

RELATO DE PESQUISA

# Retração de acento: produção dos parâmetros acústicos do acento em inglês por aprendizes brasileiros



OPEN ACCESS

EDITADO POR

- Luma da Silva Miranda (ELTE)
- Carolina Gomes da Silva (UFPB)
- Manuela Carnaval (UFRJ)

AVALIADO POR

- Ronaldo Manguiera Lima Jr. (UnB)
- Amanda Post (UFJ)

SOBRE OS AUTORES

- Carina Fragozo  
Coleta de dados, Análise de dados, Discussão e Escrita.
- Raquel Santos  
Análise de dados, Discussão e Escrita.
- Andressa Toni  
Análise de dados, Discussão e Escrita.

DATAS

- Recebido: 15/10/2023
- Aceito: 22/03/2023
- Publicado: 27/05/2024

COMO CITAR

Fragozo, C.; Santos, R.; Toni, A. (2024). Retração de acento: produção dos parâmetros acústicos do acento em inglês por aprendizes brasileiros. *Revista da Abralín*, v. 22, n. 1, p. 6-36, 2024.

Carina FRAGOZO

Universidade de São Paulo (USP)

Raquel SANTOS

Universidade de São Paulo (USP)

Andressa TONI

Universidade Estadual do Centro-Oeste (UNICENTRO)

RESUMO

Esta pesquisa investiga como a retração de acento se manifesta acusticamente na fala de aprendizes brasileiros de inglês, de diferentes níveis de proficiência, e na fala de nativos de inglês. Especificamente, nosso objetivo é observar se os aprendizes falantes de português brasileiro modificariam a duração (uma pista de acento mais proeminente no português que no inglês) de sílabas em contexto de choque acentual (e.g. *thirTEEN MEN* produzido com a sílaba forte na primeira sílaba, *THIRteen MEN*). Foram analisados contextos de choque acentual, contextos sem choque acentual (*thirTEEN baNAnas*) e contextos-controle de palavra isolada (*thirTEEN*). A amostra é composta por trinta falantes brasileiros de inglês divididos em três níveis de proficiência (básico, intermediário e avançado), além de sete falantes nativos, que constituíram o grupo de controle. A classificação dos dados foi realizada em duas etapas: uma verificação perceptual e uma verificação acústica. Os resultados apontam que a duração foi manipulada pelos aprendizes de inglês e pelos nativos; que há uma diferença estatisticamente significativa entre o nível básico e demais níveis; e que nos contextos de encontro acentual tanto a sílaba

pretônica quanto a tônica eram mais longas do que em suas contrapartes em contexto sem encontro acentual.

ABSTRACT

This paper investigates how stress shift manifests acoustically in the speech of Brazilian learners of English. The study compares different levels of proficiency versus the speech of native English speakers in three contexts: stress clash (e.g. thirTEEN MEN), no stress clash (thirTEEN baNAnas), and isolated words as control contexts (thirTEEN). Given the rhythmic and stress differences between English and Brazilian Portuguese, our goal is to observe whether Brazilian Portuguese speakers would manipulate the duration (the main acoustic cue for word stress in Portuguese) of syllables in a stress clash context in English (e.g. thirTEEN MEN, where stress can shift to the first syllable, THIRteen MEN). We analyze data of 30 Brazilian learners of English, divided into three proficiency levels (basic, intermediate and advanced), as well as data of 7 native speakers. Data classification was carried out in two phases: first a perceptual analysis and then an acoustic. The results point out that duration was manipulated by both by English learners and native speakers; that there is a statistically significant difference between the basic level and other levels; and that in contexts with an accentual cluster, both the stressed and the pretonic syllable were longer than in its counterpart in a context without an accentual cluster.

PALAVRAS-CHAVE

Retração de acento. Inglês como L2. Duração. Fonética Acústica.

KEYWORDS

Stress shift. English as a foreign language. Duration. Acoustics.

RESUMO PARA NÃO ESPECIALISTAS

Este texto discute a produção de um fenômeno sonoro conhecido do inglês, a retração acentual (*thirTEEN MEN* produzido com a sílaba forte na primeira sílaba, *THIRteen MEN* 'treze homens'), quando ocorre um encontro de acentos de palavras. Em inglês, embora o acento de palavra seja principalmente marcado por variações de altura (pitch), ocorrem também alterações na duração que permitem definir a posição do acento. Utilizamos a duração, então, para identificar a posição do acento produzidas por brasileiros aprendendo inglês, em três níveis de proficiência (básico, intermediário e avançado). Nossos resultados mostraram que desde o nível básico ocorre a

manipulação da duração, mas que há uma diferença significativa entre este grupo e os demais (e também em relação a falantes nativos de inglês). Entretanto, o fenômeno da retração não foi encontrado como descrito na literatura (com a primeira sílaba tornando-se mais longa do que a sílaba final da palavra quando há um encontro acentual); de fato encontramos que ambas as sílabas se tornam mais longas quando comparadas com contexto em que não havia encontro acentual.

## Introdução

A produção de aspectos suprasegmentais de uma língua – como seu ritmo, seus processos de junção de palavras, seus parâmetros acentuais, sua entonação, e mesmo suas variações melódicas atitudinais (CRYSTAL, 1986) – configura-se, na aprendizagem de uma língua estrangeira, como uma acomodação de novos padrões prosódicos. A prosódia trata dos aspectos gramaticalizáveis de uma língua que se estendem por mais de um segmento, podendo abranger diferentes objetos fonológicos, como a sílaba, a palavra prosódica e os processos fonológicos. Muitas vezes, esses aspectos se mostram também em interface com outros componentes da gramática, como a sintaxe, a semântica, a pragmática e mesmo o discurso.

Neste artigo, discutiremos o fenômeno da retração de acento, que é um processo fonológico que modifica a posição do acento das palavras quando estes se encontram em sequência dentro de uma frase fonológica (NESPOR e VOGEL, 1986). A retração acentual é um processo de reajuste rítmico que visa garantir a eurritmia – isto é, a alternância entre sílabas fortes e fracas, evitando um choque acentual, i.e., duas sílabas fortes adjacentes (LIBERMAN e PRINCE, 1977; SELKIRK, 1984; NESPOR e VOGEL, 1986). Por exemplo, os sintagmas JeSUS CRISTo<sup>1</sup> e thirTEEN MEN ‘treze homens’ podem ser produzidos e percebidos como JEsus CRISto e THIRteen MEN, antecipando a proeminência acentual da primeira palavra a fim de evitar o choque acentual.

De acordo com Roach (2000, p.93), o acento pode ser estudado foneticamente de duas maneiras: sob o ponto de vista da produção ou da percepção. O acento é entendido como uma propriedade relacional, pois uma sílaba só pode ser considerada acentuada se comparada a outras sílabas não acentuadas. No que concerne à percepção, todas as sílabas acentuadas possuem como característica a proeminência, isto é, as sílabas acentuadas são reconhecidas como tais por serem mais proeminentes do que as sílabas átonas. Já no que concerne à produção, em termos articulatórios, o acento de uma sílaba pode realizar-se por meio da frequência fundamental (F0 ou *pitch*), da duração, da intensidade ou, menos comumente, dos formantes (REETZ e JONGMAN, 2009, p. 210). Uma sílaba acentuada apresenta tipicamente frequência fundamental mais alta, maior duração e/ou maior

---

<sup>1</sup> As maiúsculas representam a sílaba acentuada das palavras.

intensidade em comparação com sílabas átonas. Fry (1958) e Bolinger (1986) afirmam que o correlato mais importante para a definição de uma sílaba acentuada no inglês é o *pitch*. Reetz e Jongman (2009), estudando dados de fala sintética, apontam que o F0 (*pitch*) e a duração são os correlatos mais importantes para a identificação do acento no inglês; e Van Heuven (2018), analisando os dados de Fry (1955), mostra que a duração é um bom correlato para o acento e seus resultados mostram 98% de alterações na duração a depender da tonicidade da sílaba. Por outro lado, no português brasileiro (doravante PB), a duração é apontada como o parâmetro acústico mais importante para a definição do acento lexical (MAJOR, 1992; MASSINI, 1992; BARBOSA, 2000; FERREIRA, 2008). Isto faz com que brasileiros muitas vezes tenham dificuldade em reconhecer a sílaba tônica de uma palavra em inglês, assim como falantes de inglês têm a mesma dificuldade com o português (ALBINI; KLUGE; SILVA, 2019; MODESTO; BARBOSA, 2019).

Além dos parâmetros acústicos de acento, as línguas também diferem com relação ao ritmo. Há línguas que apresentam ritmo acentual, as quais apresentam intervalos regulares entre os acentos independentemente do número de sílabas entre eles, como o russo e o árabe; e outras que possuem ritmo silábico, nas quais os intervalos entre acentos aumentam de acordo com o número de sílabas entre eles, como o espanhol, o francês e o italiano. Diferentemente do inglês, cujo ritmo é normalmente caracterizado como acentual, o PB apresenta características tanto do ritmo acentual quanto do silábico (ABAURRE-GNERRE, 1981; MASSINI, 1992).

Considerando as diferenças rítmicas e acentuais do português e do inglês,<sup>2</sup> este estudo visa observar como a retração de acento se manifesta na fala de aprendizes brasileiros de inglês, de diferentes níveis de proficiência, comparando os aspectos acústicos do processo com aqueles constatados na fala de falantes nativos de inglês. Nosso objetivo é observar se os aprendizes falantes de português brasileiro modificariam a duração de sílabas em contexto de choque acentual, uma vez que este parâmetro não é o principal para a acentuação de palavra no inglês.

O artigo está organizado como se segue: a seção 1 traz um breve panorama sobre os estudos anteriores sobre a retração acentual em inglês e sua aquisição por falantes brasileiros. A seção 2 descreve a metodologia utilizada no estudo experimental conduzido pela presente pesquisa para comparar as produções acentuais de falantes nativos e dos aprendizes de L2. Na seção 3, os resultados são descritos e analisados estatisticamente enquanto que, na seção 4, discutimos as

---

<sup>2</sup> Partimos neste trabalho das constatações de Abousalh (1997), Sândalo e Truckenbrodt (2002), Gravina e Fernandes-Svartman (2013) de que a retração de acento existe no PB.

Gravina e Fernandes-Svartman (2013) analisam a retração acentual no português brasileiro em contextos como *jornal hoje*. As autoras observam a duração em um contexto de encontro acentual a depender se houve a percepção de uma retração acentual. Seus resultados apontam que nos casos em que perceptualmente se indicava uma retração acentual, ocorreu uma menor duração da primeira sílaba do contexto (no exemplo, *nal*) e, inversamente, uma duração maior da sílaba imediatamente anterior ao contexto de encontro (no exemplo, *jor*). Tenani (2002) traz resultados semelhantes sobre a duração nos casos de degeminação em contexto de encontro acentual.

Referimos o leitor aos trabalhos de Barbosa (2002), Gayer e Collischoon (2007) e Madureira (2002) para uma diferente visão sobre o assunto, mas chamamos a atenção de que assumir ou não a existência da retração acentual no PB não invalida a discussão aqui desenvolvida que é saber se os falantes de PB apresentam esta manipulação de duração no inglês como L2.



Entretanto, quando a vogal da sílaba que receberia o acento movido é preenchida por um *schwa*, como na sequência *maroon coat* (/məRU:N KOʊT/), o choque acentual não pode ser desfeito via retração acentual porque em inglês a vogal reduzida não pode receber o acento (cf. Levey 1999).

Para verificar se há uma preferência pela alternância entre sílabas fortes e fracas sob o ponto de vista fonético, Levey (1999) buscou investigar as características fonéticas do choque acentual com base em três possibilidades: a) o acento primário da última sílaba da primeira palavra do choque seria movido para a sílaba anterior, caracterizando a retração acentual (ex: *RAcoon COAT* ‘casaco de guaxinin’); b) o acento primário na sílaba final seria reduzido (isto é, o acento da sílaba final seria produzido de forma mais fraca quando se compara a mesma palavra produzida sem uma outra que crie o contexto de choque); e c) o choque acentual seria evitado através do uso de acento de *pitch* em uma sílaba que antecede ou sucede o choque na frase. A autora analisou tanto contextos de choque acentual quanto contextos sem choque, e sua verificação perceptual encontrou que a resolução do choque acentual ocorreu em menos de 30% dos casos, ou seja, temos aqui um processo opcional, que ocorre em um terço das produções. Já a verificação acústica apontou que a frequência fundamental foi a única pista que poderia influenciar a percepção da retração de acento, mas não foram encontradas evidências suficientes para afirmar que o F0 é, de fato, o correlato acústico responsável pela retração. Esses resultados vão ao encontro da sugestão de Grabe e Warren (1995) de que a retração de acento é, na verdade, um fenômeno de natureza perceptual, e não acústica. Os autores desenvolveram um experimento de percepção em que os participantes deveriam apontar a sílaba acentuada em uma série de palavras. Quando solicitados para identificar o acento em sequências de choque acentual como *thirTEEN MEN*, os participantes perceberam uma retração acentual, ou seja, *THIRteen MEN*. Quando a mesma gravação da palavra *THIRteen* foi apresentada isoladamente, entretanto, os participantes identificaram o acento como *thirTEEN*. Segundo Kimball e Cole (2014), esse resultado sugere que os falantes de inglês estão condicionados a perceber uma alternância entre sílabas fortes e fracas, mesmo que uma verificação acústica indique que essa alternância não tenha sido, de fato, produzida.

Em relação à aprendizagem da retração de acento do inglês por falantes brasileiros, Silva Jr. (2013) e Silva Jr e Scarpa (2019) são os únicos, até onde sabemos, que compararam a produção de sequências de choque acentual nas duas línguas em questão, buscando verificar a aplicação ou não da retração acentual. Silva Jr (2013) investigou as produções de cinco informantes: três brasileiros falantes de inglês como L2 (grupo experimental) e dois norte-americanos falantes de português como L2 (grupo de controle). Os informantes brasileiros foram divididos em três níveis de proficiência: C1 (*advanced*), B2 (*high intermediate*) e B1 (*low intermediate*). O corpus contou com sequências de palavras com choque acentual em inglês (ex: *thirTEEN BALLS*) ou em português (ex: *MULHER Ótima*). Os resultados mostraram que a inserção de um curto pulso rítmico silencioso (*silent demi-beat*) (uma pausa, em termos não técnicos) foi a estratégia mais empregada pelos falantes brasileiros para a resolução do choque em ambas as línguas. Os falantes norteamericanos, por outro lado, preferiram a estratégia de retração de acento. Em um segundo experimento, um americano (falante de português) julgou a produção dos brasileiros produzindo as sequências em inglês, e um brasileiro (falante de inglês) julgou a produção de norteamericanos produzindo as sequências em português.

Os resultados indicaram que esses ouvintes tenderam a perceber os intervalos acentuais como sendo mais regulares do que eles realmente são quando analisados acusticamente. Silva Jr e Scarpa (2019) discutem especificamente a duração das vogais no encontro acentual e na sílaba pré-encontro, nos dados de produção de três brasileiros (nível avançado e intermediário-avançado) e três norteamericanos. Os autores fazem uma análise sintagmática, comparando a duração de V1/V2 (vogal pré-encontro/vogal da primeira sílaba do encontro). Seus resultados apontam que os falantes brasileiros, além de acrescentar uma batida silenciosa entre as palavras do contexto de encontro acentual, aumentam a duração da vogal da primeira sílaba do encontro (como os autores não comparam com contextos sem encontro, infere-se que este alongamento deve ser em relação ao que fazem os falantes norteamericanos). Os autores concluem que falantes nativos de português não assumem os traços de línguas acentuais, como é o caso do inglês, mantendo as características de uma língua mais silábica, como é o caso do português.

Os resultados de Silva Jr (2013) e Silva Jr e Scarpa (2019) são instigantes, mas derivam de um número reduzido de informantes de inglês como L2. Com o estudo experimental descrito a seguir, expandimos a quantidade de informantes. Além disso, esperamos observar se, do ponto de vista perceptual, a retração de acento é apreendida nos mesmos contextos na fala de falantes nativos e não nativos. Do ponto de vista acústico, analisamos se o uso da duração sofre algum tipo de alteração em contextos de choque acentual em relação a um contexto sem encontro, isto é, comparamos a produção dos brasileiros com suas próprias produções, em outros contextos sonoros; e também se a manipulação deste parâmetro varia a depender da fluência dos aprendizes.

## 2. Metodologia

Para verificar se haveria alguma alteração na duração, um experimento eliciou a leitura de vinte e uma palavras oxítonas em contextos de frase com choque acentual (ex: *fifTEEN MEN* ‘quinze homens’), de frase sem choque acentual (ex: *fifTEEN poTatoes* ‘quinze batatas’) e de palavra isolada (ex: *fifTEEN* ‘quinze’) – cf. Quadro 1. Com as palavras isoladas era possível saber se o informante sabia a posição do acento; e a produção em sequências com e sem choque acentual permitia observar as modificações na palavra alvo.<sup>3</sup>

Somente foram consideradas para análise as palavras que o informante produzia corretamente a acentuação oxítona na palavra isolada – assim, casos em que era inserida uma epêntese final (como *complete* produzida como [kõ.'pli.tjĩ] – além da vogal plena na primeira sílaba) não foram analisados

<sup>3</sup> Apesar de a retração não ser possível em palavras como *police* e *complete* na fala de nativos por conterem a vogal reduzida *schwa* [ə] na sílaba que receberia o acento em caso de retração ([kəm'pli:t], [pə'li:s]), incluímos esses exemplos no experimento porque os falantes brasileiros, sobretudo no nível básico, tendem a produzir vogais plenas nesses contextos (*complete* [kom'pli:t] - ['kom.plit]; *police* [po'li:s] - [pə'li:s]), o que possibilita a análise da retração nesses casos.

pois ao introduzir a epêntese, o acento deixava de ser final (esses casos ocorreram majoritariamente no nível básico).

Palavra isolada	Sequências com choque acentual	Sequências sem choque acentual
1. thirTEEN	1. thirTEEN PENCils	1. thirTEEN poTAtos
2. fourTEEN	2. fourTEEN WomEn	2. fourTEEN baNAnas
3. fifTEEN	3. fifTEEN GIRLS	3. fifTEEN toMAtos
4. sixTEEN	4. sixTEEN CHAIRS	4. sixTEEN imiTAtions
5. sevenTEEN	5. sevenTEEN YEARS	5. sevenTEEN paPAYas
6. eighTEEN	6. eighTEEN CHILDren	6. eighTEEN adVENTures
7. nineTEEN	7. nineTEEN BOYS	7. nineTEEN eXAMples
8. unKIND	8. unKIND Comment	8. unKIND reVENge
9. kangaROO	9. kangaROO KIM	9. kangaROO MeLlssa
10. TenneSSEE	10. TenneSSEE PEOPle	10. TenneSSEE volcAno
11. groTESQUE	11. groTESQUE PICTures	11. groTESQUE deCEption
12. roBUST	12. roBUST BABies	12. roBUST umbRElla
13. bamBOO	13. bamBOO BRACelets	13. bamBOO maTERials
14. poLICE	14. poLICE Officer	14. poLICE conVENTion
15. disCRETE	15. disCRETE Areas	15. disCRETE beGInning
16. comPLETE	16. comPLETE PAPER	16. comPLETE compUTers
17. Bel-AIR	17. Bel-AIR BOY	17. Bel-AIR ceLEbrity
18. U2	18. U2 SONG	18. U2 celeBRation
19. reTAKE (v.)	19. reTAKE COURSE	19. reTAKE examiNAtion
20. exPRESS	20. exPRESS TRAIN	20. exPRESS transforMAtions
21. disLIKE	21. disLIKE CHOcolate	21. disLIKE perFEctionism
	22. disLIKE POWer	22. disLIKE conFUSion
	23. disLIKE PROblems	23. disLIKE poliTiCIans

Quadro 1 – Palavras e Sequências de Palavras para a Verificação da Retração de Acento  
 Fonte: Adaptado de Fragozo (2017)

As sequências de palavras foram inseridas em frases-veículo com frases fonológicas do mesmo tamanho, conforme (3),<sup>4</sup> de forma a controlar a influência do contexto sintático ou prosódico sobre a retração acentual.

- (3) a) [I saw] [grotESQUE PICTures] [in the park]. ‘eu vi desenhos grotescos no parque.’  
 b) [I have] [thirTEEN poTAtoes] [at home]. ‘eu tenho treze batatas em casa.’  
 c) [I saw] [kangaROO KIM] [last night]. ‘eu vi o canguru Kim ontem à noite.’  
 d) [I saw] [bamBOO BRACelets] [last night]. ‘eu vi pulseiras de bambu ontem à noite.’

As frases-alvo foram apresentadas em ordem aleatória e representavam um terço das sentenças apresentadas no experimento. A apresentação se deu através de sentenças escritas em *slides* do *Power Point*, cada sentença em um *slide*. Dada a quantidade de sentenças, os participantes faziam pausas (estabelecidas pelo experimentador, quando percebia que uma entonação da forma de lista

<sup>4</sup> Cumpre notar que uma vez que as palavras deveriam ser conhecidas pelos participantes do experimento, nem sempre os sintagmas analisados tinham a mesma quantidade de sílabas. De qualquer forma, a quantidade de sílabas até a primeira sílaba da segunda palavra do contexto em análise era sempre a mesma. Além disso, a quantidade de sílabas das palavras analisadas, se dissílabas ou trissílabas, não apresentou efeito significativo.



era utilizada). Foram consideradas ocorrências de retração os casos em que o informante apresentou o padrão oxítono na palavra isolada (ex: fifTEEN, sevenTEEN) e o padrão paroxítono (ex: FIFteen MEN) ou proparoxítono (ex: SEventeen MEN) na mesma palavra em frases com ou sem sequência de choque acentual. Foi considerado como não retração quando perceptualmente o acento mantinha-se na última sílaba ou quando acusticamente encontramos valores maiores na última sílaba da palavra que nas anteriores. Para a verificação perceptual, foram utilizados os mesmos procedimentos aplicados por Levey (1999) e Stander (2007): duas pessoas (a primeira autora e uma professora de inglês, também brasileira, com alto grau de treinamento em fonética e fonologia do inglês, mas alheia a pesquisas linguísticas) sinalizaram independentemente sua percepção de onde ocorria o acento em cada palavra do corpus. Nos casos em que houve discordância, ambas discutiram até que se chegasse um consenso a respeito das sílabas acentuadas.

Além da variação no contexto em análise (apresentar ou não encontro acentual), foi controlada a inserção de pausa (que solucionaria o encontro acentual de uma outra maneira diferente da retração).

Foi controlado também o nível de fluência fonológica dos participantes, e a idade de início dos estudos de inglês (infância ou adolescência).<sup>5</sup> Participaram do experimento sete falantes nativos de inglês e trinta falantes de PB aprendizes de inglês como L2. A amostra analisada é do tipo aleatória estratificada, em que a população é dividida em subgrupos (estratos) e uma subamostra é selecionada a partir de cada estrato da população (LEVIN e FOX, 2004). Os falantes nativos de inglês deveriam ter vivido a maior parte de suas vidas em seus países de origem. Dentre esses falantes nativos, dois são ingleses e vivem no Brasil há cerca de quinze anos, dois são australianos e vivem no Brasil há cerca de oito anos e três são estadunidenses e nunca vieram ao Brasil.<sup>6</sup> Já os aprendizes de L2 deveriam ser filhos de pais brasileiros monolíngues; não falar nenhuma outra língua estrangeira além do inglês; e não deveriam ter morado em país falante de inglês por mais de um mês. Eles foram selecionados a partir de um teste realizado com base no *Common European Framework of Reference for Languages* (documento elaborado pelo Conselho Europeu que permite a avaliação do conhecimento linguístico de aprendizes de qualquer idioma - COUNCIL OF EUROPE, 2001) para comporem três grupos (básico, intermediário e avançado, com dez participantes cada). Este nivelamento ocorreu em duas etapas: inicialmente, um teste *online* com vinte e cinco questões de múltipla escolha indicou uma estimativa do nível do participante; em seguida, um teste oral avaliou aspectos fonológicos, a partir do qual os grupos foram estabelecidos (isto é, em casos em que os informantes tivessem um nível intermediário em sintaxe, mas básico em fonologia, os informantes foram classificados como de nível básico).

Para a análise das medidas acústicas, decidiu-se por medir as duas sílabas da primeira palavra e trabalhar com dois tipos de duração, absoluta e relativa, para a visualização do fenômeno sob diferentes perspectivas. A duração absoluta permite controlar as diferenças inerentes às consoantes e

---

<sup>5</sup> Para a discussão sobre os efeitos do nível de fluência e da idade de início dos estudos em inglês, cf. Santos e Fragozo (2020).

<sup>6</sup> Até onde sabemos, para o estudo dos fenômenos em questão, as diferentes variantes da língua inglesa não apresentam qualquer tipo de influência.

vogais utilizadas nas palavras do teste (já que o tipo de segmento e quantidade de segmentos por sílaba não foi um fator controlado na escolha dos estímulos) e permite observar qual tipo de estratégia de alongamento os falantes aplicam nos contextos de retração: se encurtam a sílaba originalmente tônica; se alongam a sílaba originalmente átona; ou se alongam uma sílaba e encurtam a outra.<sup>7</sup> Tais estratégias podem ser visualizadas de duas formas: 1) comparando a duração das sílabas de uma mesma palavra e verificando qual delas se mostra mais longa em cada contexto acentual (ou seja, comparar *bam* com *boo*) – por exemplo, como em Silveira Jr e Scarpa 2019; 2) comparando a duração de uma mesma sílaba nos diferentes contextos acentuais (ou seja, comparar *bam* em contexto de choque, sem choque e isolado; e o mesmo com *boo*) – como em Gravina e Fernandes-Svartman (2013). Em seguida, trabalhamos com a duração relativa (sílabas/palavra), que permite controlar efeitos como a taxa de elocução do falante e os alongamentos característicos do fim de enunciado nos contextos de palavra isolada.

Na seção a seguir, inicialmente apresentamos os resultados da estatística descritiva utilizando valores de duração absolutos, que foram explorados via testes de Wilcoxon e Kruskal-Wallis. Nestes, comparamos as propriedades acústicas das sílabas nos três contextos acentuais analisados. Em seguida, um modelo linear de efeitos mistos (com falante como variável aleatória) compara a duração relativa das sílabas em relação a toda palavra (por exemplo, *bam/bamboo* e *boo/bamboo*) em cada contexto. As análises foram realizadas no software estatístico SAS 9.4 e foi fixado o nível de significância  $\alpha = 0,05$  para todas as análises.<sup>8</sup>

### 3. Resultados

Como explicitado na metodologia, previamente à análise acústica, uma análise perceptual foi conduzida a fim de determinar a taxa de aplicação da regra (cf. FRAGOZO e SANTOS, 2021). Interessantemente, mesmo nos contextos em que não havia choque acentual (ex: *thirTEEN poTatoes*), a retração foi percebida em taxas expressivas, tanto na fala dos falantes nativos (51,6%) quanto na dos aprendizes (27,6%). De forma análoga, nos casos de contextos de encontro acentual (ex: *thirTEEN MEN*), perceptualmente nem todos os contextos foram resolvidos – seja por meio da retração acentual, seja pelo uso de outras estratégias, tanto na fala nativa quanto não nativa. Percebeu-se retração acentual em contextos de encontro em 21,7% do total de casos (396/1378) na fala dos aprendizes,

---

<sup>7</sup> Observe-se que para a retração acentual ocorrer, não é necessário que a última sílaba da primeira palavra seja encurtada, basta que a duração da primeira sílaba da primeira palavra seja alongada. No caso de *unKIND COMment*, por exemplo, basta que *un* seja produzida com duração maior, tendo *kind* sido produzida com menor duração ou não.

Um parecerista sugere que a duração relativa seja calculada por meio da subtração das durações silábicas (por exemplo, *KIND – un*). Embora esta seja uma medida interessante, já que revelaria pelo valor positivo ou negativo do resultado qual sílaba teve sua duração alterada pelo falante, a subtração dificultaria a compreensão dos resultados se ambas as sílabas tiverem sido alteradas. Optamos, assim, pela observação dos valores absolutos em combinação com o cálculo relativo sílaba/palavra.

<sup>8</sup> A análise estatística foi conduzida pela Dra. Andressa Kutschenko Nahas, CONRE 9066-A.

contra 50,06% dos casos (133/322) na fala nativa. Tal resultado, embora seja inconsistente com modelo métrico de Hayes (1984), foi também encontrado em outros estudos, como em Levey (1999), sugerindo que a resolução do choque acentual pode depender de outros fatores além da métrica e do domínio prosódico.

Os resultados perceptuais também mostraram que a retração passa a ser mais frequente conforme o avanço do nível de proficiência, ou seja, na produção dos falantes de nível avançado a retração de acento é mais perceptível do que nas produções dos falantes de nível básico e intermediário. Observa-se que, no nível avançado, os falantes brasileiros apresentam uma taxa de retração muito aproximada à dos falantes nativos, conforme a Tabela 1:

Nível de Proficiência	Contexto por nível de proficiência			
	Com choque		Sem choque	
	n	%	n	%
<b>Básico</b>	42	18%	38	17%
<b>Intermediário</b>	57	25%	57	25%
<b>Avançado</b>	107	47%	95	41%
<b>Nativo</b>	80	49,7%	83	51,6%

Tabela 1 – Taxas de retração acentual na análise perceptual por contexto: com ou sem choque acentual  
Fonte: Adaptado de Fragozo (2017)

A explicação para esses resultados está relacionada à fluência na produção: como a retração de acento é uma regra que ocorre na fala natural, e informantes de nível básico nem sempre são capazes de realizar a leitura das frases de modo natural, a fala mais pausada acaba bloqueando a aplicação da regra. Mas interessa-nos saber se os informantes estão se comportando da mesma forma que os falantes nativos. Como dissemos, embora a retração acentual seja um processo de aplicação variável, portanto nem sempre realizado, é relevante investigar como esse processo se revela acusticamente nos casos em que há ou não a percepção da retração, já que, conforme a literatura, português e inglês se baseiam em pistas acústicas majoritárias distintas para a realização do acento. Vejamos então, a seguir, como a duração se comporta na fala de falantes nativos e não nativos.

### 3.1 Medidas de duração na retração acentual por aprendizes de inglês como L2 e falantes nativos

Para tratar de verificar a ocorrência ou não da retração de acento, inicialmente examinou-se se todas as palavras do teste seriam produzidas com duração maior na última sílaba – ou seja, com um padrão oxítono – no contexto de palavra isolada. Em todos os níveis de proficiência dos aprendizes de L2, confirmou-se o padrão oxítono também observado no contexto controle.

Primeiramente, dada a opcionalidade do processo de retração acentual, os resultados abaixo descrevem, palavra a palavra, o que acontece com as medidas de duração. A Tabela 2 mostra os nativos, as médias absolutas e o desvio padrão (DP) da duração nas sílabas das mesmas palavras

produzidas nos contextos de sequência com choque acentual (ex: *bamBOO BRACElets*), de sequência sem choque acentual (ex: *bamBOO maTErials*) e a palavra isolada (ex: *bamBOO*). Os asteriscos na coluna à direita indicam os casos em que o valor de p, obtido através do teste Kruskal-Wallis, se mostrou significativo. Como se observa, a diferença de duração na última sílaba de todas as palavras nos três contextos foi estatisticamente significativa, exceto na palavra *seventeen*. Tal diferença aponta que a duração em um (ou mais) dos três contextos (palavra isolada, com choque, sem choque) se destaca em relação aos demais – ainda não sendo possível, portanto, apontar se houve ou não manipulação significativa em virtude da retração de acento. Em todos os casos, a duração foi maior em palavra isolada.

Duração (em milissegundos) - Falantes Nativos								
Palavra	Sílaba	Frase com Choque		Frase sem Choque		Palavra Isolada		valor-p
		Média	DP	Média	DP	Média	DP	
<i>bamboo</i>	bam	212,29	67,16	189,00	38,75	226,71	40,50	0,0857
	boo	185,57	33,70	179,86	70,54	311,14	54,95	0,0037 *
<i>Bel-Air</i>	Bel	173,71	28,94	162,29	36,04	228,86	42,71	0,0101 *
	Air	194,43	37,58	196,00	60,48	298,71	31,34	0,0035 *
<i>complete</i>	com	163,43	36,71	143,43	20,61	162,71	26,25	0,3876
	plete	242,29	26,85	270,71	39,25	454,14	80,28	0,0008 *
<i>discrete</i>	dis	189,29	35,26	187,86	28,61	208,00	34,42	0,5735
	crete	263,71	54,17	261,71	46,52	434,86	60,63	0,0016 *
<i>Dislike</i>	dis	200,14	27,91	200,62	23,32	250,00	45,88	0,0240 *
	like	228,00	30,06	243,05	35,69	375,86	70,32	0,0002 *
<i>eighteen</i>	eigh	143,86	35,27	132,57	23,86	169,86	22,27	0,0413 *
	teen	255,71	33,79	214,29	39,35	394,43	44,73	0,0005 *
<i>express</i>	ex	167,00	52,81	176,43	28,26	188,29	46,51	0,7322
	press	321,00	57,49	326,14	66,46	521,29	53,91	0,0012 *
<i>Fifteen</i>	fif	203,86	53,65	208,57	50,18	227,00	59,02	0,7341
	teen	272,14	37,86	252,14	67,99	397,00	84,71	0,0030 *
<i>fourteen</i>	four	213,71	50,93	256,43	40,36	232,00	44,43	0,3526
	teen	261,29	45,54	258,86	49,54	414,71	44,01	0,0012 *
<i>grotesque</i>	gro	132,43	24,00	154,71	42,50	139,57	25,95	0,5372
	tesque	356,00	73,89	413,29	60,50	503,71	34,10	0,0021 *
<i>kangaroo</i>	kan	215,86	28,40	208,29	30,77	203,43	39,84	0,7627
	ga roo	106,00 149,43	20,72 29,02	108,29 134,00	21,91 26,87	122,43 275,14	27,77 53,06	0,4037 0,0011 *
<i>nineteen</i>	nine	230,57	37,81	215,71	50,18	244,00	47,37	0,4571
	teen	306,71	47,64	207,86	42,24	387,43	55,47	0,0008 *
<i>Police</i>	po	70,14	18,55	110,86	30,54	135,14	31,70	0,0030 *
	lice	208,00	51,60	241,00	49,35	471,00	76,87	0,0007 *
<i>Retake</i>	re	144,14	16,06	132,86	19,82	150,29	40,86	0,3967
	take	289,29	34,55	276,00	25,83	413,57	45,88	0,0011 *
<i>robust</i>	ro	130,71	22,46	141,71	26,18	132,71	29,81	0,5743
	bust	332,14	63,16	397,86	108,22	499,29	47,30	0,0047 *
<i>seventeen</i>	se	184,43	32,70	195,86	34,46	198,71	29,05	0,6819

	ven	142,86	18,99	138,86	19,65	157,14	27,17	0,3333	
	teen	253,14	51,64	289,29	86,75	347,29	61,71	0,0539	
sixteen	six	302,71	64,73	305,14	58,76	303,43	32,05	0,9564	
	teen	249,71	51,71	225,43	73,00	353,57	23,99	0,0050	*
Tennessee	te	125,43	28,24	130,71	19,31	125,00	23,61	0,8175	
	nne	103,00	14,36	106,14	13,72	115,00	10,30	0,2849	
	ssee	241,57	64,89	239,00	31,03	369,29	39,71	0,0027	*
thirteen	thir	207,71	57,64	212,00	77,54	178,29	20,85	0,6223	
	teen	257,14	49,79	250,29	50,98	441,29	80,31	0,0021	*
U2	U	137,57	17,39	129,14	29,99	155,71	33,67	0,5274	
	2	223,71	31,34	251,43	37,58	387,29	50,03	0,0007	*
unkind	un	178,71	35,65	187,00	50,71	167,86	39,65	0,6196	
	kind	353,00	78,43	372,86	103,62	494,43	58,51	0,0180	*

Tabela 2 – Médias de duração em contexto isolado, com choque acentual ou sem choque – Nativos

Fonte: Adaptado de Fragozo (2017)

Vejamos a diferença de duração na última sílaba em contextos de frase com choque e sem choque acentual. Das vinte e uma palavras, onze apresentaram média de duração absoluta maior na última sílaba em contexto de choque, indicando uma não-retração, enquanto outras dez palavras apresentaram uma duração um pouco maior na última sílaba em contexto *sem* choque, indicando uma retração acentual.

Consideremos agora a duração das sílabas que, no caso de retração acentual, receberiam o acento, isto é, a penúltima e antepenúltima sílabas - como em *Bel-Air* e *kangaroo*. Se observarmos todas as palavras que não apresentaram diferença significativa neste contexto (todas exceto *Bel-Air*, *dislike*, *eighteen* e *police*), sete palavras apresentaram duração maior em casos de choque acentual, enquanto que em outras nove, a duração nessas sílabas foi maior em frases sem sequência de choque acentual. Nos casos de *complete* e *police*, que não deveriam sofrer retração acentual já que a vogal pretônica é um *schwa*, observamos que, enquanto em *complete* a diferença não foi significativa em contexto com e sem encontro acentual (de acordo com o esperado), em *police* encontrou-se uma diferença significativa - mas não na direção esperada, já que a sílaba foi mais curta em contexto de choque.

Em relação à estratégia de alongamento utilizada, percebe-se que, no geral, a sílaba originalmente tônica permanece com duração maior tanto nos contextos sem choque quanto nos contextos com choque. Entretanto, se comparamos a duração das sílabas pretônicas nos contextos com e sem choque acentual, encontramos uma tendência à maior duração nos contextos de choque, mesmo que esta duração seja menor do que a duração da sílaba tônica.

As tabelas 3, 4 e 5 a seguir trazem os resultados, separados por nível de proficiência, referentes às médias de duração empregadas pelos trinta falantes de inglês como L2. Chamamos a atenção de que uma estratégia comum de brasileiros adquirindo inglês é acrescentar uma epêntese, o que reorganiza a estrutura silábica do inglês para o PB. Lembramos que palavras que foram produzidas com inserção de epêntese (ex: [kō.'pli.t.jí] *complete*) não foram incluídas nas análises dos correlatos acústicos de retração de acento, pois as médias tiveram de ser calculadas com base em sílabas *iguais* para a comparação dos correlatos acústicos nos três contextos.

Observa-se que, no nível básico – Tabela 3, as ocorrências de diferença estatisticamente significativa nos três contextos analisados são poucas: apenas cinco de vinte e uma palavras. As palavras *robust*, *kangaroo*, *express* e *police* foram as únicas que apresentaram diferença significativa na última sílaba, e nos três casos, a duração mais longa ocorreu no contexto de palavra isolada. No caso das palavras em que a sílaba anterior era preenchida por um *schwa*, observe que, mesmo não sendo estatisticamente significativa a diferença na duração, a última sílaba é mais curta em contexto de encontro acentual do que nos demais contextos.

Duração (em Milissegundos) - Falantes de Nível Básico								
Palavra	Sílaba	Frase com Choque		Frase sem Choque		Palavra Isolada		valor-p
		Média	DP	Média	DP	Média	DP	
bamboo	bam	240,20	45,86	237,30	57,76	246,70	24,49	0,9399
	boo	253,00	78,43	211,80	85,19	220,70	62,21	0,3764
Bel-Air	Bel	278,60	74,76	247,70	47,56	254,40	36,99	0,7158
	Air	369,80	112,72	321,30	63,51	347,80	74,01	0,4535
complete	com	170,30	18,67	157,10	14,82	175,00	27,29	0,3235
	plete	311,40	86,63	340,60	38,44	414,90	96,00	0,0524
discrete	dis	183,00	36,68	171,60	24,79	194,80	43,27	0,3946
	crete	408,40	67,97	389,10	71,20	453,30	53,73	0,0652
dislike	dis	192,52	59,17	207,72	42,61	217,40	59,44	0,2310
	like	385,77	88,81	409,28	80,29	441,20	62,69	0,0924
eighteen	eigh	189,33	38,35	205,70	42,14	205,90	48,73	0,7216
	teen	357,56	78,84	330,20	59,40	371,50	95,60	0,4924
express	ex	224,00	62,28	210,40	66,40	186,70	62,87	0,5318
	press	420,20	81,80	429,70	65,48	554,90	94,42	0,0052 *
fifteen	fif	265,80	59,51	263,90	52,95	292,00	58,10	0,4534
	teen	297,10	96,77	260,90	41,79	315,60	76,45	0,2867
fourteen	four	281,40	40,36	318,10	47,19	264,90	50,27	0,0708
	teen	380,10	66,71	385,60	65,75	331,70	61,23	0,1875
grotesque	gro	178,33	35,64	221,78	54,87	161,89	34,21	0,0417 *
	tesque	565,44	123,01	600,11	109,36	551,67	74,34	0,5975
Kangaroo	kan	234,50	44,26	203,50	26,97	211,50	16,36	0,1750
	ga	152,60	65,66	121,70	38,02	146,80	54,40	0,4935
	roo	209,80	61,44	122,70	36,16	213,90	43,65	0,0007 *
Nineteen	nine	367,45	108,58	380,90	73,08	330,80	57,75	0,2586
	teen	333,27	107,36	303,60	69,44	355,50	92,55	0,4124
Police	po	146,40	26,75	175,00	31,31	155,70	26,01	0,1398
	lice	282,40	50,91	332,70	62,63	367,10	42,88	0,0082 *
Retake	re	148,30	35,54	142,40	28,07	142,10	45,65	0,9900
	take	470,00	80,13	425,70	76,12	513,20	102,49	0,1357
Robust	ro	163,20	43,89	233,90	62,59	162,40	45,06	0,0269 *
	bust	405,20	75,58	390,60	80,61	472,00	56,26	0,0405 *
Seventeen	se	228,00	43,81	257,50	55,63	234,90	42,00	0,4061
	ven	173,90	21,14	181,00	31,08	178,10	19,36	0,5649
Sixteen	teen	265,90	72,54	302,60	59,93	305,50	89,08	0,5820
	six	358,00	70,85	389,50	44,37	359,20	53,39	0,3556
Tennessee	teen	257,00	56,37	316,50	34,73	282,10	65,32	0,1329
	te	106,00	24,53	108,10	22,66	106,20	20,85	0,9598
	nne	175,80	76,66	167,20	60,71	168,60	67,03	0,9900
Thirteen	ssee	272,90	92,06	265,30	79,06	329,60	58,36	0,1332
	thir	198,09	22,06	211,00	37,96	189,60	28,96	0,3286
U2	teen	321,64	103,48	288,33	68,16	319,70	68,75	0,5448
	U	193,00	36,74	172,33	41,30	163,50	33,78	0,3002

	2	294,09	63,27	344,22	75,65	334,20	54,66	0,1937
<i>Unkind</i>	un	188,70	47,59	181,10	47,71	189,10	41,66	0,7174
	kind	416,10	144,24	441,60	127,42	444,60	102,53	0,8269

Tabela 3 – Médias de duração no nível Básico - Inglês como L2  
Fonte: Adaptado de Fragozo (2017)

A Tabela 4 apresenta as médias de duração dos falantes de nível intermediário. Percebe-se que há diferença significativa em dezoito palavras do experimento, bem mais que no nível básico, e essa diferença significativa ocorreu na última sílaba em dezesseis palavras. Em todas essas palavras a última sílaba é mais longa no contexto de palavra isolada e que, na maioria dos casos, é mais longa no contexto de frase com choque do que no contexto sem choque. Nas palavras em que a penúltima sílaba apresentou diferença de duração significativa, a média mais alta ocorreu no contexto de palavra isolada (exceto na palavra *unkind*) e, dentre os dois contextos de frase, em sete casos foi mais alta no contexto de choque (exceto *fourteen*). Nos casos das palavras que apresentavam *schwa* na sílaba anterior, encontramos que *police* não apresentou diferença entre a duração nem da sílaba portadora do *schwa*, nem na última sílaba da palavra, ao se comparar os contextos com e sem encontro acentual. No caso de *complete*, a primeira sílaba foi mais longa em contexto com encontro acentual.

Duração (em Milissegundos) - Falantes de Nível Intermediário									
Palavra	Sílaba	Frase com Choque		Frase sem Choque		Palavra Isolada		valor-p	
		Média	DP	Média	DP	Média	DP		
<i>Bamboo</i>	bam	206,30	26,07	198,00	29,77	261,00	51,71	0,0023	*
	boo	217,10	82,20	141,10	47,33	246,60	55,29	0,0029	*
<i>Bel-Air</i>	Bel	240,00	57,34	220,70	30,88	280,30	50,05	0,0280	*
	Air	268,00	89,32	266,30	96,08	336,50	75,20	0,0886	
<i>Complete</i>	com	159,50	25,05	145,70	23,50	180,80	17,96	0,0152	*
	plete	269,60	47,44	296,70	76,31	423,50	77,80	0,0017	*
<i>Discrete</i>	dis	189,80	39,50	171,60	44,94	213,10	23,13	0,0492	*
	crete	315,20	36,87	316,20	115,34	444,10	72,54	0,0014	*
<i>Dislike</i>	dis	169,35	41,69	165,24	39,53	216,30	55,82	0,0228	*
	like	280,77	62,88	280,07	59,98	424,50	72,80	0,0001	*
<i>Eighteen</i>	eigh	152,30	37,63	160,40	27,31	185,10	26,39	0,0615	
	teen	335,20	104,95	269,80	47,69	362,70	57,50	0,0120	*
<i>Express</i>	ex	212,70	30,53	245,20	35,16	253,80	61,51	0,1023	
	press	361,40	43,42	360,80	74,47	534,50	96,77	0,0002	*
<i>Fifteen</i>	fif	229,30	25,85	250,09	43,68	243,90	42,08	0,5211	
	teen	303,50	51,96	257,09	58,90	367,40	75,39	0,0025	*
<i>fourteen</i>	four	220,45	22,16	279,55	56,97	235,90	40,00	0,0153	*
	teen	317,91	100,50	298,91	71,48	360,60	40,82	0,0682	
<i>grotesque</i>	gro	197,30	54,14	265,80	90,74	197,10	65,27	0,1591	
	tesque	416,50	72,55	498,00	163,99	529,00	85,80	0,0500	
<i>kangaroo</i>	kan	195,10	26,95	191,90	26,41	210,10	22,59	0,0966	
	ga	116,60	38,53	113,50	26,96	136,10	26,83	0,1220	
	roo	244,00	90,92	151,30	42,83	267,50	42,59	0,0005	*
<i>nineteen</i>	nine	296,40	58,48	291,70	31,96	310,70	73,62	0,9252	
	teen	291,20	44,68	279,70	76,89	326,40	55,02	0,0876	
<i>Police</i>	po	122,30	30,18	134,80	47,38	146,90	35,87	0,3020	
	lice	260,10	43,62	267,70	69,72	403,20	68,56	0,0003	*
<i>Retake</i>	re	150,20	46,02	154,90	42,35	173,10	55,82	0,5916	
	take	340,10	66,61	325,50	39,74	506,10	70,88	0,0001	*
<i>Robust</i>	ro	145,00	36,94	165,80	43,69	180,10	38,58	0,0569	

	bust	353,70	55,78	375,00	93,59	495,70	81,67	0,0030	*
seventeen	se	202,40	31,57	216,40	36,75	225,00	47,00	0,5364	
	ven	156,20	25,45	155,90	36,42	190,10	30,19	0,0541	
	teen	247,60	55,56	253,10	58,99	300,90	45,85	0,0894	
Sixteen	six	332,17	49,91	301,57	24,62	327,50	43,11	0,2003	
	teen	249,50	62,20	229,43	28,77	315,90	54,45	0,0091	*
Tennessee	te	99,90	26,29	107,70	30,96	119,80	42,90	0,7013	
	nne	153,60	46,91	128,10	49,62	162,60	81,14	0,0504	
	ssee	222,90	115,72	213,40	61,90	347,50	80,14	0,0042	*
thirteen	thir	176,56	25,73	174,56	52,08	194,80	33,70	0,1349	
	teen	269,67	47,22	271,00	48,48	356,90	52,51	0,0011	*
U2	U	174,30	28,80	158,60	29,54	201,00	33,48	0,0130	*
	2	216,50	34,43	229,10	43,39	337,30	66,23	0,0004	*
Unkind	un	193,10	29,53	156,40	28,73	178,20	38,75	0,0477	*
	kind	440,10	127,70	368,10	57,49	499,80	70,35	0,0108	*

Tabela 4 – Médias de duração no nível Intermediário - Inglês como L2

Fonte: Adaptado de Fragozo (2017)

A Tabela 5 apresenta as médias de duração dos informantes do nível avançado, que apresentaram diferença significativa em todas as vinte e uma palavras do experimento, sobretudo na última sílaba. Nas palavras que apresentaram diferença estatisticamente significativa na última sílaba, observa-se que, assim como no nível intermediário, as durações são mais altas no contexto de palavra isolada. Ao comparar os contextos com mais de uma palavra, em onze delas a última sílaba é mais longa no contexto de frase com choque, enquanto em outras dez, a sílaba mais longa ocorre em contexto sem choque. Observa-se que para *complete*, a duração das duas sílabas é bem próxima nos contextos com e sem encontro, o mesmo que ocorre com a última sílaba de *police*. A primeira sílaba de *police* é mais longa em contexto sem choque. Em ambos os casos, observa-se um comportamento mais próximo das regras do inglês que nos níveis básico e intermediário.

Duração (em Milissegundos) - Falantes de Nível Avançado								
Palavra	Sílaba	Frase com Choque		Frase sem Choque		Palavra Isolada		valor-p
		Média	DP	Média	DP	Média	DP	
bamboo	bam	243,30	34,87	241,10	70,28	297,40	56,23	0,1066
	boo	197,80	68,12	172,40	61,99	286,30	62,01	0,0031
Bel-Air	Bel	231,60	43,62	210,00	47,06	267,10	30,66	0,0162
	Air	216,00	72,23	220,80	72,81	307,80	71,53	0,0212
complete	com	162,20	31,72	160,70	39,77	184,10	30,49	0,1624
	plete	245,40	55,51	272,20	68,41	503,50	76,72	0,0001
discrete	dis	203,00	36,66	223,00	86,14	210,90	41,36	0,8304
	crete	328,60	79,20	310,20	90,45	497,80	63,71	0,0003
Dislike	dis	188,10	41,80	187,43	37,36	228,10	37,42	0,0218
	like	268,77	60,25	299,27	80,46	467,10	43,94	0,0000
eighteen	eigh	172,40	40,27	164,90	31,94	175,10	34,47	0,7067
	teen	288,20	58,16	280,90	60,35	444,90	64,82	0,0001
express	ex	241,20	29,45	239,20	24,47	233,10	54,58	0,4790
	press	369,80	51,13	386,50	76,24	596,80	83,59	0,0001
Fifteen	fif	224,40	53,69	224,10	47,36	252,70	55,56	0,3911
	teen	311,60	88,18	243,50	40,48	417,20	83,68	0,0001
fourteen	four	236,10	46,10	249,90	69,18	239,90	38,86	0,9451
	teen	305,70	53,46	303,90	78,45	431,70	75,53	0,0016



<i>grotesque</i>	gro	162,80	36,74	192,20	44,57	187,90	41,25	0,2177	*
	tesque	427,10	57,71	470,20	71,63	589,90	69,93	0,0002	
<i>kangaroo</i>	kan	224,56	37,26	216,10	35,66	227,00	30,88	0,6413	*
	ga	117,56	19,68	101,40	21,34	122,40	32,64	0,1581	
	roo	163,22	57,66	144,10	37,32	283,00	87,06	0,0007	
<i>nineteen</i>	nine	299,30	37,42	304,78	41,55	322,10	50,35	0,5078	*
	teen	307,50	66,30	272,33	55,47	386,30	53,99	0,0024	
<i>Police</i>	po	88,10	23,71	101,20	29,42	122,50	27,05	0,0330	*
	lice	284,70	54,17	282,10	68,70	530,50	113,52	0,0001	
<i>Retake</i>	re	137,40	54,40	139,60	25,92	138,70	34,79	0,7404	*
	take	397,10	40,60	356,80	63,23	537,60	59,09	0,0001	
<i>Robust</i>	ro	186,50	39,56	208,60	50,59	177,30	49,45	0,3988	*
	bust	284,80	36,85	351,60	70,11	514,70	78,63	0,0000	
<i>seventeen</i>	se	213,70	50,64	231,90	43,35	217,90	40,31	0,5854	*
	ven	181,80	19,29	169,80	26,40	203,20	16,80	0,0050	
<i>Sixteen</i>	teen	254,70	74,40	283,50	72,71	372,40	64,25	0,0060	*
	six	337,91	70,00	354,56	76,67	349,40	49,75	0,6363	
<i>Tennessee</i>	teen	237,08	42,86	248,50	46,28	369,00	72,80	0,0003	*
	te	125,70	35,24	122,30	31,05	143,80	38,96	0,5226	
	nne	130,50	24,38	125,80	22,11	136,20	19,39	0,7534	
<i>thirteen</i>	ssee	251,90	73,68	246,90	76,85	392,70	89,07	0,0027	*
	thir	199,40	56,99	187,30	46,01	221,90	64,01	0,4253	
<i>U2</i>	teen	275,10	63,26	264,00	54,81	419,00	73,32	0,0004	*
	U	191,40	57,01	168,50	32,72	184,20	52,01	0,6827	
<i>Unkind</i>	2	234,50	64,45	271,60	51,50	396,80	57,13	0,0001	*
	un	169,33	29,38	181,63	20,74	174,10	31,17	0,6416	
	kind	323,67	45,19	380,75	66,09	535,00	69,55	0,0000	*

Tabela 5 - Médias de duração no nível Avançado - Inglês como L2

Fonte: Adaptado de Fragozo (2017)

Em suma, ao comparar a duração absoluta das sílabas dentro de uma mesma palavra (comparação sintagmática), notamos que foram poucos os casos em que a penúltima ou antepenúltima sílabas apresentaram duração maior que aquela que originalmente deveria ser a sílaba mais longa – e parte dessas palavras apresentou tal alongamento mesmo em contexto sem choque. A comparação mais produtiva para examinar as manipulações acústicas realizadas pelos falantes foi a comparação paradigmática, em que se observa a duração de uma mesma sílaba em contexto de choque *versus* em contexto sem choque. Neste cotejo, foi possível observar a ocorrência variável do fenômeno da retração acentual: cerca de metade dos estímulos apresentou pretônicas com duração absoluta maior nos contextos com choque. Destaca-se, então, que a duração foi um correlato acústico manipulado pelos aprendizes brasileiros de inglês nos contextos de choque acentual. Destaca-se, também, que nesta manipulação a sílaba passível de alongamento não se torna mais longa que a sílaba tônica, e nem a sílaba tônica torna-se mais curta; a sílaba passível de retração torna-se mais longa em relação a ela mesma em contexto sem choque acentual. Além disso, notou-se que esta manipulação tem sua produtividade aumentada de acordo com o nível de proficiência do aprendiz. A seguir, utilizamos medidas de duração relativa para verificar, de maneira mais controlada, os comportamentos duracionais aqui observados, além de verificar sua relação com o fator nível de proficiência.

Inicialmente observem-se os Gráficos 1 e 2, que trazem a média da sílaba/palavra (tanto das pré-tônicas quanto das tônicas) nos contextos sem encontro acentual (EA 0) e com encontro acentual (EA 1), para todos os aprendizes:

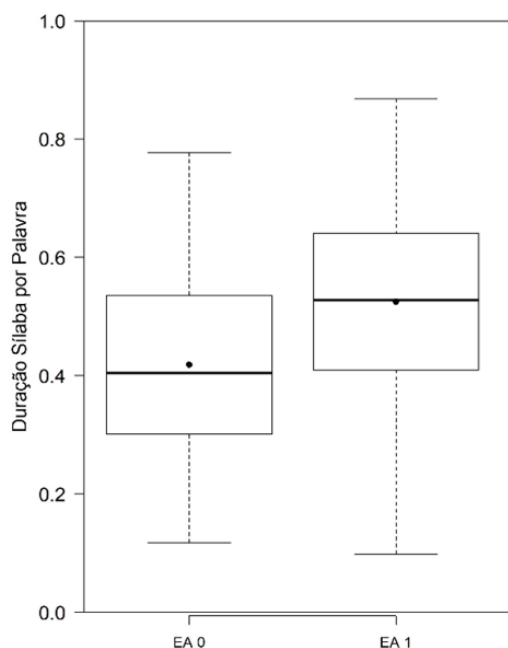


Gráfico 1 - Média de duração relativa por contexto acentual - Inglês como L2  
 Fonte: Adaptado de Fragozo (2017)

No Gráfico 1, é notável que os contextos de choque acentual (EA1) apresentam-se mais longos que os contextos sem choque (EA0) - e com base no observado também nas tabelas de duração absoluta, isso se deve a um alongamento tanto da sílaba passível de receber alongamento via retração acentual quanto da sílaba tônica. No contexto de choque acentual, portanto, todas as sílabas da palavra se mostram mais alongadas em comparação com o contexto sem choque.

Observe-se agora o Gráfico 2, que traz a média de duração de ambas as sílabas de algumas das palavras do experimento para ilustrar o comportamento da duração:

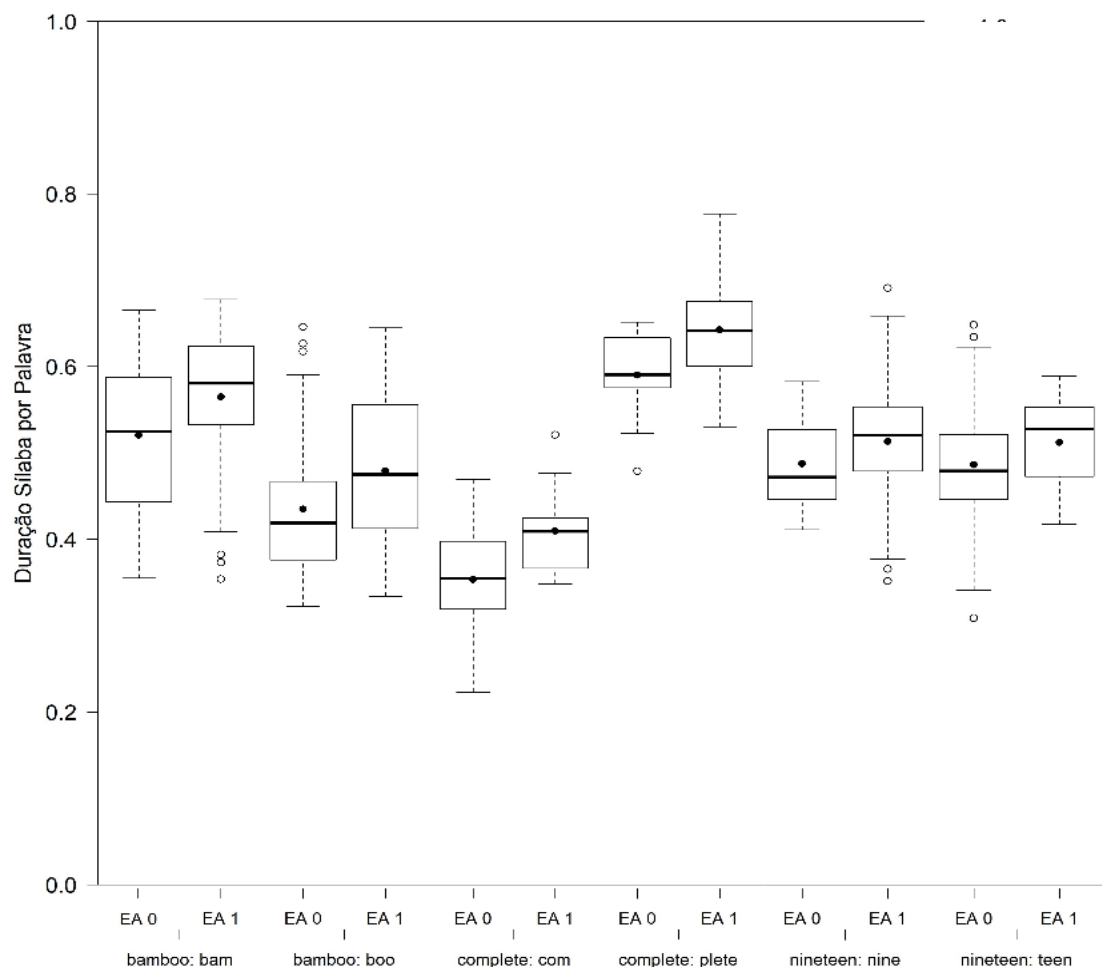


Gráfico 2 – Exemplos da média de duração relativa por contexto acentual - Inglês como L2  
 Fonte: Adaptado de Fragozo (2017)

No Gráfico 2 temos exemplo deste comportamento em três palavras: *bamboo*, *complete* e *nineteen*, para as duas sílabas em análise. Observe que para os dois exemplos, *Sílaba1/palavra* e *Sílaba2/palavra* são mais longas em contexto EA1 que em EA0. Entretanto, cada um dos exemplos ilustra uma possibilidade de diferente manipulação acústica observada nos dados: em *bamboo*, nota-se que tanto em EA0 quanto em EA1 *bam* é mais longo que *boo*, e ambas as sílabas são mais longas em EA1 que em EA0 (o que ocorreu em cinco palavras nos dados de duração relativa); em *complete*, embora as sílabas em contexto EA1 sejam mais longas que em EA0, a sílaba tônica, *plete*, se mantém mais longa que a pretônica, *com* (o que ocorreu em quinze palavras); e por fim, em *nineteen*, nota-se duração semelhante entre *nine* e *teen*, havendo alongamento de ambas em EA1 (somente nesta palavra).

Na análise inferencial, um modelo de efeitos mistos foi ajustado aos dados tomando a duração relativa (sílabas/palavra) como a variável dependente, “Participante” como efeito aleatório e os fatores “Encontro de Acento” (com ou sem choque), “Retração” (percebida na análise perceptual ou não), “Nível de Aquisição” (Básico, Intermediário, Avançado, Nativo), “Sílabas” (última, penúltima,

antepenúltima), “Tipo de Segmento Final” e “Inserção de Pausa” como variáveis predictoras. No presente artigo, enfocamos as variáveis “Encontro de Acento”, “Retração” e “Nível de aquisição”.

Vejamos inicialmente os resultados referentes à interação entre contexto acentual e percepção de retração (Tabela 6). Valores significativos foram observados tanto para o fator “Encontro de Acento” ( $p < 0,0001$ ) quanto para “Retração Percebida” ( $p < 0,0001$ ), e também para a interação entre estes fatores (Encontro de Acento \* Retração Percebida) ( $p < 0,0001$ ). Um teste *post hoc* apontou diferenças significativas entre todas as combinações da interação entre “Encontro de Acento” e “Retração Percebida”. Ou seja, existe manipulação acústica significativamente diferente na duração relativa dos contextos com choque e sem choque acentual - e que foram ou não percebidos com retração.

Least Squares Means							
Effect	Retração Percebida	Encontro de Acento	Estimate	Standard Error	DF	t Value	Pr >  t
Retração	Não		0.4900	0.003007	3481	162.98	<.0001
Retração	Sim		0.4488	0.002696	3481	166.47	<.0001
Encontro de*Retração	Não	Sem choque	0.3471	0.004313	3481	80.47	<.0001
Encontro de*Retração	Sim	Sem choque	0.4736	0.003637	3481	130.19	<.0001
Encontro de*Retração	Não	Com choque	0.6330	0.004190	3481	151.05	<.0001
Encontro de*Retração	Sim	Com choque	0.4241	0.003981	3481	106.55	<.0001

Tabela 6 – Resultados do modelo de efeitos mistos: Retração e Encontro de Acento.

Para todos os efeitos: DF: 3574; Alpha: 0.05

Fonte: as autoras (2024)

Vejamos agora os resultados no que se refere ao “Nível de Fluência”. O efeito isolado do “Nível” não é significativo ( $p = 0,992$ ), porém há diferença significativa nos contextos de “Encontro de Acento” ( $p < 0,0001$ ) e entre a interação “Encontro de Acento” e “Nível”. Ou seja, falantes de diferentes níveis de fluência apresentam duração relativa diferente nos contextos acentuais com choque e sem choque. A Tabela 7 mostra que essa significância ocorre tanto para os casos de encontro acentual quanto para os casos de não encontro acentual, para todos os níveis de fluência, e a Tabela 8 mostra a direção das interações.

Least Squares Means							
Effect	Nível	Encontro de Acento	Estimate	Standard Error	DF	t Value	Pr >  t
Nível	Avançado		0.4716	0.004631	3574	101.84	<.0001
Nível	Básico		0.4735	0.004626	3574	102.37	<.0001
Nível	Intermediário		0.4720	0.004616	3574	102.25	<.0001
Nível	Nativo		0.4722	0.005522	3574	85.50	<.0001
Encontro de Ac*Nível	Avançado	sem choque	0.4336	0.006344	3574	68.34	<.0001
Encontro de Ac*Nível	Básico	sem choque	0.4000	0.006338	3574	63.11	<.0001
Encontro de Ac*Nível	Intermediário	sem choque	0.4236	0.006326	3574	66.97	<.0001
Encontro de Ac*Nível	Nativo	sem choque	0.4323	0.007560	3574	57.18	<.0001
Encontro de Ac*Nível	Avançado	com choque	0.5096	0.006748	3574	75.53	<.0001

<b>Encontro de Ac*Nível</b>	Básico	com choque	0.5471	0.006740	3574	81.18	<.0001
<b>Encontro de Ac*Nível</b>	Intermediário	com choque	0.5204	0.006725	3574	77.38	<.0001
<b>Encontro de Ac*Nível</b>	Nativo	com choque	0.5120	0.008051	3574	63.59	<.0001

Tabela 7 – Resultados do modelo de efeitos mistos: Nível e Encontro de Acento. Para todos os efeitos: DF: 3574; Alpha: 0.05  
 Fonte: as autoras (2024)

Differences of Least Squares Means									
Effect	Nível	Encontro Acento	Nível	Encontro Acento	Estimate	Standard Error	DF	t Value	Pr >  t
<b>Encontro de Ac*Nível</b>	Avançado	com choque	Básico	com choque	-0.03749	0.009537	3574	-3.93	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Avançado	com choque	Intermediário	com choque	-0.01076	0.009527	3574	-1.13	0.2589
<b>Encontro de Ac*Nível</b>	Avançado	com choque	Nativo	com choque	-0.00235	0.01050	3574	-0.22	0.8228
<b>Encontro de Ac*Nível</b>	Básico	com choque	Intermediário	com choque	0.02673	0.009522	3574	2.81	<b>0.0050</b>
<b>Encontro de Ac*Nível</b>	Básico	com choque	Nativo	com choque	0.03514	0.01050	3574	3.35	<b>0.0008</b>
<b>Encontro de Ac*Nível</b>	Intermediário	com choque	Nativo	com choque	0.008404	0.01049	3574	0.80	0.4231
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Básico	sem choque	0.03360	0.008967	3574	3.75	<b>0.0002</b>
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Intermediário	sem choque	0.009932	0.008959	3574	1.11	0.2676
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Nativo	sem choque	0.001249	0.009869	3574	0.13	0.8993
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Intermediário	sem choque	-0.02367	0.008954	3574	-2.64	<b>0.0083</b>
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Nativo	sem choque	-0.03235	0.009866	3574	-3.28	<b>0.0011</b>
<b>Encontro de Ac*Nível</b>	Intermediário	sem choque	Nativo	sem choque	-0.00868	0.009858	3574	-0.88	0.3785
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Avançado	com choque	-0.07607	0.009261	3574	-8.21	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Básico	com choque	-0.1136	0.009256	3574	-12.27	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Intermediário	com choque	-0.08683	0.009245	3574	-9.39	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Avançado	sem choque	Nativo	com choque	-0.07843	0.01025	3574	-7.65	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Avançado	com choque	-0.1097	0.009257	3574	-11.85	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Básico	com choque	-0.1472	0.009252	3574	-15.91	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Intermediário	com choque	-0.1204	0.009241	3574	-13.03	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Básico	sem choque	Nativo	com choque	-0.1120	0.01025	3574	-10.93	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Intermediário	sem choque	Avançado	com choque	-0.08601	0.009249	3574	-9.30	<b>&lt;.0001</b>
<b>Encontro de Ac*Nível</b>	Intermediário	sem choque	Básico	com choque	-0.1235	0.009243	3574	-13.36	<b>&lt;.0001</b>

Encontro de Ac*Nível	Intermediário	sem choque	Intermediário	com choque	-0.09676	0.009233	3574	-10.48	<.0001
Encontro de Ac*Nível	Intermediário	sem choque	Nativo	com choque	-0.08836	0.01024	3574	-8.63	<.0001
Encontro de Ac*Nível	Nativo	sem choque	Avançado	com choque	-0.07732	0.01013	3574	-7.63	<.0001
Encontro de Ac*Nível	Nativo	sem choque	Básico	com choque	-0.1148	0.01013	3574	-11.34	<.0001
Encontro de Ac*Nível	Nativo	sem choque	Intermediário	com choque	-0.08808	0.01012	3574	-8.70	<.0001
Encontro de Ac*Nível	Nativo	sem choque	Nativo	com choque	-0.07968	0.01104	3574	-7.21	<.0001

Tabela 8 – Resultados da interação entre Nível e Encontro de Acento  
Fonte: as autoras (2024)

Na Tabela 8, tem-se uma comparação entre o nível de fluência e o contexto acentual, com choque (Encontro Acento com choque) e sem choque (Encontro Acento sem choque). Vejamos inicialmente os casos que comparam a duração nos contextos de encontro acentual. Observa-se que o nível Básico apresenta diferenças significativas na duração em relação a todos os demais níveis e ao modelo nativo. Já os níveis Intermediário e Avançado não apresentam diferenças significativas nem entre si, nem em relação aos falantes nativos. O mesmo se repete em contextos sem choque: novamente, somente as comparações com o nível Básico apresentam valores de duração significativamente diferentes. Entretanto, ao compararmos os contextos com choque e sem choque, observamos diferenças significativas entre todos os níveis, e mesmo dentro de cada nível – em todos os níveis e nos nativos, as sílabas analisadas foram mais longas nos contextos com encontro do que no contexto sem encontro acentual.

## 4. Discussão

Duas questões direcionam esta pesquisa: a manipulação da duração por brasileiros adquirindo o inglês em contextos com encontro acentual (*vs.* sua contraparte sem encontro) e se há variações nesta manipulação a depender do nível de fluência (proficiência fonológica).

Consideremos, inicialmente, os resultados referentes à duração absoluta, em que foram comparados os contextos de palavra isolada, contexto com encontro acentual e sem encontro acentual.

Observou-se que, em todas as palavras, a duração da última sílaba (ex: *nineteen*) foi significativamente mais alta nos contextos de palavra isolada do que em frases com e sem choque acentual. De acordo com Reetz e Jongman (2009, p. 216), esse aumento na duração da última sílaba de palavras em final de frase ou antes de uma pausa não é uma característica específica da língua inglesa, mas um fenômeno comum em muitas línguas.

De fato, estes resultados estão de acordo com os encontrados por Fougeron e Keating (1997), de que há uma relação entre a fronteira de domínio prosódico (cf. Nespor e Vogel 1986) e o alongamento

silaba final (mas ocorrendo em uma janela de três sílabas). Segundo os autores, uma sílaba que está em fronteira máxima de palavra é mais curta do que esta mesma sílaba que esteja em uma fronteira máxima mais alta, de frase fonológica, por exemplo, e este alongamento é incremental (quanto mais alto o nível prosódico, mais alongada a sílaba). Estes resultados indicam que o efeito de prolongamento não pode ser evitado, enquanto fator de confusão em um experimento, quando se inclui a palavra em uma sentença veículo, a menos que haja também o controle das fronteiras prosódicas.

Como explicitado na metodologia, a palavra isolada foi produzida não só sem uma outra palavra para contexto, mas também foi produzida fora de uma sentença-veículo. A razão para isso foi que as palavras alvo eram normalmente adjetivos ou numerais, e tentamos criar frases-veículos que não fossem artificiais (do tipo ‘diga\_de novo’), mas sim frases-veículos que formassem orações com os contextos analisados. Assim, analisando os contextos dentro das frases-veículos tínhamos:

- (4) a. [kangaroo]<sub>sintagma entoacional</sub>  
 b. [I saw [[kangaroo]<sub>palavra fonológica</sub> Kim]<sub>frase fonológica</sub> last night.]<sub>sintagma entoacional</sub>  
 c. [I saw [[kangaroo]<sub>palavra fonológica</sub> Melissa]<sub>frase fonológica</sub> last night.]<sub>sintagma entoacional</sub>

Em (4) acima observamos que a sílaba *roo* está em fronteira máxima de palavra fonológica (ou de grupo clítico, a depender da proposta de hierarquia prosódica que se assuma), nos contextos com e sem encontro, enquanto isoladamente está em fronteira de sintagma entoacional. Mesmo a sílaba *ka*, receptora de um movimento de acento, está na janela de alongamento detectada por Fougeron e Keating. Mesmo para o português brasileiro, onde este alongamento só foi significativamente encontrado no sintagma entoacional (e não nos níveis prosódicos mais baixos), este processo pode estar afetando os resultados aqui encontrados (cf. Santos e Leal 2008).

Ao analisarmos a duração relativa das sílabas em contexto com choque acentual entre os diferentes níveis de proficiência, observamos que somente o nível Básico apresentou diferenças significativas em relação aos demais. O mesmo ocorreu ao compararmos os contextos sem choque acentual. Num primeiro momento, seria possível imaginar que as diferenças significativas de duração encontradas nas tabelas 7 e 8 só são observadas no nível Básico porque neste nível os falantes naturalmente apresentam leitura mais lenta das palavras. Este, de fato, é um fenômeno observado mesmo nos dados de duração absoluta (como pode ser notado ao compararmos cada palavra entre os níveis nas Tabelas 3, 4 e 5).

Entretanto, ao observar os dados que comparam os contextos com choque e sem choque na Tabela 8 (que apresenta a análise por duração relativa), notamos que em todos os níveis a diferença passou a ser significativa – tanto ao compararmos os níveis Básico, Intermediário, Avançado e Nativo entre si quanto também ao realizarmos uma comparação intranível, Básico com Básico, Intermediário com Intermediário, Avançado com Avançado, Nativo com Nativo. Isso mostra que em todos os níveis de fluência existe manipulação significativa da duração quando comparamos contextos com e sem choque acentual. Inicialmente tratemos de como se deu esta manipulação.

Os resultados foram trazidos de duas maneiras: analisamos a duração absoluta e a duração relativa (sílabas/palavra) das sílabas da primeira palavra em contexto com encontro e sem encontro acentual. Utilizamos a duração absoluta para verificar qual tipo de manipulação acústica ocorreria em caso de retração acentual: se um encurtamento da sílaba tônica da palavra; se um alongamento da sílaba pretônica passível de receber a retração acentual; ou se ambas. Além disso, também lançamos mão de dois tipos de comparação entre as sílabas: comparar a duração entre as sílabas tônica e pretônica (comparação sintagmática), como em Silveira Jr & Scarpa (2019); e comparar a duração de cada sílaba em contexto com choque e sem choque acentual (comparação paradigmática), como em Gravina & Fernandes-Svartman (2013). Nossos resultados de duração absoluta indicam que a comparação paradigmática é aquela que melhor captura a retração acentual, havendo um aumento da duração da sílaba pretônica em contextos de choque em comparação a contextos sem choque. Passando à análise inferencial, que mobilizou a duração relativa, trouxemos resultados interessantes para a questão. O primeiro resultado, como no resultado com duração absoluta, é que todas as sílabas de todas as palavras em contexto de choque estão mais longas do que em contexto sem choque.

Silva Jr (2013) e Silva Jr & Scarpa (2019) sugerem que, para resolver o encontro acentual, os brasileiros fazem uso da inserção de um *silent demibeat* (curto pulso rítmico silencioso), que pode ser preenchido por uma pausa ou mesmo como alongamento da primeira sílaba tônica do contexto. Entretanto, esta estratégia afeta apenas a primeira sílaba do contexto e poderia, inclusive, ser resultado do nível de fluência dos informantes, em comparação com os nativos – e não falamos aqui simplesmente da taxa de elocução, mas de que o nível de fluência influi no mapeamento prosódico, resultando em alongamento em fronteiras prosódicas, como discutido acima. A maneira de descobrir se se trata realmente de uma estratégia de resolução de encontro acentual passa necessariamente por ver o que os aprendizes fazem em contexto sem encontro, o que fazemos aqui. E os resultados são de que há sim uma maior duração da primeira sílaba do contexto, corroborando os resultados de Silva Jr (2013) e Silva Jr & Scarpa (2019), mas também a maior duração da sílaba pré-encontro – o que impede então, que analisemos o fenômeno como uma inserção de *silent demibeat*. Também não podemos simplesmente dizer que nada ocorreu, que os aprendizes não fizeram nada, nenhuma modificação, pois neste caso os resultados não seriam estatisticamente significativos na comparação entre contexto com encontro e sem encontro, como efetivamente ocorreu (cf. Tabela 8 e Gráfico 1).

Sobre esta questão, temos hipóteses a serem perseguidas, mas que, dado o *corpus* aqui analisado, não podem ser investigadas e devem ficar como sugestão de investigações futuras. Esperávamos que, de alguma forma, em contexto de encontro acentual a primeira sílaba (a sílaba pretônica) acabasse por tornar-se mais longa do que a primeira sílaba do contexto. O que tivemos foi uma maior duração não em relação à tônica, mas sim entre sua contraparte em contexto sem encontro. Alguém poderia então sugerir que o que temos aqui é simplesmente uma forma de realçar esta palavra em um contexto especial. Sugerimos, entretanto, que pode ser o caso de que, mesmo que a primeira sílaba não seja mais longa do que a sílaba tônica, pode ser o caso de que a relação entre elas faça com que o ouvinte interprete como uma retração. Para isso, seriam necessários experimentos que manipulassem a duração da pretônica, alongando-a a diferentes proporções em relação à tônica, de



modo a buscar se existe uma relação de proporção entre estas duas sílabas para o que se percebe como retração acentual. Crucialmente, se a percepção do que é o acento se dá nesta relação de proporção entre as duas sílabas em uma palavra, esta percepção também ocorrerá em contextos sem encontro acentual.

Voltemo-nos agora para a segunda questão deste artigo, que era observar se haveria diferenças a depender do nível de fluência dos aprendizes. Lembramos que por nível de fluência os aprendizes foram classificados quanto a serem básico, intermediário e avançado quanto ao domínio fonológico.

Perceptualmente, observamos que mesmo os nativos não apresentaram uma alta taxa de retração acentual (Tabela 1). Além disso, dois outros fatos se destacaram: um crescendo na quantidade de retração (contexto de choque: nível básico 18%, intermediário 25%, avançado 47%, nativos 49,7%). Outro é que as taxas de percepção eram muito parecidas em contexto sem encontro, mas também apresentando este aumento na proporção de aplicação da retração (contexto sem choque: nível básico 17%, intermediário 25%, avançado 41%, nativos 51,6%). Como se observa, os resultados perceptuais para o nível avançado são bem parecidos com os resultados dos nativos.

Entretanto, ao olhar para a duração nos contextos com encontro versus sem encontro, diferenças significativas foram observadas em todos os níveis de fluência, tanto na duração absoluta com comparação paradigmática (cf. Tabelas 3, 4 e 5), quanto na análise da duração relativa (Tabelas 6, 7 e 8). Neste último caso, houve diferenças significativas na duração e na interação entre duração e nível. Na comparação entre níveis em contexto de encontro ou em contexto sem encontro, somente o nível Básico apresentou diferenças significativas com os demais níveis (cf. Tabela 8), tanto no contexto com encontro quanto no contexto sem encontro. Em suma, o que as medidas nos dizem é que a grande diferença no comportamento entre os diversos níveis em relação ao encontro acentual ocorre no nível básico.

## Considerações finais e caminhos futuros

O objetivo desta pesquisa foi observar se brasileiros adquirindo inglês como L2 manipulam da mesma forma que os nativos o parâmetro da duração em contextos de encontro acentual, onde estes aplicam o *stress shift*. Especificamente, estamos interessados se há alguma diferença no domínio deste parâmetro conforme o domínio do inglês dos aprendizes, medido aqui em três níveis: básico, intermediário e avançado, em comparação com um grupo de nativos.

Observaram-se muitas semelhanças entre os resultados dos falantes nativos e os falantes não nativos: 1) do ponto de vista perceptual, a retração ocorre tanto em contexto com choque ou sem choque em proporções muito aproximadas; 2) perceptualmente, a taxa de aplicação da retração pelos falantes de nível avançado (43,9%) foi bastante aproximada à taxa de aplicação dos falantes nativos (50%); 3) palavras lidas isoladamente apresentam durações mais longas do que palavras lidas em contexto de frase com choque ou sem choque; 4) acusticamente, os contextos de choque acentual

apresentam sílabas mais alongadas que os contextos sem choque – tanto a sílaba pretônica quanto a sílaba tônica da primeira palavra do encontro acentual.

Apesar de perceptualmente encontrarmos um movimento de crescimento na manipulação da duração, mostramos que acusticamente só há diferença significativa nos resultados apenas entre os participantes de nível Básico *vs.* Intermediário, Avançado e Nativos, quando comparamos cada contexto acentual isoladamente. Já quando cotejamos os contextos com *versus* sem choque, todos os níveis de fluência apresentam diferenças significativas, bem como os nativos. O crescimento perceptual não se reflete, portanto, num crescimento acústico da manipulação da duração.

Especificamente sobre a ocorrência da retração acentual, quando observamos a Tabela 2, com dados dos falantes nativos de inglês, somente em um caso a duração da sílaba pretônica excedeu a da sílaba tônica. Ainda assim, da mesma forma que para os aprendizes, não encontramos uma inversão na proporção da duração das sílabas dentro da palavra, mas uma maior duração tanto de sílabas pretônicas quanto tônicas em contexto de encontro acentual. Assim, diferentemente de Silva Jr (2019), não foram encontradas estratégias diferentes entre nativos e aprendizes. O fato de as pretônicas também serem alongadas tampouco nos permite concordar com a análise de Silva Jr e Scarpa (2019) de que os aprendizes fazem uso de uma estratégia de inserção de um *silent demibeat* preenchido por alongamento da última sílaba. Alternativamente, sugerimos que a explicação possa estar na relação de proporcionalidade de preenchimento da sílaba pretônica em relação à palavra, o que só pode ser investigado via um experimento futuro, com manipulação da duração.

## Informações complementares

Avaliação

Avaliação: <https://doi.org/10.25189/rabralin.v23i2.2193.R>

Editoras

Luma Miranda

Afiliação: Universidade Eötvös Loránd

ORCID: <https://orcid.org/0000-0002-5529-0338>

Manuella Carnaval

Afiliação: Universidade Federal do Rio de Janeiro

ORCID: <https://orcid.org/0000-0002-4321-5859>

Carolina Gomes da Silva

Afiliação: Universidade Federal da Paraíba

ORCID: <https://orcid.org/0000-0002-1490-0814>

## RODADAS DE AVALIAÇÃO

Avaliador 1: Ronaldo Manguiera Lima Júnior

Afiliação: Universidade de Brasília

ORCID: <https://orcid.org/0000-0002-8610-0306>

Avaliador 2: Amanda Post da Silveira

Afiliação: Universidade Federal de Jataí

ORCID: <https://orcid.org/0000-0002-9451-7005>

## AVALIADOR 1

Trata-se de um trabalho com tema relevante para a área de desenvolvimento fonético-fonológico de L2, com potencial contribuição teórica e empírica. No entanto, há questões metodológicas que podem ser revisadas e/ou esclarecidas a fim de conferir ao texto o grau de relevância que ele merece. Deixei comentários e sugestões ao longo do texto, mas registro aqui que 3 os pontos que considero mais importantes:

- É preciso especificar de modo mais explícito o(s) modelo(s) estatístico(s) ajustado(s), inclusive explicando as direções das interações encontradas;
- o ideal seria trabalhar com durações relativas (conforme prometido no início do texto) e valores de f0 normalizados;
- avalio que uma comparação das diferenças de duração e de f0 das sílabas que poderiam receber o acento em cada contexto seria uma análise mais direta do que comparar a duração e o f0 de cada sílaba com ela mesma em cada contexto.

Fico à disposição para qualquer esclarecimento necessário.

AVALIADOR 2

Caras autoras e editores,

Começo agradecendo a oportunidade de ler este trabalho que traz um tema interessantíssimo, que é o clash de acento em inglês como L1 e como L2. Concordo muito com as autoras, que descobriram que o fenômeno se trata muito mais de um aspecto perceptual da fonologia, talvez gerado pelos padrões fonológicos que trazemos no nosso inventário. Enfim, foi um prazer poder ler e levantar algumas questões na forma de diálogo com vocês. Para isso, faço o upload do arquivo com a minha revisão: o que está destacado em amarelo, tratam-se de palavras ou passagens a serem revistas, ou reformuladas; enquanto que o que está destacado em azul, são comentários e questionamentos meus a vocês, ou sugestões de acréscimo, ou substituição. Parabéns a todas pelo resultado e agradeço mais uma vez pela honra e satisfação de fazer essa leitura. Espero ter podido trazer alguma contribuição.

Grata mais uma vez, Amanda.

Conflito de Interesse

As autoras não têm conflitos de interesse a declarar.

Declaração de Disponibilidade de Dados

Os dados, códigos e materiais que suportam os resultados deste estudo estão disponíveis para consulta sob demanda em drive institucional de responsabilidade das autoras.

Ética

Esta pesquisa foi registrada e aprovada pelo Comitê de Ética do Instituto de Psicologia da Universidade de São Paulo sob o título “Aquisição Fonológica de Segunda Língua”, número CAAE 46139815.1.0000.5561.

Agradecimentos

Agradecemos ao professor Dr. Plínio Barbosa (UNICAMP) por ter gentilmente cedido o script que ajudou na realização desta pesquisa; aos professores que participaram da banca de defesa e discutiram a versão do trabalho de Fragozo (2017), Dr. Ubiratã Kickhofel Alves, Dr. Paulo Chagas, Dr<sup>a</sup> Eneida de Goes Leal e Dr<sup>a</sup> Ester Miriam Scarpa; aos participantes dos eventos MiniEnapol de Fonética e

Fonologia, EPIP4 4th International Conference on English Pronunciation: Issues & Practices, XX APIRS Convention on Regular Schools X Language Schools: EFL teachers' beliefs and practices, que contribuíram com perguntas e sugestões a versões anteriores deste trabalho, e aos dois pareceristas, Ronaldo Manguiera Lima Jr e Amanda Post, pelo levantamento de questões que enriqueceram ainda mais a discussão. Todos os erros remanescentes são de nossa responsabilidade.

### REFERÊNCIAS

ABAURRE-GNERRE, M. B. Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. *Caderno de Estudos Linguísticos*, n. 2, p. 23-34, 1981.

ABOUSALH, E. *Resolução de choques de acento no português brasileiro: elementos para uma reflexão sobre a interface sintaxe/ fonologia*. Dissertação de mestrado. Campinas: Instituto de Estudos da Linguagem, 1997.

ALBINI, A. B., KLUGE, D. C.; SILVA, A. H. P. (2019). Os efeitos de um treinamento de percepção na aquisição de pré-proparoxítonas inglesas por brasileiros. *Revista Linguagem & Ensino*, 16(1), 183-211. <https://doi.org/10.15210/rle.v16i1.15436>.

BARBOSA, P. A. Explaining Brazilian Portuguese resistance to stress shift with a coupled-oscillator model of speech rhythm production. *Cadernos de Estudos Linguísticos*, v. 43, p. 71-92. Campinas, 2002.

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

BOLINGER, D. *Two kinds of vowels, two kinds of rhythm*. Bloomington: Indiana University Press, 1986.

COOPER, W. E.; EADY, S. J. Metrical phonology in speech production. *Journal of Memory and Language*. 25. 369-384, 1986

COUNCIL OF EUROPE. *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press, 2001.

CRYSTAL, D. Prosodic development. In: Fletcher, P. J.; Garman, M. (eds), *Studies in first language development*, New York: Cambridge University Press, 1986, p. 174-197.

FERREIRA, L. *High initial tones and plateaux in Spanish and Brazilian Portuguese neutral declaratives: Consequences to the relevance of F0, duration and vowel quality as stress correlates*. Tese (Doutorado em Linguística) – University of Illinois at Urbana-Champaign, Urbana, 2008.

FOUGERON, C.; KEATING, P. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustic Society of America*, v. 101, n.6, p. 3728-3740, 1997.

FRAGOZO, C. Aquisição das regras fonológicas do inglês por falantes de português brasileiro. 2017. 251f. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2017.

FRAGOZO, C.; SANTOS, R. S. Aquisição da retração acentual do inglês por falantes de português brasileiro. *ALFA: REVISTA DE LINGUÍSTICA (UNESP. ONLINE)*, v. 65, p. e11797, 2021.

FRY, D.B. Duration and Intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765–8. 1955

FRY, D.B. Experiments in the perception of stress. *Language and Speech*, v. 1, p. 126-152, 1958

GAYER, J. E. L.; COLLISCHONN, G. Análise variacionista da resolução de choque de acento. *Revista Virtual de Estudos da Linguagem – ReVEL*, v. 5, n. 9, agosto de 2007.

GRABE, E.; WARREN, P. Stress Shift: Do Speakers Do It or Do Listeners Hear It?. In: CONNELL, B.; ARVANITI, A. [Ed.], *Papers in Laboratory Phonology IV*, 1995

HAYES, B. The Phonology of Rhythm in English. *Linguistic Inquiry*, v. 15, n. 1, 1984, pp. 33-54.

HEUVEN, V. J. Acoustic Correlates and perceptual cues of word and sentence stress – towards a cross-linguistic perspective. In: GOEDMANS, R.; J. HEINZ; HULST, H. V. *The study of word stress and accent – theories, methods, and data. Part I – phonetic correlates and prominence distinctions*. Cambridge University press. Pp. 15-59. 2018

KIMBALL, A. COLE, J. Avoidance of Stress Clash in Perception of American English *Proceedings of Speech Prosody* 7, Dublin, Ireland, 2014

LEVEY, S. LAWRENCE, R. Stress clash: Frequency and strategies of resolution. *Acoustical Society of America Journal*, Volume 111, Issue 5, pp. 2476-2476 (2002)

LEVIN, J.; FOX, J. A. *Estatística para ciências humanas*. 9a ed. São Paulo: Pearson. Prentice Hall, 2004.

LIBERMAN, Mark. & PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249336, 1977. Disponível em: <http://idiom.ucsd.edu/~bakovic/211B/Lieberman-Prince-1977.pdf>, acessado em 30/05/2013

MADUREIRA, S. An acoustic study of sequences of words with adjacent primary stressed syllables: does stress shift occur in Brazilian Portuguese? *Cadernos de Estudos Linguísticos (UNICAMP)*, Campinas, v. 43, p. 109-126, 2002.

MAJOR, R. Stress and rhythm in Brazilian Portuguese. In: KOIKE, D. A.; MACEDO, D. P. (Orgs.). *Romance Linguistics: The Portuguese Context*. Westport, Connecticut; London: Bergin Garvey, 1992, p. 3-30.

MASSINI, G. *Acento e ritmo*. São Paulo: Contexto, 1992.

NESPOR, Marina & VOGEL, Irene. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986. 327p.

MODESTO, F.; BARBOSA, P. A. Desvendando a prosódia do sotaque estrangeiro: produção e percepção do acento tônico no inglês por falantes brasileiros. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 27, n. 1, p. 165-189, 2019.

REETZ, H; JONGMAN, A. *Phonetics: transcription, production, acoustics and perception*. Oxford: Wiley-Blackwell, 2009. 316p.

ROACH, P. *English Phonetics and Phonology: A practical course*. 3ª ed. Cambridge: Cambridge University Press, 2000. 286p.

SANTOS, R. S.; LEAL, E. G. Os domínios prosódicos e a duração de sílaba no português brasileiro. *Revista da ABRALIN*, Florianópolis, v.5, n.1, p.143-167, 2008.

SANTOS, R. S.; FRAGOZO, C. Período crítico e aquisição fonológica do inglês por falantes brasileiros. *ALFA: REVISTA DE LINGÜÍSTICA (UNESP. ONLINE)*, v. 64, p. 1-25, 2020.

SELKIRK, Elisabeth. The syllable. In: *The structure of phonological representations (part II)*, H. van der Hulst & N. Smith (eds.). Dordrecht: Foris, 1982. 337-83.

SILVA Jr. *Interferências rítmicas do português no inglês como L2: o choque acentual*. Tese (Doutorado em Linguística). Joao Pessoa: Universidade Federal da Paraíba, 2013.

SILVA, JR.; SCARPA, E.M. Stress clash resolution in English as L1 and L2. *Revista Prolingua*. Vol. 14, n.1. mai/ago. p. 79-93. 2019.