

STRESS CLASH RESOLUTION IN ENGLISH AS L1 AND L2**RESOLUÇÃO DE CHOQUE ACENTUAL NO INGLÊS COMO L1 E L2**Leônidas J. Silva Jr³¹Ester M. Scarpa³²

ABSTRACT: This paper aims to analyze which strategy speakers of English as L1 and Brazilian speakers of English as L2 use to solve stress clash, a phenomenon in which two syllables bearing primary stress are adjacent in different words forming a phonological phrase such as [*thirteen men*]. The representation of stress clash, as well as the operation that allows its undoing, is one of the justifications whereby Liberman & Prince (1977) propose the metrical grid. The clash depends on information about the metrical level in which it occurs. The simple phonetic adjacency is not enough to characterize a clash. This paper is based on the work of Pike (1945), Selkirk (1984), Major (1985) to show distinctions between language types of rhythm (stress/syllable-timed ones) as well as a non-categorical polarization for rhythm in languages (Barbosa, 2000, 2002) and Silva Jr, (2013), Fragozo (2017) when comparing native and Brazilian speakers of English for the choice of stress clash solving strategies. For the Methods, we have done acoustic analysis of what we extracted and normalized the vowel duration values and applied to a statistical analysis using ANOVA test to check the degree of variance between Brazilians and native speakers of English. Our results show that native and Brazilian speakers of English use distinct strategies to solve stress clash: the former using stress retraction, which undoes primary stress clashes by moving the first stress of the clash to the left and the latter using silent demibeat addition, inserting a short pause in between the clash environment.

KEYWORDS: Stress clash. Speech Rhythm. English as L2.

RESUMO: Este artigo tem como objetivo analisar que estratégia falantes de inglês como L1 e falantes brasileiros usam para resolver o choque acentual, fenômeno em que duas sílabas com acento primário são adjacentes em palavras diferentes formando uma frase fonológica como [*thirteenmen*]. A representação do choque acentual, bem como a operação que permite a sua dissolução, é uma das justificativas pelas quais Liberman e Prince (1977) propõem a grade métrica. O choque depende de informações sobre o nível métrico em que ocorre. A simples adjacência fonética não é suficiente para caracterizar um choque. Este trabalho baseia-se nos estudos de Pike (1945), Selkirk (1984), Major (1985) para mostrar distinções entre os tipos de ritmo linguísticos (accentual/silábico), bem como a não-polarização de categorias para determinar o ritmo da fala (Barbosa, 2000, 2002) além de SILVA JR, (2013) e Fragozo (2017) comparando falantes nativos de inglês e brasileiros de inglês/L2 quando da escolha de estratégias de resolução de choque acentual. Na Metodologia, realizamos análise acústica da qual extraímos e normalizamos os valores da duração da vogal e aplicamos a uma análise estatística utilizando o teste ANOVA para verificar o grau de variância entre brasileiros e falantes nativos de inglês. Nossos resultados mostram que falantes nativos e brasileiros usam estratégias distintas para resolver o choque acentual: o primeiro usando a retração acentual, que desfaz os choques, movendo o primeiro acento do choque para a esquerda e o último, usando a adição de batida silenciosa, inserindo uma breve pausa no ambiente do choque.

PALAVRAS-CHAVE: Choque acentual. Ritmo da fala. Inglês como L2.

1 Introduction

In the study of the transfer of a speaker's mother tongue on a foreign language pronunciation, prosodic features stand out as a very conspicuous characteristic of a foreign accent. In fact, it has been observed that prosody (rhythm, stress, tone, intonation) plays a role that is as important as segmental aspects in conversational processes with regard to L2

³¹ Doutor em Linguística (UFPB) com Pós-Doutorado em Linguística (UNICAMP). Professor adjunto no Departamento de Letras da Universidade Estadual da Paraíba (DL/UEPB). E-mail: leonidas.silvajr@gmail.com

³² Doutora em Linguística (University of London) com Pós-doutorado em Linguística (Oxford University; Newcastle University e Reading University). Professora titular no Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (IEL/UNICAMP) e Bolsista de Produtividade em Pesquisa do CNPq - Nível 1C. E-mail: ester.scarpa@gmail.com

intelligibility (Crystal, 1999; Roach, 2009; Celce-Murcia et al. (2010); Avery & Ehrlich, 2012) and poses a key challenge to L2 teaching. Pike's work on intonation patterns in American English (Pike, 1945) sets the tone for prosody teaching and inspired much of didactic material for the teaching of English as L2 for decades to come, mainly by establishing pronunciation drills based on four relative pitch levels of F0 contours. Yet, the segmental aspect of L1 transfer on L2 acquisition has had a much closer attention than the prosodic ones in analyses both of production and of perception.

This paper intends to be a contribution to the study of rhythmic influence of L1 on L2 in prosodic contexts of stress clash. It aims to analyze how native speakers of Brazilian Portuguese and Americans as native speakers of English produce stress clash during the phonetic realization of English as a foreign language and mother tongue respectively and also which strategies are used to solve these clashes. We assume that stress clash is a phenomenon in which two syllables bearing primary stress are adjacent in different words forming a phonological phrase.

The paper is organized as follows. In section 2, we briefly present the key assumptions about the hierarchical theory of rhythm and the structure of stress clash. We also discuss some controversies about Brazilian Portuguese rhythm and the nature of stress clash as well as some proposals for stress clash resolution. In section 3 we present the methods and data used for the analysis. Section 4 shows analyses of the data, where we test the hypothesis that stress clash is mostly solved by Americans and Brazilians speakers distinctively: the former through stress retraction and the latter through the silent demibeat strategy (Selkirk, 1984), rather than stress retraction favored by native speakers of English (Lieberman & Prince, 1977; Hayes, 1995). They do so regardless of syntactic (within phonological phrases) or intonational factors, that is, in sentences with no F0 variation due to marked or narrow focus. Section 5 presents some concluding remarks and section 6 the references used along this paper.

2 Literature review

In this section we will see previous works about stress clash; definition and strategies to solve it, the concept of Rhythm in Linguistics, the prosodic hierarchy and the domain for solving stress clashes and some studies of Brazilian Portuguese solving strategies claimed in clash environment.

2.1 Rhythm and stress clash

Rhythm refers to a given movement that is structured by regular successions of time (isochrony) of strong and weak beats in a unit of speech. Rhythm denotes, then, interstress events that occur in regular speech intervals (Roach, 2005). Rhythm seen as an interstress-related architecture was a major motivation for the classical classification of languages into two categories, *stress-timed* (when rhythm is based on stressed syllables of words in an utterance that occur at apparently regular intervals) - and as *syllable-timed* (when rhythm is based on the syllable, each syllable taking a quasi-similar amount of time in its production).

More than one century ago, Jones (1918) had already noticed that there is a tendency of connected speech to consist of strong and equidistant syllables (the Rhythm Unit) as he poses: "(...) *The (rhythm) units tend to follow one another in such a way that the lapse of time between the beginning of their prominent syllables is somewhat uniform*" (p. 34). This tendency would explain the variation in duration on the stressed syllables in English.

Some decades later, Pike (1945) proposed that this kind of Rhythm Unit is a *stress-timed* rhythm typical of languages such as English. It contrasts with *syllable-timed* languages,

in which the duration of the syllable depends on a given number of syllables and not on the presence of a strong stress. Abercrombie (1967) established a link between “chest-pulse” and “breath-pulse” in the stress-producing mechanism. Those account for *stress-timed* and *syllable-timed* rhythm, respectively. Coalescence processes (segment or syllable deletion, weakening of unstressed positions, amalgamation, assimilation) tend to appear more commonly in stress-timed languages that bring about variation in the temporal structure of the utterance. Both authors, albeit extending the dichotomy to all languages generically, point out that difference in rhythm may depend on other factors, such as pragmatic intention.

As soon as phonetic measurements started to be taken in several languages in order to bring evidence for one or the other type of rhythm, controversy related to the conflict among acoustic phonetic approaches, on the one hand, and phonological and perceptual/auditory views on the other hand, started to spring up in the literature. The isochrony controversy, therefore, seems to be almost as old as the concept itself. As stated by Couper-Kuhlen (1993), already in 1939 Classe had used the chimograph to measure the interval between stressed syllables in sentences. The results show that a precise isochrony is only met in very special circumstances or conditions. In order to attain a near-perfect isochrony, the rhythmic groups that were tested through specific reading techniques had to have similar number of syllables with a similar phonetic and grammatical structure, read with a slow tempo.

Isochrony, in this perspective, should then be considered an “illusion” acoustically speaking, but a real and concrete perceptual phenomenon (Couper-Kuhlen, 1993). Apparently, isochronic stretches that are created by prominent syllables at regular intervals, should be initially considered as an identifiable and independent phenomenon, which does not preclude that, once identified, such phenomena should not be related to a syntactic, semantic, pragmatic or intonational organization of speech.

Moreover, the exclusive emphasis on the acoustic signal may betray the fallacy that acoustic signals translate themselves into linguistic unities. These are constructed cognitively, linguistically, including the prosodic-rhythmic unities of languages. The autosegmental approach to rhythm intends to bridge this gap.

2.2 Stress clash and hierarchical organization of speech rhythm

The metrical approach puts forward the notion of stress clash in an attempt to adjust it in the process of eurhythmy at the levels of word and phonological phrasing of the relative stress of prosodic constituents. Words and phrases are organized by means of relative prominence among syllables, words and phrases. A binary “strong–weak” (*s-w*) relationship is defined in syllables and groups of syllables underlying the level of the word. Above it, the relationship *s-w* is defined on morphosyntactic constituents. The prominence relationships may be represented either with *tree-like* or with a *text-to-grid* binary structure, according to the principles of rhythmic alternance, based on eurhythmy or rhythmic euphony (Selkirk, 1984; Hayes, 1984). The notion of stress clash stems from the principles of rhythmic alternance and euphonic optimization of the metrical structure of languages. The relations of prominence internally to words tend to keep the stress syllable constant, but within phonological phrases the stress pattern resulting from the primary accent algorithm may be modified.

The classical example by Selkirk (1984) is the sequence “[thirteenmèn]”. Normally, when pronounced in isolation, [*thirTÉEN*] bears the primary stress in the last syllable. However, if it is syntactically concatenated to another word, whose first syllable bears a stress of relative force of equal or higher strength, such as [*mén*], the tendency is that a stress reversion occurs, which transfers the primary stress of “thirteen” to its initial syllable, thus generating [*THÍRteen MÉN*].

In other words, stress clash is a phenomenon in which two syllables bearing primary stress are adjacent in different words forming a phonological phrase (P-phrase) such as [marébrava] (high tide). The representation of stress clash, as well as the operation that allows its undoing, is one of the justifications whereby Liberman & Prince (1977) propose the metrical grid. The clash depends on information about the metrical level in which it occurs. The simple phonetic adjacency is not enough to characterize a clash. Only adjacent primary stresses on the same line of the grid are felt as a clash and are subject to rhythmic adjustment. One of the euphony rules is *beat movement* (BM), which undoes primary stress clashes by moving the first stress of the clash to the left (as above [thirTÉEN MÉN]* becomes [THÍrteen MÉN]). This is the case of English and other stress-timed languages.

When there is no adjustment in the grid by retraction nor beat deletion, Selkirk (1984) proposes *silent demibeats* (SDB) occupied also by pauses that interrupt the clash adjacency (as above [maRÉ BRAva]* becomes [maRÉ^(SDB) BRAva] (high tide). They are more frequent in syllable-timed languages.

2.3 Some considerations on rhythm and stress clash in Brazilian Portuguese

The same controversy between periodicity and synchronicity about stress-timing and syllable-timing rhythm is found in studies about Brazilian Portuguese (BP), depending on the methodology used. Work based on acoustic analysis of interstress intervals interprets BP as syllable-timed rhythm (Major; 1985) and some theoretical frameworks based on phonological analysis lead to the conclusion that BP is stress-timed (Dauer, 1983). Evidence of both types of rhythm or of a “mixed rhythm” has been the recent trends (Bisol, 2000), as well as the assertion that the synchronicity polarization is a false one (Barbosa, 2000).

This fact is also attested in Barbosa (2002) when he assumes the coupled-oscillator model³³ of speech rhythm production for BP. The author claims that speech rhythm is understood as a consequence of the variation of the perceived duration along the entire utterance.

The autosegmental approach of BP rhythm has been based on the interface between phonology and syntax in the determination of which prosodic domains on which rhythmic rules and stress clash should be applied (Sândalo&Truckenbrodt, 2001; Santos, 2002; 2003). As well as Selkirk (1984) and Hayes (1995), Aboussalh (1997) shows evidence that the phenomenon of stress clash and clash resolution are domain-sensitive and the P-phrase is the elected level for rhythmic rules to be applied in BP.

Another trend in autosegmental studies about BP rhythm is to determine which prosodic strategy for clash solving is followed. Here too there is controversy. Sândalo&Truckenbrodt (2001) and Santos (2002, 2003) claim that stress clash resolution rules in BP, accounting P-phrase level, are the same as in English, that is, through retraction. On the other hand, Silva Jr. (2013) finds acoustic evidence in favor of demibeat insertion rather than BM, it is to say, from the first primary stressed vowel in the syllable being moved to the immediately preceding vowel in the syllable. Fragozo (2017) finds perceptual evidence in favor of stress clash rather than acoustic evidence and reinforces that both the former and the latter act similarly for native and Brazilian speakers of English.

³³ The coupled-oscillator model has a “phrase-stress and a syllabic oscillator”. The former is implemented by a train of pulses whereas the latter is implemented by a sinusoidal function working bi-directionally.

In this work we follow the conclusions of Nespor (1990), Sândalo&Truckenbrodt (2001) and Santos (2002, 2003) that stress clash and clash resolution have to do with the interface between phonology and syntax (through prosodic structure). Thus the matter may have a stronger impact in language learning beyond a mere pronunciation reach. The transfer of rhythmic structure from a native language (L1) to a foreign language(L2) may affect the meaning and interpretation of the sentences themselves.

3 METHODS

In this section we will describe the subjects' screening process, the procedure taken and the determination of the acoustic and statistical analyses protocols.

3.1 The subjects

For this study, we collected audio data of 03 BP speakers of English as L2 (*experimental group*- EG) and 03 American English (AmE) native speakers (*control group* - CG). In order to test our hypothesis, it was necessary to accurately determine the English proficiency level of the EG. For this, we applied the Oxford Online Placement Test (OOPT) (Purpura, 2018)³⁴.

After applying the test, the subjects were separated into two distinct categories of OOPT proficiency levels: **C1** (*Advanced Proficient User – Effective operational command*) – and **B2** (*High intermediate Independent User - Vantage*). The speakers' proficiency levels were normalized³⁵ into a unique proficiency group (see COUNCIL OF EUROPE, 2001) from different common abilities that intersected both B2 and C1 and established only one proficiency level as shows table 1:

Table 1: Normalized proficiency level of Brazilian speakers (Source: Personal collection).

<i>Listening</i>	<i>Grammar/Vocabulary</i>	<i>Normalized Proficiency Level</i>
C1/B2 → B2/B1	C1/ B2	NPL

3.2 Procedures

In the procedure of our research, speakers were told that their task would be to read aloud a text and this would be recorded. The text contained 05 clash-triggered P-phrases. They were shown the text in advance to be familiarized with and to avoid anxiety while reading. Speakers were recorded at a quiet room using a *Zoom H1 Recorder*, a sampling rate of 44.1 kHz and 16-bit quantization to ensure high quality of dataset that would be later used in the acoustic analysis. Speakers perform a total of 30 samples³⁶ (15 samples each group).

³⁴ The Oxford Online Placement Test (OOPT) helps in establishing the learner's proficiency place so that s/he is properly allocated in a consistent environment with the reality of their English language skills). The acronyms used to determine the degree of proficiency of subjects - C1, B1 and B2 - indicate the level is from highest to lowest proficiency according to OOPT.

³⁵ Normalization procedure was conducted through the application of the Swiss National Science Research Council Survey (1994; 1995 cited in COUNCIL OF EUROPE, 2001) that developed a descriptor-level scale identifying a band of language use. This 9-point scale provides more or less equally sized, coherent levels for the determination of groups (see COUNCIL OF EUROPE, 2001, p. 31).

³⁶ The minimum amount of samples required to run a probabilistic-statistical test such as ANOVA, is 10 (*ten*) samples per group (see Rietveld & Hout, 2005; Anderson *et al*, 2005; Lowie & Seton, 2013).

3.3 Acoustic analysis

All of the acoustic analysis and acoustic image plots were carried out through free software Praat (Boersma & Weenink, 2019). In the first moment, we segmented our data into phonological syllables at *P-phrase* prosodic domain as being the most suitable prosodic level for stress clash and stress clash alleviation (see Nespor & Vogel, 2007). Then, we extracted the vowel of each clashing and preceding-clash syllable environment. For each P-phrase of our data, we based the analysis on the vowel duration in milliseconds (*ms*) of the first vowel in stress clash condition (*in-clash* vowel – V2) and of the vowel at its immediate preceding syllable (*pre-clash* vowel – V1) to determine which strategy would be used for clash alleviation by the speakers. After extracting the raw vowel durations, we normalized our data using two different rhythm metrics: *percentual* for vowels (%V) to check the amount of vowel by means of percentage is phonetically produced in terms of duration for both AmE and BP and *delta* for vowels (ΔV) to check the variability of %V for both languages as well. These metrics were proposed by Ramus et al. (1999) and detected automatically through a script for Praat (Siva Jr. and Barbosa, 2019). The vowels were segmented dynamically and taken its full portion (the pure vowel) into account. Rhotics and nasals in onset and coda positions were separated from their F3 values and semivowels from their F1 and F2 values (see Hillenbrand, 2013). The acoustic analysis procedures graphically described below in figures 1 and 2 for an AmE and BP speaker respectively:

Figure 1: Waveform and spectrogram of the P-phrase: [commit moments] ϕ (first tier) and for V1 and V2 (second tier) from an AmE speaker. $V1_{duration} = 129,8ms$ | $V2_{duration} = 69ms$ (Source: Personal collection).

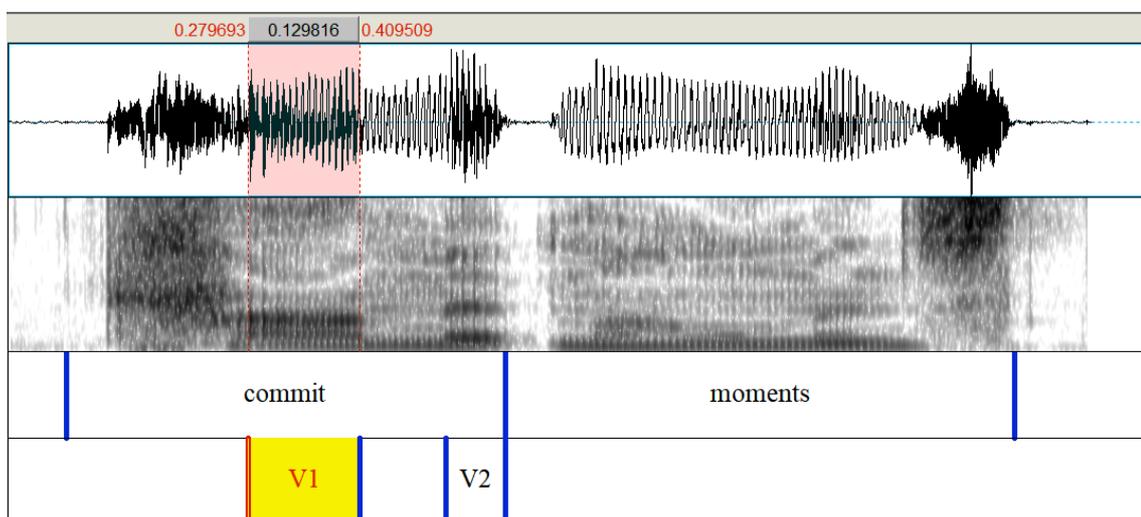
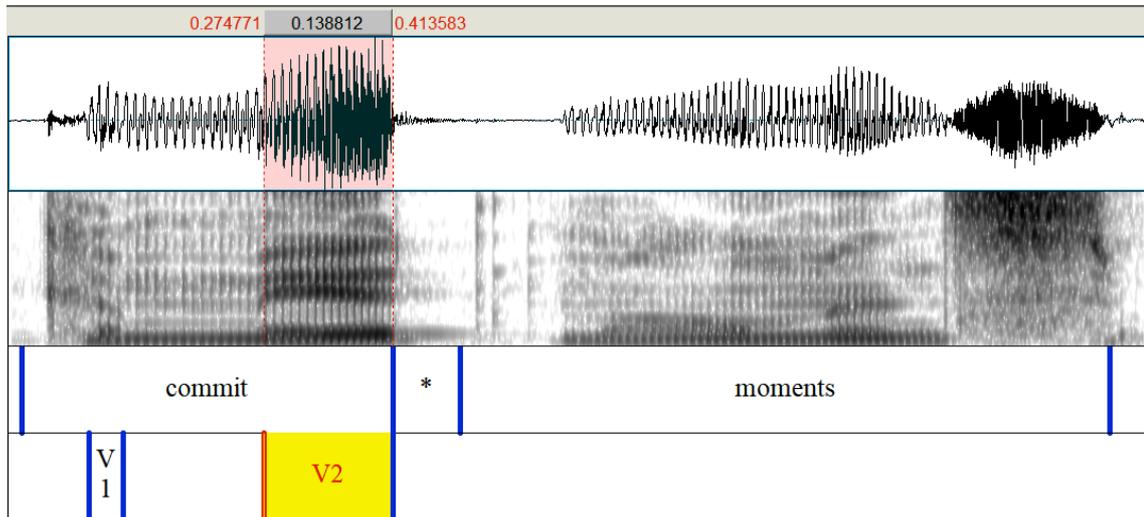


Figure 2: Waveform and spectrogram of the P-phrase: [commit moments]φ (first tier) and for V1 and V2 (second tier) from a BP speaker. $V1_{duration} = 36\text{ ms}$ | $V2_{duration} = 138,8\text{ ms}$ (Source: Personal collection).



As already mentioned, the %V is the proportion of how much of relative V1 and V2 duration is calculated at speech rate as we can see in (1) the %V formula and its detailed version for our purpose. The ΔV is the *standard deviation* of the V1 and V2 as we can see in formula 2 and its detailed version for our purpose:

(1) *Percentual of vowels (%V)*

$$\%V = \sum_{i=1}^n \frac{V_i}{(V_i + C_i)} (100)$$

$$\%V = \frac{\sum_{i=\text{each vowel duration}}^{\text{all vowel durations}} \text{each vowel duration}_i}{(\text{each vowel duration}_i + \text{each consonant duration}_i)} \times 100$$

(2) *Delta for vowels (ΔV)*

$$\Delta V = \sqrt{\frac{\sum_{i=1}^n (V_i - V)^2}{n-1}}$$

$$\Delta V = \sqrt{\frac{\sum_{i=\text{each vowel sample}}^{\text{number of samples}} (\text{each vowel duration}_i - \text{vowel duration mean})^2}{\text{number of samples} - 1}}$$

To describe acoustic evidence of stress clash as well as defining solving strategies, *vowel duration* values were taken into account as being the most robust acoustic parameter for vowel measurement. Our choice to use normalized vowel nuclei duration arose from the assumption that these members have stability when it comes to defining the length of syllables in the determination of clash-solving strategies as attested by Yao (2007). Moreover, we have a significant quantity of monosyllabic and disyllabic words in English that highly onset and coda-variate compounding our P-phrases. This fact brings inconsistency taking *p*-

*centers*³⁷ (see Marcus, 1976) or phonetic syllable³⁸ (see Barbosa, 2006) rather than vowels for clash-solving contexts.

Previous studies chose vowel duration for analyzing rhythm in English as L2 such as Sarmah et. al. (2009) and Gibbon & Gut (2001). The former claims that the duration of vowels acts directly as a vector of rhythmic value showing great reliability for measuring English as L2 speakers' production. The latter, used vowel duration as the most accurate acoustic correlate measure and lists some factors to support this hypothesis:

- Rhythmic durations are directly associated with the duration of vowels and not with syllables;
- The length of consonant clusters is irrelevant, either in the onset or in the coda position;
- Only one parameter is sufficient to measure the syllable.

In favor of vowel duration for our research, we also considered the experiment of Mehler et. al. (1996) when they assert that: “*Vowels own most of the energy in the speech signal; Vowels are longer than consonants; Vowels function as a stress carrier; Vowels indicate whether the syllable is strong or weak*”.

3.4 Statistical analysis

For the statistical analysis, we ran a model for a 1-way ANOVA test for our *control group* and a model for the *Kruskal-Wallis* test for the *experimental group* since – for the latter - the data failed to meet the three conditions for the conventional ANOVA (*normality of the residuals, homoscedasticity of variances, and independence of the samples*). So, a non-parametric correlate method had to be used instead. A significance level (*alpha*) of 5% was established to test whether the % means of the duration values from V1 and V2 of our two groups of speakers varied significantly or not. The purpose of applying ANOVA is to test the hypothesis that BP speakers of English as L2 solve stress clash with the SDB addition rather than stress shifting.

For our analysis, the *Language* with two levels: *American English* (AmE) speakers and *Brazilian Portuguese* (BP) speakers of English as L2 was the independent/explanatory variable and the %V1/V2 and ΔV were the dependent/response variables. All statistical tests, the graph plots for statistical purposes as well as the effect size and estimation of the samples were carried out in R environment (R CORE TEAM, 2019) through a script for R (Silva Jr, 2019). For the effect size and estimation of the samples, we used the *Effects* package for R (Fox and Weisberg, 2019).

4 RESULTS AND DISCUSSION

In this section we will discuss our results and draw inferences from the observed data behavior.

4.1 Vowel duration

Table 2 brings the %V1/V2 and the mean (μ)%V1/V2 and their significance level (*p-value*) ran by ANOVA and (and its non-parametric version *Kruskal-Wallis*) for all of the P-phrases produced by each group:

³⁷ “**P-center**” (*perceptual center*) is understood as an instant of time perceived in a consonant-vowel (CV) transition.

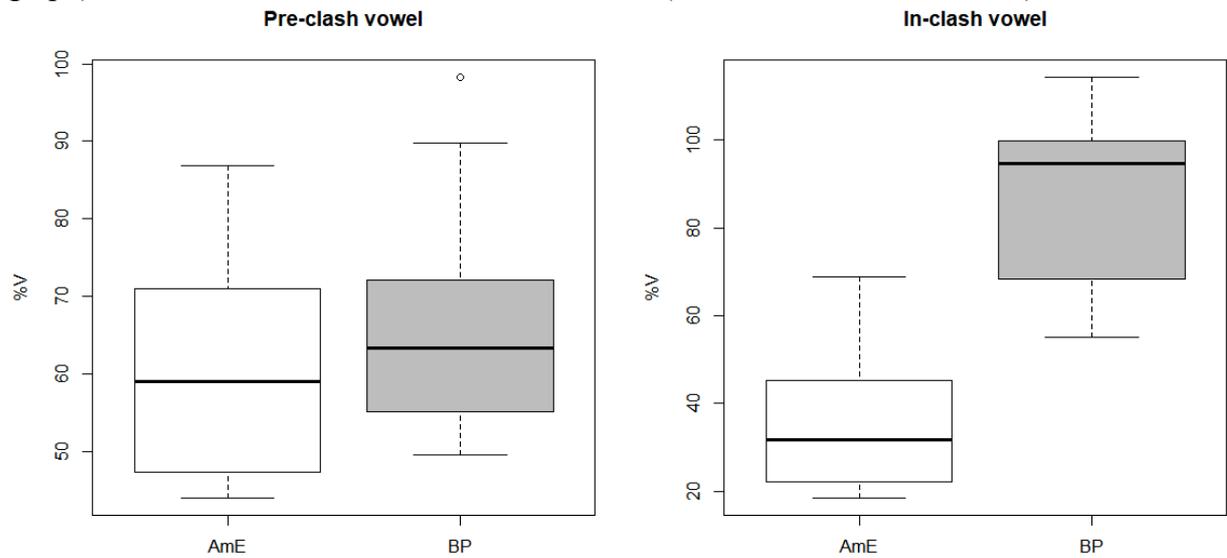
³⁸ We assume phonetic syllable as being determined from vowel onset to the next right-aligned vowel onset.

Table 2: English P-phrases, % values and $\mu\%$ values of V1 and V2, ΔV and $\mu\Delta V$ values, P-phrases p-values and collation p-values for AmE and BP speakers (Source: Personal collection).

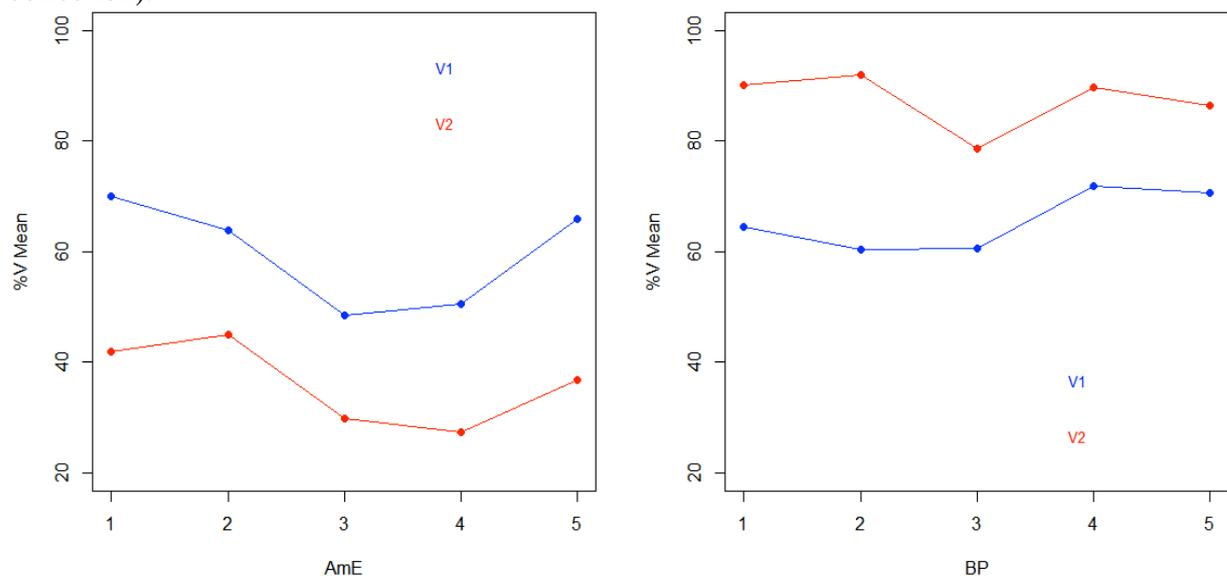
P-PHRASE	LANGUAGE												P-VALUE AmE↔ BP $\mu\%V1$ $\mu\%V2$
	AmE						BP						
	%V1	%V2	$\mu\%V1$	$\mu\%V2$	ΔV	$\mu\Delta V$	%V1	%V2	$\mu\%V1$	$\mu\%V2$	ΔV	$\mu\Delta V$	
<i>bel-air boy</i>	59	22,6			58		56	69,3			46,1		>0,7 <0,0001***
	75,2	43,7			53,8		63,4	99,1			50,8		
	76	59,8	70.06	42,03	49	53,6	73,8	101,7	64,40	90,03	51,2	49,36	
	49,9	19,8			59		59,5	67,6			42,4		
<i>thirteen balls</i>	71	46,6			50,4		51,2	95,7			50,4		>0,1 <0,0001***
	70,9	68,9	63.93	45,10	40,1	49,83	70,5	112,5	60,40	91,93	60,1	50,97	
	46,1	31,7			49,3		58,2	58,6			49,5		
	54,3	39,4			51,2		49,5	88,7			39,2		
<i>commit moments</i>	45,2	18,5	48.53	29,86	55,6	52,03	73,8	88,7	60,50	78,66	48,5	45,73	>0.5 <0,0001***
	48,4	23,7			56,2		53,9	55			33,4		
	59,5	37,6			57,3		98,3	114,3			48,2		
	44	20,6	50.63	27,30	50	54,5	63,3	100,1	71,83	89,80	59,9	47,17	
<i>recall process</i>	64,2	21,7			69,7		54,2	65,2			40,7		>0,7 <0,0001***
	86,9	65,1			46,9		89,8	99,5			39,6		
	46,4	23,9	65.83	36,90	59,9	58,83	67,8	94,7	70,60	86,46	54,1	44,8	
LANGUAGE COLLATION (AmE↔ BP)													
$\mu\%V1$													>0,8
$\mu\%V2$													<0,0001***
$\mu\%V1 - \mu\%V2$													<0,0001***
$\mu\Delta V$													>0,1

From the information shown at table 3, both AmE and BP speakers present significant variability when collating their %V1 to their %V2 productions and between %V2 of both groups as we shall see the frequency distribution of the samples along the boxplots in graph 1; centrality and extreme values. On the one hand, American speakers reduce their *in-clash* vowel and retract stress to the preceding vowel using BM for clash alleviation. On the other hand, Brazilian speakers prefer to solve the clashes via SDB strategy along their productions, as shown at table 2 through values and mean values and at graph 2 through the linear tracks of mean values:

Graph 1: Boxplots for the %V of AmE and BP speakers of the pre-clash vowel (V1, on the left portion of the graph) and the in-cash vowel (V2, on the right portion of the graph). Whiskers indicate 95% confidence interval (Source: Personal collection).



Graph 2: %V1 and %V2 mean values for the AmE and BP speakers (Source: Personal collection).



We can infer from these observations that besides the choice for SDB strategy, BP speakers significantly increase the duration of the in-clash vowel. We corroborate similar results found by Barbosa (2002) who acoustically verified the production of stress clash

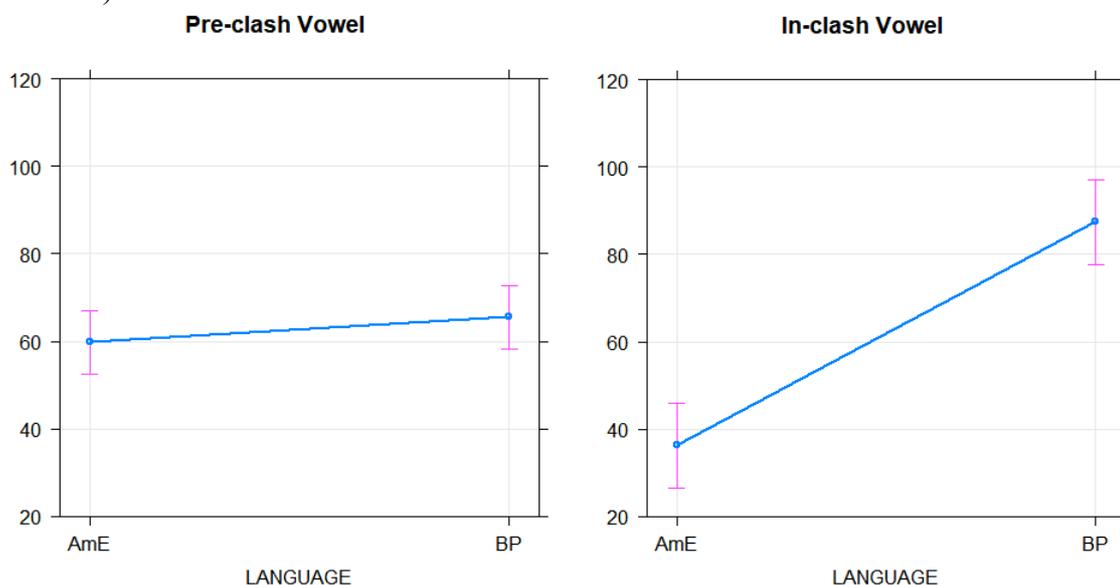
inBrazilian Portuguese P-phrases through duration. The author claims for the fact that in dynamical speech, the more we approximate to the phrasal stress which is the immediate syllable (vowel) succeeding the *in-clash vowel*, BP speakers realizes a deceleration of the jaw to produce the phrasal stress and consequently the *in-clash* vowel becomes longer in duration. Yet at graph 2 as before mentioned, it is possible to compare that the amount of vowel of the P-phrases for BP production is robustly higher and longer in V2 than AmE production, that is, as far as our data are concerned, BP speakers are sensitive to stress shift and prefer SDB resolution.

When we assume the study of Silva Jr. (2013), it is interesting to bring up a comparison between the phonetic production – not only for English, but for BP productions as well. The author attests that BP P-phrases are spoken longer than English P-phrases by BP speakers and AmE speakers prefers stress-shifting resolution in both languages and when speaking BP they reduce *in-clash* vowels even more than they do in English so that it is evidence that native speakers of English are sensitive to SDB. Either stress-shifting (used for clash solving by the so-called *stress-timed* languages) – by AmE speakers or SDB (used for clash solving by the so-called *syllable-timed* languages) for BP speakers fulfill the *Obligatory Contour Principle* (OCP) rule and the *Destressing Clash* algorithm (Nespor & Vogel, 2007; Hayes, 1995).

Our results show evidence against the hypothesis that there is stress shift in a clash environment in Brazilian Portuguese (Major, 1981; Dauer, 1983; 1987; Abousalh, 1997; Sândalo and Truckenbrodt, 2001; Santos, 2002; 2003; 2004).

They also show somewhat different results than the ones put forward by Fragozo (2017) who states that “*results of both native speakers of English and BP speakers regarding the application of rule for solving stress clash act similarly either acoustically or perceptually; especially the advanced proficiency level group of speakers*” (p. 158). The author refers to a multicolinearrelation by means of duration for the choice of stress clash resolution. Fragozo (2017) also asserts that for native speakers of English “*stress clash is a phenomenon of perceptual nature but not an acoustic one*” (p. 154). Against this statement our acoustic parameters show orthogonality relation between %V1 and %V2 and their estimate values for both native and non-native speakers of English as we may see in graph 3:

Graph 3: Effect size of %V1 and %V2 estimate values for the AmE and BP speakers (left and right portion respectively). Whiskers indicate 95% confidence interval (Source: Personal collection).



4.2 The percentual and delta metrics

Comparing AmE and BP data in table 2, we conclude that %V1 does not bring distant values for the two groups. On the contrary it brings low variability between the data and language. We may either see that, even applying SDB in a clash environment, BP speakers present a more balanced syllable-size pattern as attested by Barbosa (2002), to the detriment of V1 duration whereas AmE speakers tend to reduce the clashing vowel duration significantly in a form of a stress-sized pattern and increase V1. We must highlight the fact that even increasing V1 for the choice of BM as clash resolution, %V1 for both languages does not seem to make a difference. On the other hand, %V2 presents very significant differences between both AmE and BP; the former in vowel reduction and the latter in vowel lengthening.

As mentioned earlier in this paper, ΔV represents the standard deviation, that is to say, the variability that happens in speech rhythm (“vowels” herein). Ramus et. al. (1999) attests that ΔV , for example, is smaller if vocalic intervals are less variable in duration, in other words, more regular. In this research, we present a significant rate of variability for the $\mu\Delta V2$ when comparing AmE and BP speakers. On the contrary, the $\mu\Delta V1$, does not produce a significant effect in terms of variability for both groups as shown in table 2.

We shall conclude so far that BP speakers produced a highly-variate speech rate when *in-clash vowel* is into account. This might have reflected a prosodic transfer from BP in terms of the choice for clash resolution strategy even considering the NPL. For V1 production, BP and AmE speakers are normally-distributed as pointed out in table 2. We take the assumptions for the following remarks:

- The effect of AmE V1s shows that even receiving the stress from V2 via BM, they maintain low ΔV s and so, low V-rates;
- AmE %V2s are significantly reduced (table 2) and also present low variability on their rates as well as V1s;
- BP keeps low $\mu\%V1$ and high $\mu\%V2$ values and significant high variability results are only reflected *between* both categories but not *in* the same vowel one. There is a lack of vowel reduction in V2 from BP but not for AmE speakers which reflects their choice for SDB in clash environment;
- In the concern of $\mu\Delta V$, AmE and BP speakers do not vary significantly, on the contrary, the means in table 2 show low variability mainly because AmE reduce significantly V2 as long as BP speakers lengthen it keeping equivalence when collated each other’s values.

5 Conclusions

We conclude from this study that there is preference from BP speakers in favor of SDB rather than stress retraction (AmE speakers) when facing stress clash environment - in English as L2 and L1 respectively. These early findings show that through acoustic phonetic analysis, there is a tendency of BP to align syllable-timing in relation to clash solving strategies for speech rhythm opposing the stress-timing alignment for AmE language.

We primarily conclude that there exists an influence of rhythm on the production of English from BP speakers. Perception and acoustic analysis may differ in some degree especially when we take into account the discrete *vs.* dynamic measurements of rhythm in languages. We do agree that for future work it would be necessary to run a perception test in order to check if there are significant differences between production and perception by means of choosing a strategy for stress clash resolution.

In relation to the percept of eurhythmy, Grabe & Warren, (1995); Vogel, Bunnell & Hoskins, (1995) tried to find phonetic evidence for the application of the Rhythm Rule rather

than beat addition intervening clash. Both studies agree that the perception of stress shift may be explained by the relational character of stress: besides reinforcing the primary stress features of the clash-triggering syllable of the first word in sequences such as “thirTEEN MEN” (clash condition), which would allow the action of phonetic parameters for the first syllable of the first word, the alternative decrease in prominence in the second syllable of the first word would also favor the perception of an iamb. To testify these postulates, it is also suggested an addition of another response variable; the fundamental frequency ($F_{zero} - F_0$) and once again, we are able to analyze perception and acoustic evidence of BM/SDB for solving clashes.

Vogel et al, (1995) show that the duration effect is mainly restricted to the rhymes as universal properties (and that is where the nuclei play the most important role). The authors suggested that prior shift of stress does not seem to be required for implementing pitch accent alternation.

Taking stress clash into account, and the way it is phonetically solved by Brazilian speakers of English, it would be safe to conclude from our study that BP does not assume stress-timing features as was presumed from phonological literature since the early 1980s. On the contrary, BP accounts for syllable-size duration as a perceptual cue that can be attested acoustically.

Thus, the speaker that has L1 with syllable-timing characteristics – such as BP - demonstrates a strong tendency to carry out all the syllables, words and P-phrases of utterances equally in terms of duration (and that includes not only content words but also, function words such as clitics and determinants which are weakened in English).

Our results confirm the prosodic transfer from L1 to L2 in cases of stress clash and stress resolution. We agree with Crystal (1976); Nooteboom (1997); Dupoux&Peperkamp (1999); Nunan (1999); Jenkins (2000); Gong (2002); Roach (2009); Gut, 2012; Avery & Ehrlich (2012), who claim that prosodic aspects may be fundamental in foreign language acquisition as prosodic cues operate at the interface with the syntax and are responsible for discrimination and segmentation of larger phonetic and phonological units in a language. In this sense, incorrect tracking location of prosodic borders – such as stress - can make an impact on speech intelligibility.

References

- ABOUSSALH, E. **Resolução de choques de acento no português brasileiro**: elementos para uma reflexão sobre a interface sintaxe-fonologia. Dissertação (Mestrado). Universidade Estadual de Campinas, 1997.
- ANDERSON, D.; SWEENEY, D.; WILLIAMS, T. **Statistics for Business and Economics** 9ed, Thomson Learning, Mason, 2005.
- AVERY, P.; EHRLICH, S. Problems of Selected Language Groups. In: P. Avery & S. Ehrlich (Eds), **Teaching American English Pronunciation**, pp. 111-157. New York: Oxford University Press, 2012.
- BARBOSA, P. “**Syllable-timing in Brazilian Portuguese**”: umacrítica a Roy Major. *DELTA*, 16/2 pp 369-402, 2000.
- BARBOSA, P. **Explaining Brazilian Portuguese Resistance to Stress Shift with a coupled-oscillator Model**. Caderno de Estudos Linguísticos, 43, pp 71-92, 2002.
- BARBOSA, P. **Incursões em torno do ritmo da Fala**. Campinas, FAPESP/Pontes Editores, 2006.
- BOERSMA, P; WEENINK, D. **Praat**: doing phonetics by computer (Version 6.0) <http://www.praat.org>, 2018.

- BISOL, L. **O Troqueu Silábico no Sistema Fonológico - Um Adendo ao Artigo de Plínio Barbosa**, *DELTA*, 16/2, pp. 403-413, 2000.
- CELCE-MURCIA, M.; BRINTON, D.; GOODWIN, J. **Teaching Pronunciation: A course book and reference guide** (2ed). New York: Cambridge University Press, 2010.
- COUNCIL OF EUROPE. **Common European framework of reference for languages: Learning, teaching, assessment**. Cambridge, U.K: Press Syndicate of the University of Cambridge. 2001.
- COUPER-KUHLEN, E. **English Speech Rhythm: form and function in everyday verbal interaction**, Philadelphia: John Benjamins Publishing, 1993.
- CRYSTAL, D. **The future of Englishes**. *English Today*, 15/2, pp. 10-20, 1999.
- DAUER, R. **Phonetics and phonological components of rhythm**. In: Proceedings of 11th ICPPhS, pp. 447-450, 1987.
- DAUER, R. **Stress-timing and syllable-timing reanalyzed**, *Journal of Phonetics* 11, pp. 51-62, 1983.
- DUPOUX, E.; PEPPERKAMP, S. **Fossil markers of language development: phonological 'deafnesses' in adult speech processing**. Oxford: Oxford University Press, 1999.
- GIBBON, D.; GUT, U. **Measuring Speech Rhythm**. In: Proceedings of Eurospeech, Aalborg, pp 91-94, 2001.
- GONG, J. **Introducing English rhythm in Chinese EFL classrooms: a literature review**. *Post-Script*, 31, pp. 26-42, 2002.
- GRABE, E.; WARREN, P. **Stress Shift: do speakers do it or do listeners hear it?** In: B. Connell & A. Arvaniti (Eds), **Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV**, pp. 95-110, Cambridge: Cambridge University Press, 1995.
- FOX, J.; WEISBERG, S. **An R Companion to Applied Regression**, 3ed, Los Angeles, SAGE Publications Inc., 2019.
- FRAGOZO, C. **Aquisição de Regras Fonológicas do Inglês por Falantes de Português Brasileiro**. Tese (Doutorado), Universidade de São Paulo, São Paulo, 2017.
- GUT, U. **Rhythm in L2 speech**. *Speech and Language Technology*, 14/15, pp. 83-94, 2012.
- HAYES, B. **Metrical Stress Theory: Principles and Case Studies**. Chicago: University of Chicago Press, 1995.
- HAYES, B. **The phonology of rhythm in English**. *Linguistic Inquiry*, 15/1, pp.33-74, 1984.
- HILLENBRAND, J. **Static and Dynamic Approaches to Vowel Perception**. In: G. S. Morrison and P. F. Assmann (Eds.), **Vowel Inherent Spectral Change**, pp. 9-30. Berlin: Heidelberg, 2013.
- JENKINS, J. **The phonology of English as an International Language**. Oxford: Oxford University Press, 2000.
- JONES, D. **An Outline of English Phonetics**. Cambridge: W. Heffer & Sons, 1918.
- LIBERMAN, M.; PRINCE, A. **On stress and linguistic rhythm**. *Linguistics Inquiry* 8. pp. 249-336, 1977.
- LOWIE, W.; SELTON, B. **Essential Statistics for Applied Linguistics**, Palgrave, Macmillan, New York, 2013.
- MAJOR, R. **Stress and rhythm in Brazilian Portuguese**. *Language*, 61, pp. 259-282, 1985.
- MAJOR, R. **Stress-Timing in Brazilian Portuguese**. *Journal of Phonetics*, 9, pp. 343-351, 1981.
- MARCUS, S. **Perceptual centres**. Tese (Doutorado), Cambridge, University of Cambridge, 1976
- MEHLER, J.; DUPOUX, E.; NAZZI, T.; DEHAENE-LAMBERTZ, G. **Coping with linguistic diversity: The infant's viewpoint**. In: J. Morgan & K. Demuth, **Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition**, pp 101-106. Mahwah:

- Lawrence Erlbaum Associates. University of Alberta. Indiana University Linguistics Club, 1996.
- NESPOR, M. On the separation of prosodic and rhythmic phonology. In: S. Inkelas & D. Zec (Eds), **The phonology-syntax connection**, pp. 243-258. Chicago: UCP, 1990.
- NESPOR, M.; VOGEL, I. **Prosodic Phonology: with a new forward**. Mouton de Gruyter, Berlin, 2007.
- NOOTEBOON, S. The Prosody of Speech: Melody and Rhythm. In: W. Hardalstle and J. Laver (Eds), **The handbook of phonetic sciences**, pp. 640-673. Oxford: Blackwell, 1997.
- NUNAN, D. **Second Language Teaching & Learning**. London: Heinle & Heinle, 1999.
- PIKE, K. **The Intonation of American English**, Ann Arbor: University of Michigan Publications, 1945.
- PURPURA, J. **Oxford Online Placement Test**. <https://www.oxfordenglishtesting.com/>, 2018.
- RAMUS, F., NESPOR, M. & MEHLER, J. **Correlates of linguistic rhythm in the speech signal**. *Cognition*, 73, pp 265–292, 1999.
- RIETVELD, T. HOUT, R. **Statistics in Language Research: Analysis of Variance**. Mouton de Gruyter, Berlin, 2005.
- ROACH, P. **Rhythm and Production: A lecture by Peter Roach**. https://www.youtube.com/watch?v=Xh_fiP6AT_I, 2009.
- ROACH, P. **English Phonetics and Phonology: A Practical Course**, 3ed. Cambridge: Cambridge University Press, 2005.
- SANDALO, F.; TRUCKENBRODT, H. **Some Notes on Phonological Phrasing in Brazilian Portuguese**. *MIT Working Papers In Linguistics*, 42, pp. 285-310, 2001.
- SANTOS, R. **Categorias Sintáticas Vazias e Retração de Acento em Português Brasileiro**. *DELTA*, 18, pp 67-86, 2002.
- SANTOS, R. **Traces, pro and Stress Retraction in Brazilian Portuguese**. *Journal of Portuguese Linguistics*, 2/2, pp 101-113, 2003.
- SANTOS, R. **Retração acentual e verbos transitivos com leitura intransitiva**. *Revista de Estudos da Linguagem*, 12/2 pp. 101-113, 2004.
- SARMAH, P.; GOGOI, V.; WILSHIRE, C. **Thai English: Rhythm and Vowels**. *English World-Wide*, 30, pp. 196–217, 2009.
- SELKIRK, E. **Phonology and Syntax: The relation between sound and structure**. Cambridge Mass: MIT Press, 1984.
- SILVA Jr, L. **Regression and Variance Model Maker version 1.0**. Script for R, 2019.
- SILVA Jr, L. **Interferências Rítmicas do Português Brasileiro no Inglês como L2. O choque acentual**. Tese (Doutorado), Universidade Federal da Paraíba, 2013.
- SILVA Jr, L.; BARBOSA, P. **Metrics and Acoustics Extractor version 1.0**. Script for Praat, 2019.
- VOGEL, I.; BUNNELI, T.; HOSKINS, S. The phonology and phonetics of the Rhythm Rule. In: B. Connell & A. Arvaniti (Eds), **Phonology and Phonetics Evidence**, *Papers in Laboratory Phonology IV*, pp. 111-127. Cambridge: Cambridge University Press, 1995.
- YAO, Y. Closure duration and VOT of word-initial voiceless plosives in English in spontaneous connected speech. In: **UC Berkeley Phonology Lab Annual Report**, pp. 183-225, 2007.

Recebido em 02/07/2019

Aceito em 10/08/2019