alegre, v. 46, n. 3, p. 71-80, set. 2011.
- ARVANITI, A. The usefulness of metrics in the quantification of speech rhythm. **Journal of Phonetics**, Amsterdã, v. 40, n. 3, p. 351–373, maio. 2012. Disponível em: <http://dx.doi.org/10.1016/j.wocn.2012.02.003>. Acesso em: 10 fev. 2021.
- AVERY, P.; EHRLICH, S. Common pronunciation problems. *In*: AVERY, P.; EHRLICH, S. **Teaching American English pronunciation**. Oxford: Oxford University Press, 2012. p. 96-109.
- BALTAZANI, M. Prosodic rhythm and the status of vowel reduction in Greek. *In*: TOPINTIZI, N.; LAVIDAS, N.; MOUMTZI, M. (ed.). **Selected papers on theoretical and applied linguistics from the 17th international symposium on theoretical & applied linguistics**. Thessaloniki: Aristotle University of Thessaloniki, 2007. p. 31–43. Disponível em: <https://www.enl.auth.gr/symposium/publication.html>. Acesso em: 16 abr. 2021.
- BARBOSA, P. A. Análise e modelamento dinâmicos da prosódia do português brasileiro. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 15, n. 2, p. 127-137, jul./dez. 2007.
- BARBOSA, P. A. **Incursões em torno do ritmo da fala**. Campinas: Pontes, 2006.
- BARBOSA, P.A. “Syllable-timing in Brazilian Portuguese”: uma crítica a Roy Major. **Documentação de Estudos em Lingüística Teórica e Aplicada**, São Paulo, v. 16, n. 2, p. 369–402, nov. 2000.

BARBOSA, P.A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 20, n. 1, p. 11-27, jan./jun. 2012.

BARBOSA, P.A.; BAILLY, G. Characterisation of rhythmic patterns for text-to-speech synthesis. **Speech Communication**, Amsterdã, v. 15, n. 1-2, p. 127-137, out. 1994.

BERTINETTO, P.M; BERTINI, C. On modeling the rhythm of natural languages. *In*: SPEECH PROSODY 2008, 4., 2008, Campinas. **Anais [...]**. Londres: International Phonetic Association, 2010. p. 427-430.

BERTINETTO, P.M; BERTINI, C. Towards a unified predictive model of speech rhythm. **Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa**, Pisa, v. 7, n.8, p.1-24, 2008.

BERTINETTO, P.M.; FOWLER, C.A.1989. On sensitivity to durational modifications in Italian and English. **Rivista di Linguistica**. Pisa, v. 1, n. 1, p. 69-94, 1989.

BEST, C. T. A direct-realist view of cross-language speech perception. *In*: STRANGE, W. (ed.). **Speech perception and linguistic experience: issues in cross-language research**. Timonium: York Press, 1995. p. 171-204.

BEST, C. T.; TYLER, M. D. Nonnative and second-language speech perception: Commonalities and complementarities. *In*: BOHN, O. S.; MUNRO, M. J. (org.). **Language experience in second language speech learning: in honor of James Emil Flege**. Philadelphia: John Benjamins Publishing Company, 2007. p. 13-34.

BISOL, L. O troqueu silábico no sistema fonológico (um adendo ao artigo de Plínio Barbosa). **Documentação de Estudos em Lingüística Teórica e Aplicada**, São Paulo, v. 16, n. 2, p. 403- 413, nov. 2000.

BOERSMA, P.; WEENINK, D. **Praat: doing phonetics by computer**. (Version 6.0). <http://www.praat.org>, 2020.

BORZONE DE MANRIQUE, A. M.; SIGNORINI, A. Segmental duration and rhythm in Spanish. **Journal of Phonetics**, Amsterdã, v. 11, n. 2, p. 117-128, abr. 1983.

BOTINIS, A., GRANSTROM, B., MOBIUS, B. Developments and paradigms in intonation research. **Speech Communication**, Amsterdã, v. 33, n. 4, p. 263-296, mar. 2001.

BUNTA, F.; INGRAM, D. The acquisition of speech rhythm by bilingual Spanish- and English speaking 4- and 5-year-old children. **Journal of Speech, Language, and Hearing Research**, Rockville, v. 50, n. 4, p. 999-1014, ago. 2007.

CAGLIARI, L. C. **Elementos de fonética do português brasileiro**. Tese (Livre Docência em Lingüística) – UNICAMP, Campinas, 1981.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. GRINER, B. **Teaching Pronunciation: a course book and reference guide**. Cambridge: Cambridge University Press, 2010.

CRYSTAL, D. **Dicionário de linguística e fonética**. Tradução de Maria Carmelita Pádua Dias. Rio de Janeiro: Jorge Zahar Editor, 2000.

CUMBERS, B. A. **Perceptual correlates of acoustic measures of vocal variability**. 2013. 97 p. Tese de Mestrado (Master of Science) - College of Health Sciences, University of Wisconsin Milwaukee, Milwaukee, 2013. Disponível em: <https://dc.uwm.edu/etd/214/>. Acesso em: 30 jun. 2021.

CUMMING, R. E. **Speech rhythm: the language-specific integration of pitch and duration**. 2010. 268 f. Tese (Doutorado em Filosofia) – Downing College. Universidade de Cambridge, Cambridge, 2010. Disponível em: <https://www.repository.cam.ac.uk/handle/1810/228685>. Acesso em: 27 mar. 2021.

DAUER R. M. Phonetic and phonological components of language rhythm. *In: 11TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES*, 11., 1987, Tallinn. **Anais [...]**. Londres: International Phonetic Association, 1987. p. 447-450.

DAUER, R. M. Stress-timing and syllable-timing reanalyzed. **Journal of Phonetics**, Amsterdã, v. 11, n. 1, p. 51–62, jan. 1983. Disponível em: [https://doi.org/10.1016/S0095-4470\(19\)30776-4](https://doi.org/10.1016/S0095-4470(19)30776-4). Acesso em: 30 jan. 2021.

DE BOT, K. Introduction: second language development as a dynamic process. **The Modern Language Journal**, Hoboken, v. 92, n. 2, p. 166-178, maio. 2008.

DE BOT, K.; LOWIE, W.; VERSPOOR, M. A dynamic systems theory approach to second language acquisition. **Bilingualism: Language and Cognition**, Cambridge, v. 10, n. 1, p. 7-21, mar. 2007.

DELLWO, V. Rhythm and speech rate: a variation coefficient for deltaC. *In: KARNOWSKI, P.; SZIGETI, I. (ed.). Language and language-processing*. Frankfurt: Peter Lang, 2006. p. 231–241.

DELLWO, V. The role of speech rate in perceiving speech rhythm. *In: SPEECH PROSODY*, 4., 2008, Campinas. **Anais [...]**. Londres: International Speech Communication Association, 2008. p. 375-378.

DELLWO, V.; FOURCIN, A. J.; ABBERTON, E. R. Rhythmical classification of languages based on voice parameters. *In: 16TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES*, 16., 2007, Saarbrücken. **Anais [...]**. Londres: International Phonetic Association, 2007. p. 1129–1132.

DELLWO, V.; WAGNER, P. Relationships between speech rate and rhythm. *In: 15TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES*, 15., 2003, Barcelona. **Anais [...]**. Londres: International Phonetic Association, 2003.

DERWING, T.M. The role of phonological awareness. *In: GARRET, P.; COTS, J.M. (org.). The Routledge Handbook of Language Awareness*. Abingdon: Routledge, 2018. p. 339–354.

- DETERDING, D. The rhythm of Singapore English. *In: FIFTH AUSTRALIAN INTERNATIONAL CONFERENCE ON SPEECH SCIENCE AND TECHNOLOGY*, 5., 1994, Perth. **Anais [...]**. Canberra: Australasian Speech Science and Technology Association, 1994. p. 6–8.
- DING, H.; XU, X. L2 English rhythm in read speech by Chinese students. *In: INTERSPEECH*, 17., 2016, São Francisco. **Anais [...]**. Londres: International Speech Communication Association, 2016. p. 2696-2700.
- FITCH, W. T.; HAUSER, M. D.; CHOMSKY, N. The evolution of the language faculty: clarifications and implications. **Cognition**, Amsterdã, v. 97, n. 2, p. 179–210, dez. 2005.
- FLEGE, J. Second language speech learning: theory, findings and problems. *In: STRANGE, W. (ed.). **Speech perception and linguistic experience: issues in cross-language research***. Timonium: York Press, 1995. p. 233-277.
- FLEGE, J.; BOHN, O. The revised speech learning model (SLM-r). *In: WAYLAND, R. (ed.). **Second language speech learning: theoretical and empirical progress***. Cambridge: Cambridge University Press, 2021. p. 03-83.
- FOWLER, C.A. **Timing control in speech production**. Indiana University Linguistics Club: Bloomington, 1977.
- FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. *In: XV ENCONTRO DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA*, 15., 2000, Faro. **Anais [...]**. Porto: Associação Portuguesa de Linguística, 2000. p. 533–555.
- FROTA, Sónia; VIGÁRIO, Marina; MARTINS, Fernando. Discriminação entre línguas: evidência para classes rítmicas. *In: XVI ENCONTRO DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA*, 16., 2002, Lisboa. **Anais [...]**. Porto: Associação Portuguesa de Linguística, 2002. p. 189–199.
- FUCHS, R. **Speech rhythm in varieties of English**. Singapore: Springer, 2016.
- GIBBON, D.; GUT, U. Measuring speech rhythm. *In: 7th EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY*, 7., 2001, Aalborg. **Anais [...]**. Londres: International Speech Communication Association, 2001. p. 91–94.
- GONÇALVES, J.S.S. Algumas observações sobre a prosódia da representação do discurso outro (RDA). **Linguagem & Ensino**, Pelotas, v.21, n. esp., p. 37-72, 2018.
- GRABE, E; LOW, E. Durational variability in speech and the rhythm class hypothesis. *In: GUSSENHOVEN C.; WARNER, N. (ed.), **Papers in laboratory phonology***. Berlin: De Gruyter Mouton, 2002. p. 515–546.
- GUT, U. Rhythm in L2 speech. *In: GIBBON, D.; HIRST, N.; CAMPBELL N. (ed.). **Rhythm, melody and harmony in speech: studies in honour of Wiktor Jassem***. Cambridge: Cambridge University Press, 2021. p. 83-94.

HE, L. Syllabic intensity variations as quantification of speech rhythm: evidence from both L1 and L2. *In: 6th INTERNATIONAL CONFERENCE ON SPEECH PROSODY*, 6., 2012, Shanghai. **Anais [...]**. Londres: International Speech Communication Association, 2012. p. 466–469.

HELDNER, M. Spectral emphasis as an additional source of information in accent detection. *In: ITRW ON PROSODY IN SPEECH RECOGNITION AND UNDERSTANDING*, 2001, 1., New Jersey. **Anais [...]**. Londres: International Speech Communication Association, 2001.[4] p. Disponível em: https://www.isca-speech.org/archive_open/archive_papers/prosody_2001/prsr_010.pdf. Acesso em: 10 jul. 2021.

HOLLIMAN, A. J.; WOOD, C.; SHEEHY, K. Does speech rhythm sensitivity predict children's reading ability 1 year later? **Journal of Educational Psychology**, Washington, v. 102, n. 2, p. 356–366, maio. 2010.

JEON, H. S. **Acoustic measure of speech rhythm: Korean learners of English**. Dissertação (Mestrado em Ciências), University of Edinburgh, Edinburgh, 2006.

KUPSKE, F.F.; ALVES, U.K.; LIMA JR. R.M. **Investigando os sons de línguas não nativas: uma introdução**. Campinas: Editora da ABRALIN, 2021. *E-book*. Disponível em: <https://editora.abralin.org/publicacoes/investigando-os-sons-de-linguas-nao-nativas/>. Acesso em: 23 set. 2021.

LADEFOGED, P. **A course in phonetics**. New York: Harcourt Brace Jovanovich, Inc., 1975.

LADEFOGED, P.; JOHNSON, K. **A course in phonetics**. Boston: Cengage learning, 2014.

LARSEN-FREEMAN, D. Chaos/complexity science and second language acquisition. **Applied Linguistics**, Oxford, v. 18, n. 2, p. 146-165, jun. 1997.

LARSEN-FREEMAN, D.; CAMERON, L. **Complex systems and applied linguistics**. Oxford: Oxford University Press, 2008.

LEHISTE, I. **Suprasegmentals**. Massachusetts: The M.I.T. Press, 1977

LEVIS, J. M.; SONSAAT, S.; LINK, S.; BARRIUSO, T.A. Native and nonnative teachers of L2 pronunciation: effects on learner performance. **TESOL Quarterly**, Hoboken, v. 50, n. 4, p. 894–931, dez. 2016.

LI, A.; POST, B. L2 acquisition of prosodic properties of speech rhythm: evidence from L1 Mandarin and German learners of English. **Studies in Second Language Acquisition**, Cambridge, v. 36, n. 2, p. 223-255, dez. 2014. Disponível em: <https://doi.org/10.1017/S0272263113000752>. Acesso em: 12 abr. 2021.

LIMA JR., R. M; ALVES, U. K. A dynamic perspective on L2 pronunciation development: bridging research and communicative teaching practice. **Revista do GEL**, São Paulo, v. 16, n. 2, p. 27-56, dez. 2019.

LIMA JR., R. M. Uma investigação dos efeitos do ensino explícito da pronúncia na aula de inglês como língua estrangeira. **Revista Brasileira de Linguística Aplicada**, Belo Horizonte, v. 10, n. 3, p. 747-771, mar. 2010.

LIMA JR., R.M. A necessidade de dados individuais e longitudinais para análise do desenvolvimento fonológico de L2 como sistema complexo. **ReVEL**, [s. l.], v. 14, n. 27, ago. 2016. Disponível em: <http://www.revel.inf.br/files/ed7d4e06176e8c2e5265036095d6ab6d.pdf>. Acesso em: 23 abr. 2021.

LIMA JR., R.M.; GARCIA, G. Probing rhythmic patterns in English-L2 - a preliminary study on vowels reduction by Brazilian learners at different ages. **Journal of Speech Sciences**, Campinas, v. 6, n. 1, nov. 2017.

LIN, H. L.; WANG, Q. Vowel quantity and consonant variance: a comparison between Chinese and English. In: BETWEEN STRESS AND TONE CONFERENCE, 2005, 1., Leiden. **Anais** [...]. Leiden: International Institute for Asian Studies, 2005. p. 1-24.

LISS, J. M.; WHITE, L.; MATTYS, S. L.; LANSFORD, K.; LOTTO, A. J.; SPITZER, S. M.; CAVINESS, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. **Journal of Speech, Language and Hearing Research**, Rockville, v.52, n. 5, p. 1334–1352, out. 2009.

LLÉO, C.; RAKOW, M.; KEHOE, M. Acquiring rhythmically different languages in a bilingual context. In: 16TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 16., 2007, Saarbrücken. **Anais** [...]. Saarbrücken: Universität des Saarlandes, 2007. p. 1545-1548.

LOUKINA, A.; KOCHANSKI, G.; SHIH, C.; KEANE, E.; WATSON, I. Rhythm measures and dimensions of durational variation in speech. **The Journal of the Acoustical Society of America**, College Park, v. 129, n. 5, p. 3258–3270, maio. 2011.

LOW, E. L.; GRABE, E.; NOLAN, F. Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. **Language and Speech**, Thousand Oaks, v. 43, n. 4, p. 377–401, dez. 2000.

LOWIE, W.; VERSPOOR, M. Variability and variation in second language acquisition orders: a dynamic reevaluation. **Language Learning**, Hoboken, v. 65, n. 1, p. 63-88, mar. 2015.

MAJOR, R. C. Stress-timing in Brazilian Portuguese. **Journal of Phonetics**, Amsterdã, v. 9, n. 3, p. 343–351, jul. 1981. Disponível em: [https://doi.org/10.1016/S0095-4470\(19\)30977-5](https://doi.org/10.1016/S0095-4470(19)30977-5). Acesso em: 23 jan. 2021.

MENNEN, I.; SCHAEFFLER, F.; DOCHERTY, G. Cross-language differences in fundamental frequency range: a comparison of English and German. **The Journal of the Acoustical Society of America**, College Park, v. 131, n. 3, p. 2249–2260, mar. 2012.

MICHAELIS moderno dicionário da língua portuguesa. São Paulo: Melhoramentos. Disponível em: <http://michaelis.uol.com.br/moderno/portugues/>. Acesso em: 13 jun. 2021.

- MIGLIORINI, L.; MASSINI-CAGLIARI, G. Sobre o ritmo do português brasileiro: evidências de um padrão acentual. **ReVEL**, [s. l.], v. 8, n. 15, ago. 2010. Disponível em: http://www.revel.inf.br/files/artigos/revel_15_sobre_o_ritmo.pdf. Acesso em: 20 mar. 2021.
- MOK, P. P. K.; LEE, S. I. Korean speech rhythm using rhythmic measures. *In*: 18TH INTERNATIONAL CONGRESS OF LINGUISTS (CIL18), 18., 2008, Seoul. **Anais [...]**. Leiden: Permanent International Committee of Linguists (CIPL), 2008.
- MORENO, M. C. Sobre la adquisición de la prosodia en lengua extranjera: estado de la cuestión. **Didactica (Lengua y Literatura)**, Madri, v. 12, n. 12, p. 91–119, jan. 2000.
- MORGAN, J. L. A rhythmic bias in preverbal speech segmentation. **Journal of Memory and Language**, Amsterdã, v. 35, n. 5, p. 666–688, out. 1996.
- MORLEY, J. A multidimensional curriculum design for speech-pronunciation instruction. *In*: MORLEY, J. (ed.). **Pronunciation pedagogy and theory: new views, new directions**. Bloomington: TESOL, 1994. p. 64-91.
- MUNRO, M. J.; DERWING, T. M. Modeling perceptions of the accentedness and comprehensibility of L2 speech: the role of speaking rate. **Studies in Second Language Acquisition**, Cambridge, v. 23, n. 4, p. 451–468, dez. 2001.
- NAZZI, T.; BERTONCINI, J.; MEHLER, J. Language discrimination by newborns: toward an understanding of the role of rhythm. **Journal of Experimental Psychology: human perception and performance**, Washington, v. 24, n. 3, p. 756–766, jun. 1998.
- NESPOR, M.; SHUKLA, M.; MEHLER, J. Stress-timed vs. syllable-timed languages. *In*: VAN OOSTENDORP, M.; EWEN, C. J.; HUME, E.; RICE, K. (ed.). **The blackwell companion to phonology**, New Jersey, 2011. p. 1–13.
- ORDIN, M.; POLYANSKAYA, L. Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. **The Journal of the Acoustical Society of America**, College Park, v. 138, n. 2, p. 533–544, ago. 2015.
- PELLEGRINO, E. The perception of foreign accented speech. Segmental and suprasegmental features affecting the degree of foreign accent in L2 Italian. *In*: VIITH GSCP INTERNATIONAL CONFERENCE: SPEECH AND CORPORA, 7., 2012, Belo Horizonte. **Anais [...]**. Vercelli: Gruppo di Studio sulla Comunicazione Parlata, p. 261-267.
- PIKE, K. **The intonation of American English**. Ann Arbor: University of Michigan Publications, 1945.
- PINKER, S.; JACKENDOFF, R. The faculty of language: what's special about it? **Cognition**, Amsterdã, v. 95, n. 2, p. 201–236, mar. 2005.
- RAMUS, F.; ROSEN, S.; DAKIN, S.C.; DAY, B. L. CASTELLOTE, J.M.; WHITE, S.; FRITH, U. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. **Brain**, Oxford, v. 126, n. 4, p. 841–865, abr. 2003.

RAMUS, F.; NESPOR, M.; MEHLER, J. Correlates of linguistic rhythm in the speech signal. **Cognition**, Amsterdã, v. 73, n. 3, p. 265–292, dez. 1999.

RATHCKE, T.; SMITH, R. Exploring timing in accents of British English. *In*: 17TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 17., 2011, Hong Kong. **Anais [...]**. Londres: International Phonetic Association, 2011. p.1666–1669.

RUSSO, M.; BARRY, W. J. Isochrony reconsidered - objectifying relations between rhythm measures and speech tempo. *In*: SPEECH PROSODY, 4., 2008, Campinas. **Anais [...]**. Londres: International Phonetic Association, 2010. p. 419-422.

SILVA JR., L. J. **Interferências rítmicas do português brasileiro no inglês como L2: o choque acentual**. 2013. 223 p. Tese de Doutorado (Doutorado em Linguística) - Centro de Ciências Humanas, Letras e Artes, Universidade Federal da Paraíba, João Pessoa, 2013.

SILVA JR., L.J.; BARBOSA, P. **Metrics & Acoustics Extractor version 1.0**. Script for Praat. 2019a.

SILVA JR., L.J.; BARBOSA, P. Speech rhythm of English as L2: as investigation of prosodic variables on the production of Brazilian Portuguese speakers. **Journal of Speech Sciences**, Campinas, v. 8, n. 2, p. 37–57, ago. 2019b.

STEINER, I. On the analysis of speech rhythm through acoustic parameters. *In*: FISSANI, B.; HANS-CHRISTIAN, S.; SCHRÖDER, B.; WAGNER, P. (org). **Sprachtechnologie, mobile kommunikation und linguistische ressourcen**. Berna: Peter Lang, 2005. p. 647–658.

TEIXEIRA, L.A.S.; LIMA JR., R.M. Análise do desenvolvimento do ritmo do inglês-L2 por brasileiros por meio de três métricas rítmicas. **Revista X**, v. 16, n. 5, p. 1258-1292, set. 2021. Disponível em: <https://revistas.ufpr.br/revistax/article/view/81413>. Acesso em: 22 set. 2021.

THOMSON, R. I.; DERWING, T. M. The effectiveness of L2 pronunciation instruction: a narrative review. **Applied Linguistics**, Oxford, v. 36, n. 3, p. 326–344, jul. 2015.

TSIARTSIONI, E. **The acquisition of features of rhythm and stop voicing in Greek and English L2**. 2003. Dissertação (Mestrado em Filosofia) - Center for Language and Communication Studies, Trinity College Dublin, 2003.

URBANI, M. Pitch range in L1/L2 English. An analysis of F0 using LTD and linguistic measures. *In*: BUSÁ, M. G.; STELLA, A. (ed.). **Methodological perspectives on second language prosody**. Padova: Coop. Libreria Editrice Università di Padova, 2012. p. 79–83.

VOGEL, I.; ATHANASOPOULOU, A.; GUZZO, N. B. Timing properties of (Brazilian) Portuguese and (European) Spanish. *In*: REPETTI, L.; ORDÓÑEZ, F. (ed.). **Romance Languages and Linguistic Theory 14**. Amsterdã: John Benjamins Publishing Company, 2018. p. 325-340.

WAGNER, P. S; DELLWO, V. Introducing YARD (yet another rhythm determination) and re-introducing isochrony to rhythm research. *In*: SPEECH PROSODY, 2., 2004, Nara. **Anais [...]**. Londres: International Phonetic Association, 2004. p. 227-230.

- WHITE, D.; MOK, P. L2 Speech rhythm and language experience in new immigrants. *In: PROCEEDINGS OF THE 19TH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES*, 19., 2019, Melbourne. **Anais [...]**. Londres: International Phonetic Association, 2019. p. 1-5.
- WHITE, L.; MATTYS, S. L. Calibrating rhythm: first language and second language studies. **Journal of Phonetics**, Amsterdã, v. 35, n. 4, p. 501–522, out. 2007.
- WHITWORTH, N. Speech rhythm production in three German-English bilingual families. **Leeds Working Papers in Linguistics and Phonetics**, Leeds, v. 9, n. 3, p. 175–205, 2002.
- WIGET, L.; WHITE, L.; SCHUPPLER, B.; GRENON, I.; RAUCH, O.; MATTYS, S. L. How stable are acoustic metrics of contrastive speech rhythm? **The Journal of the Acoustical Society of America**, v. 127, n. 3, p. 1559–1569, mar. 2010.

ANEXO A – TEXTO EM INGLÊS

Is English your native language? If not, your foreign accent may show people that you come from another country. Why is it difficult to speak a foreign language without an accent? There are a couple of answers to this question. First, age is an important factor in learning to pronounce. We know that young children can learn a second language with perfect pronunciation. We also know that older learners usually have an accent, though some older individuals also have learned to speak without an accent.

ANEXO B – TEXTO EM PORTUGUÊS

O inglês é a sua língua nativa? Caso não seja, o seu sotaque estrangeiro pode mostrar para as pessoas que você vem de outro país. Por que é difícil falar uma língua estrangeira sem sotaque? Existem algumas respostas para essa pergunta. Primeiro, idade é um fator importante na aprendizagem da pronúncia. Nós sabemos que crianças pequenas conseguem aprender uma segunda língua com pronúncia perfeita. Também sabemos que aprendizes mais velhos normalmente têm sotaque, apesar de alguns aprendizes mais velhos também conseguirem aprender a falar sem sotaque algum.