

O processamento lexical do acento em pseudopalavras a partir da leitura de falantes do português brasileiro

The lexical processing of stress in pseudowords based on the reading of Brazilian Portuguese speakers

Aline de Lima Benevides*

Faculdade de Tecnologia do Estado de São Paulo, São Paulo, SP, Brasil

Resumo: Este trabalho investigou, a partir de um estudo experimental, o papel da similaridade fonológica, da frequência de ocorrência da palavra base e da sílaba final na atribuição do acento primário em pseudopalavras a partir da produção de falantes nativos do português brasileiro. A métrica de criação das pseudopalavras baseou-se no estudo de Protopapas et al. (2006). As palavras a partir das quais as pseudopalavras foram criadas tinham alta ou baixa frequência de ocorrência, três sílabas de extensão. Contrastaram-se quatro grupos experimentais: i. similar de alta frequência; ii. dissimilar de alta frequência; iii. similar de baixa frequência; e iv. dissimilar de baixa frequência. Hipotetizou-se que pseudopalavras similares de alta frequência desencadeassem mais processos analógicos do que pseudopalavras dissimilares. Participaram do estudo 34 falantes, em uma tarefa de leitura. Os resultados demonstraram que a semelhança entre palavras reais e pseudopalavras desencadeia processos analógicos com a extensão do padrão acentual e que a frequência de ocorrência da palavra alvo e a sílaba final também são relevantes. Tais resultados trazem indícios de que o acento é armazenado juntamente com a palavra (isto é, é lexical) e de que os níveis segmental, silábico e acentual interagem na determinação do local em que o acento incide.

Palavras-chave: Acento. Pseudopalavras. Português Brasileiro.

Abstract: This work investigated, based on an experimental study, the role of phonological similarity, the frequency of occurrence of the base word and the final syllable in the attribution of primary stress in pseudowords based on the production of native speakers of Brazilian Portuguese. The pseudoword creation metric was based on the study by Protopapas et al. (2006). The words from which the pseudowords were created had a high or low frequency of occurrence, three syllables in length. Four experimental groups were compared: i. similar high frequency; ii. high frequency dissimilar; iii. similar low frequency; and iv. low frequency dissimilar. It was hypothesized that similar high-frequency pseudowords would trigger more analogical processes than dissimilar pseudowords. 34 speakers participated in the study, in a reading task. The results demonstrated that the similarity between real words and pseudowords triggers analogical processes with the extension of the stress pattern and that the frequency of occurrence of the target word and the final syllable are also relevant. Such results provide evidence that the stress is stored together with the word (that is, it is lexical) and that the segmental, syllabic and stress levels interact in determining the place where the stress falls.

Keywords: Stress. Pseudowords. Brazilian Portuguese.

* Doutorado em Letras pela Universidade de São Paulo; professora na educação básica e na FATEC; São Paulo, SP; benevides.aline12@gmail.com

1 INTRODUÇÃO

Mais de meio século após a primeira hipótese sobre o acento em português brasileiro (Câmara Jr., 1970), este fenômeno continua alvo de inúmeras análises e reanálises, visto que, no nosso entendimento, não há, ainda, uma proposta que consiga compreender todos os padrões acentuais sem se valer de estratégias de excepcionalidade. Elas, em geral, consideram aspectos prosódicos e/ou morfológicos (cf. Mateus, 1983; Bisol, 1994; Lee, 1995, 2007; Massini-Cagliari, 1999; Magalhães, 2004; Wetzels, 2007), além de justificar padrões ditos irregulares como extramétricos e/ou marcados, por exemplo.

Atualmente, a proposta que melhor justifica os padrões encontrados na língua consiste na de Cantoni (2013), uma das mais recentes análises do acento em Português Brasileiro (doravante PB). A autora assume dois mecanismos centrais: especificação lexical e determinação morfológica. Estes são explicados diacronicamente. O latim, língua-mãe do português, caracterizava-se como uma língua de acento fixo, visto que era determinado pela quantidade silábica: sílabas penúltimas pesadas atraíam o acento, do contrário, o acento recuava para a sílaba precedente (antepenúltima). Por o português ter herdado grande parte do léxico do latim, há, ainda em seu léxico, resquícios dessa forte correlação entre peso silábico e padrão acentual. Entretanto, diante do fato de que a sensibilidade à quantidade silábica foi perdida na passagem do latim ao português e de que inúmeros processos fonológicos atingiram a língua nesse período, a previsibilidade do acento se perdeu e a regularidade dos padrões morfológicos se estabeleceu como o novo atrator da língua, determinando o local em que o acento incide. Com isso, a morfologia passa a regular o padrão acentual de palavras que apresentam morfemas de derivação ou de flexão. Os demais vocábulos são compreendidos como especificados lexicalmente, já que teriam herdado a tonicidade da língua da qual provém.

O ponto central da análise de Cantoni (2013), que nos faz assumi-la como ponto de partida do presente trabalho, são os mecanismos que regulam a construção das representações gramaticais. Para ela, os vocábulos, ao serem usados, têm seus padrões abstraídos, categorizados e armazenados em suas representações mentais, de forma que eles podem ser replicados a novas palavras - conforme previstos pelos Modelos baseados no Uso (Bybee, 2001, 2006, 2010) e pela Teoria dos Exemplares (Kemmer; Barlow, 1999; Langacker, 1999, 2008, 2013; Bybee, 2010; Pierrehumbert, 2001). Esses modelos também preveem que a frequência de tipo e a de ocorrência são mediadoras desse processo, já que a frequência de tipo é responsável pela fortificação e pela replicação de padrões recorrentes (silábicos, segmentais, acentuais etc.), ao passo que a frequência de ocorrência garante a estabilização e a não regularização de vocábulos irregulares.

Todos esses mecanismos se mostram relevantes para a presente pesquisa, tendo em vista que testamos, à luz do experimento realizado por Protopapas et al. (2006) na língua grega, em que medida a semelhança entre palavras reais e pseudopalavras pode desencadear processos analógicos, fazendo com que o padrão acentual da palavra real seja estendido para o novo vocábulo - por exemplo, investiga-se, se uma pseudopalavra como *catater*, ao ser associada com a palavra real *caráter*, em decorrência de suas semelhanças fonológicas, receberá também o padrão acentual paroxítono (*catáter*). Em outros termos, avaliamos a hipótese de que a produção de

pseudopalavras¹ pode resgatar e ativar padrões métricos e segmentais baseados nas palavras do PB armazenadas no léxico mental dos falantes, a partir de processos analógicos entre palavras reais e pseudopalavras. Para realizar tal análise, três variáveis são analisadas: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final. Através da similaridade fonológica, investigamos em que medida pseudopalavras que são semelhantes a palavras reais podem ter o seu padrão acentual estendido para pseudopalavras. Ou seja, a questão que está posta é: ao processarmos um vocábulo novo (pseudopalavra), associamo-lo a uma palavra existente e resgatamos o seu acento? Quanto à frequência dos itens lexicais, analisamos se a frequência de ocorrência do item lexical a partir do qual a pseudopalavra é criada pode mediar o acesso lexical, promovendo mais produções com semelhante padrão acentual em palavras de alta frequência do que as de baixa frequência. A sílaba final, por fim, permite-nos tanto controlar o *template* das pseudopalavras, como também testar se a composição da sílaba final pode ou não ser atratora do acento. Isso porque propostas métricas, como a de Bisol (1994), propõem que sílabas pesadas finais são atradoras do acento em português.

É preciso destacar, a priori, que esta pesquisa não tem o intuito de advogar a favor de que um único fator determina o acento em PB; ao contrário disso, partimos da hipótese de Cantoni (2013) de que o acento é lexical, de que padrões acentuais foram abstraídos e podem ser reproduzidos em novos vocábulos. Buscamos, portanto, mapear quais são os fatores que contribuem para que tal processo ocorra. Esses fatores, em nossa perspectiva teórica, são múltiplos. Este trabalho se dedica a avaliar apenas três: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final, a partir de modelos de processamento de palavra (Coltheart; Rastle, 1994; Rastle; Coltheart, 2000; Coltheart, 2005), dos Modelos de Uso (Bybee, 2001, 2006, 2010) e de Exemplares (Kemmer; Barlow, 1999; Langacker, 1999, 2008, 2013; Bybee, 2010; Pierrehumbert, 2001).

2 METODOLOGIA

2.1 Participantes

Participaram deste estudo 34 falantes nativos do português brasileiro, com idade entre 18 e 60 anos, de ambos os sexos e com nível de escolaridade de Ensino Fundamental completo a Mestrado. Nenhum dos participantes relatou qualquer problema cognitivo, auditivo ou fonoarticulatório².

2.2 Estímulos

Todas as pseudopalavras apresentam três sílabas com os seguintes *templates*: CV-CV-CV ou CV-CV-CVC. Selecionaram-se essas estruturas por serem as mais frequentes do português brasileiro (Benevides; Guide, 2017) e, ainda, por permitirem

¹ Pseudopalavras são palavras inventadas que respeitam a fonotática da língua, mas não possuem significado.

² Esta pesquisa foi registrada e aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH/USP), sob o processo nº 3.332.409. Todos os participantes consentiram que o estudo fosse gravado e utilizado em pesquisas acadêmicas.

que todas as tonicidades (oxítone, paroxítone e proparoxítone) pudessem ser empregadas pelos participantes. Além disso, esses *templates* permitem avaliar em que medida a composição da sílaba final, com ou sem coda, pode determinar o local em que o acento incide. Investiga-se, com esses *templates*, possíveis efeitos da sílaba final.

Tais estruturas foram preenchidas por material segmental a partir da métrica de similaridade fonológica proposta por Protopapas et al. (2006). Segundo os autores, palavras similares são aquelas que se diferenciam minimamente, numa escala de 1 a 3, considerando, para isso, modificação de vozeamento, ponto e modo de articulação - um ponto para cada mudança. A título de exemplo, a pseudopalavra *tarater* apresenta similaridade fonológica de um nível com a palavra *caráter*, pois a mudança de /k/ para /t/ acarreta a modificação apenas de ponto de articulação. Palavras dissimilares, por sua vez, são aquelas que se diferenciam de 5 a 10 níveis - novamente, considerando vozeamento, ponto e modo de articulação. A palavra *caráter* é dissimilar a *gatafer*, uma vez que há mudança no vozeamento de /k/ para /g/; troca de /r/ para /t/ com modificação de vozeamento e de ponto de articulação; e, por fim, alteração de /t/ para /f/, totalizando cinco mudanças. Há, portanto, dois níveis da variável similaridade: similar e não similar.

A frequência de ocorrência do item lexical a partir do qual a pseudopalavra foi criada consistiu no terceiro fator de investigação. Foram caracterizadas como palavras de alta frequência aquelas que apresentavam frequência igual ou superior a 100 ocorrências por milhão e palavras de baixa frequência aquelas cuja frequência era igual ou inferior a 2 ocorrências por milhão. A frequência de ocorrência foi retirada do Corpus ABG (Benevides; Guide, 2017). Para esta variável, dois níveis foram investigados: alta e baixa frequência.

As variáveis similaridade fonológica e frequência de ocorrência foram cruzadas, formando quatro grupos de investigação: i. pseudopalavras maximamente similares de alta frequência; ii. pseudopalavras minimamente similares de alta frequência; iii. pseudopalavras maximamente similares de baixa frequência; e iv. pseudopalavras minimamente similares de baixa frequência, conforme exposto no Quadro 1³.

Quadro 1 - Descrição dos grupos de pseudopalavras por similaridade e frequência com exemplos.

	Similaridade	Frequência	Exemplos
Grupo 1	Alta	Alta	Figura (<i>diçura, fidura, figuba</i>) Caráter (<i>darater, cacater, caraper</i>)
Grupo 2	Baixa	Alta	Código (<i>nojilo, zofisu, bocito</i>) Militar (<i>pitudar, dipinar, sipivar</i>)
Grupo 3	Alta	Baixa	Pateta (<i>lateta, pafeta, pateda</i>) Cabaré (<i>nabare, camare, cabate</i>)
Grupo 4	Baixa	Baixa	Pêssego (<i>lemifo, demefo, zedeno</i>) Funeral (<i>mupeval, lusipal, puvebal</i>)

Fonte: adaptado de Benevides (2022, p. 128).

³ Confira todos as palavras reais e as pseudopalavras no Apêndice.

A fim de evitar possíveis enviesamentos, restringimos o uso de consoantes palatais (/ɲ/ e /ʎ/), da fricativa /x/ e das pós-alveolares (/ʃ/ e /ʒ/) em posição de ataque da última e/ou da penúltima sílabas, visto que, segundo Câmara Jr. (1953), Cristófar-Silva (1995) e Wetzels (2000), esses segmentos impedem a emergência do acento proparoxítono - o que é contrário ao objetivo deste estudo, que é permitir que todas as tonicidades possam emergir.

2.3 Validação da métrica de (dis)similaridade

Para garantir que os estímulos criados fossem, de fato, considerados (dis)similares a uma palavra real, todas as pseudopalavras criadas foram submetidas a um teste de validação. Este consistiu na exposição de todos os vocábulos, por meio do envio de um formulário *on-line*, a um grupo de 10 falantes nativos do PB, que foram instruídos a indicar para cada pseudopalavra uma palavra real que se assemelhasse - isto é, a primeira que viesse em mente. O índice determinado para que o estímulo fosse validado como similar foi de 80% - ou seja, dos 10 participantes, 8 deveriam indicar a mesma palavra que utilizamos como base para a criação da pseudopalavra; para os dissimilares, dos 10 participantes, no máximo, 2 indivíduos podiam indicar uma mesma palavra, independente se era a palavra base ou qualquer outro vocábulo.

A taxa média de validação foi de 72,2% das palavras. Somente as pseudopalavras validadas por esse teste foram incluídas na presente análise, de forma que não foi possível um balanceamento total dos estímulos, já que alguns grupos não tiveram todas as pseudopalavras validadas⁴.

2.4 Quantidade de dados

O experimento contou, inicialmente, com 309 pseudopalavras⁵, que foram lidas por 34 falantes nativos do português brasileiro, totalizando 10.506 estímulos. Destes, foram excluídos 113 (1,1%) em decorrência inadequações nas produções⁶, resultando em 10.393 estímulos lidos.

2.5 Desenho experimental

Os estímulos foram exibidos aos participantes por meio de uma apresentação de *slides*, com uma palavra por tela. Todas as pseudopalavras foram escritas seguindo a ortografia da língua portuguesa, sem qualquer marcação de acento gráfico. É preciso destacar que é sabido que a forma escrita com ou sem gráfico dá pistas aos falantes em relação à sílaba proeminente, como em *sábia*, *sabia* e *sabiá*. Contudo, adotou-se tal metodologia, tendo em vista que diversos estudos prévios demonstram uma maior

⁴ É preciso destacar que a decisão de excluir os estímulos não validados não se deu de forma aleatória, tendo em vista que as estatísticas que realizaram a análise inferencial deste trabalho optaram por excluir os dados não validados. Segundo elas, para que a hipótese de investigação fosse analisada, era preferível desbalancear os grupos experimentais do que incluir dados não validados. Os grupos mais afetados foram, dessa forma, os grupos similares, G1 e G3, resultando em 67 pseudopalavras no G1; 86 no G2; 65 no G3 e 91 no G4.

⁵ Para mais detalhes da contabilização dos estímulos, consulte Benevides (2022).

⁶ Os estímulos excluídos decorrem de troca de segmentos, (des)vozeamento, inserção de sílabas ou segmentos, mudança na ordem das sílabas etc.

neutralidade na leitura de palavras escritas do que na produção de palavras ouvidas ou formadas a partir de imagens (Benevides, 2017).

Para minimizar possíveis interferências do conhecimento que os falantes têm em relação ao sistema ortográfico da língua portuguesa, os participantes foram orientados de que as palavras expostas na tela do computador não apresentavam qualquer marcação de diacríticos (acentuação, cedilha etc.) em decorrência de falhas do programa no reconhecimento de caracteres. A tarefa experimental foi realizada *on-line*, com o compartilhamento da tela do computador da pesquisadora com os participantes. A chamada de vídeo e áudio foi realizada por meio de aplicativos como o *Zoom* e o *Google Meet*, com captação de áudio/vídeo realizada através do *notebook NP Compact*. Todos os participantes deste estudo autorizaram previamente a gravação de áudio/vídeo e o seu uso em meios acadêmicos.

2.6 Método de análise dos dados

Os dados foram submetidos a análises descritivas e inferenciais no *software R* (R Core Team, 2015), a partir da avaliação de um estatístico. O modelo utilizado na análise foi *modelo de regressão logística multinomial misto*, assumindo a paroxítona como categoria de referência.

3 RESULTADOS E ANÁLISE

Os resultados gerais da tarefa experimental compreenderam 4.569 (44%) de produções oxítonas, 5.436 (52,3%) de paroxítonas e 388 (3,7%) de proparoxítonas, conforme exposto na Tabela 1.

Tabela 1 - Frequência e proporção de tonicidade de produção nos dados gerais.

Tonicidade de Produção	n (%)
Oxítone	4.569 (44%)
Paroxítone	5.436 (52,3%)
Proparoxítone	388 (3,7%)
Total	10.393 (100%)

Fonte: Benevides (2022, p. 143).

Esses resultados gerais já trazem indícios de dois fatos relevantes: i. embora a paroxítona seja o acento predominante, as oxítonas também emergiram de forma significativa; e ii. as proparoxítonas não podem ser consideradas um padrão excepcional da língua, tendo em vista que, até mesmo em um estudo experimental, elas foram produzidas.

Para além dos resultados mais abrangentes, interessa-nos, nesta pesquisa, avaliar se a similaridade fonológica desempenha algum papel na atribuição do acento em português. Os dados quantitativos, apresentados na Tabela 2, ajudam-nos a observar que os grupos similares (G1 e G3) apresentaram taxas mais altas de acento oxítono e proparoxítono quando a palavra base também apresentava tal tonicidade, respectivamente. Em outros termos, 85% das pseudopalavras criadas a partir de uma

palavra real oxítone receberam também acento oxítone. Essa taxa mostra-se significativamente maior do que as pseudopalavras do grupo dissimilar, que só atingiram 51,1% ($p < 0,001$). O efeito da similaridade parece ser ainda maior em relação às proparoxítonas, que atingiram uma taxa de 22,2% nos grupos similares em comparação a 4,4% nos grupos dissimilares ($p < 0,001$). A análise inferencial seleciona a variável *similaridade fonológica* como uma variável relevante para a explicação do fenômeno em análise ($p < 0,001$).

Tabela 2 - Proporção de tonicidade de produção por tonicidade da palavra base e por nível de (dis)similaridade.

Nível de Similaridade	Tonicidade de Produção = Tonicidade da Palavra Base		
	Oxítone	Paroxítone	Proparoxítone
Similaridade (G1 e G3)	85%	58,1%	22,2%
Dissimilar (G2 e G4)	51,1%	52,5%	4,4%

Fonte: Adaptação de Benevides (2022, 152).

É preciso destacar, ainda, que a análise inferencial indica que os grupos dissimilares (G2 e G4) favorecem o acento paroxítone. Este resultado mostra-se extremamente relevante, pois evidencia que a ausência de semelhança entre uma palavra real e uma pseudopalavra faz com que o falante atribua o padrão acentual *default* da língua - no caso, o paroxítone. Ou seja, uma possível interpretação é de que a ausência de um distrator, que direcione o acento para uma sílaba específica da palavra, como a similaridade fonológica, faz com que o indivíduo resgate o padrão acentual mais recorrente/robusto de sua língua, armazenado em seu léxico mental, e o atribua ao novo vocábulo. Há, assim, uma evidência importante de que o acento paroxítone é o padrão acentual do PB, já que é o mais frequente, o que leva à sua produtividade. Outra possível interpretação é de que o percentual em torno de 50% de acento oxítone e paroxítone nos grupos dissimilares pode dar indícios de uma atribuição aleatória, entretanto, como será evidenciado nos dados da Tabela 8, a sílaba final pode ser um fator explicativo.

Uma maneira de evidenciarmos como a similaridade fonológica desencadeia processos analógicos por meio da extensão do padrão acentual entre o vocábulo real e a pseudopalavra é detalharmos mais esses dados, a fim de analisarmos a relação entre atribuição acentual e tonicidade da palavra base, conforme demonstra a Tabela 3:

Tabela 3 - Frequência e proporção de tonicidade de produção por tonicidade da palavra alvo.

		Tonicidade de Produção			
		Oxítona	Paroxítona	Proparoxítona	Total
Tonicidade da Palavra Alvo	Oxítona	2.397 (63,6%)	1.331 (35,3%)	40 (1,1%)	3.768 (36,3%)
	Paroxítona	1.863 (44%)	2.371 (56%)	65 (1,5%)	4.299 (41,4%)
	Proparoxítona	309 (13,3%)	1.734 (74,5%)	283 (12,2%)	2.326 (22,4%)
	Total	4.569 (44%)	5.436 (52,3%)	388 (3,7%)	10.393 (100%)

Fonte: Benevides (2022, p. 150).

Observe, como dito, que houve um maior percentual de atribuições de acento oxítono (63,6%) quando a pseudopalavra foi criada a partir de uma palavra com tonicidade oxítona. A probabilidade dela ser oxítona, entretanto, sofre uma queda de 75% e 90% quando a palavra base é paroxítona ou proparoxítona, respectivamente. Sob a mesma ótica, a probabilidade da palavra ser paroxítona aumenta se a palavra base for paroxítona ou proparoxítona ($p < 0,001$). Há, ainda, 400% mais chances da pseudopalavra ser proparoxítona se ela tiver sido criada a partir de uma palavra proparoxítona.

Antes de adentrarmos na análise a respeito da frequência lexical da palavra base, cabe uma reflexão sobre a composição dos grupos experimentais: a fim de possibilitar que houvesse todos os contrastes de similaridade e frequência, criamos quatro grupos (G1 - Similar de alta frequência; G2 - Dissimilar de alta frequência; G3 - Similar de baixa frequência; e G4 - Dissimilar de baixa frequência). Contudo, ao observamos os resultados, expostos na Tabela 4, questionamo-nos se, de fato, tal divisão se mostra válida, tendo em vista que o pressuposto para a pseudopalavra ser enquadrada na categoria dissimilar é não apresentar semelhança com uma palavra real, de forma que não parece um raciocínio válido pensar que a frequência da palavra real que não se assemelha à pseudopalavra pode a afetar. Se tal raciocínio estiver correto, como julgamos que o seja, a melhor maneira de analisarmos o papel da frequência lexical da palavra base é isolar apenas o G1 dos demais grupos, como apresentado na Tabela 5:

Tabela 4 - Frequência e proporção de tonicidade de produção por tonicidade da palavra alvo e por nível de frequência.

Nível de Similaridade	Tonicidade da Palavra Alvo	Tonicidade de Produção			
		Oxítona	Paroxítona	Proparoxítona	Total
Grupos de Alta Frequência (1 e 2)	Oxítona	1.106 (62,1%)	667 (37,4%)	9 (0,5%)	1.782 (17,1%)
	Paroxítona	950 (41,6%)	1.298 (56,8%)	38 (1,7%)	2.286 (22%)
	Proparoxítona	53 (4,9%)	889 (82,5%)	136 (12,6%)	1.078 (10,4%)
Grupos de Baixa Frequência (3 e 4)	Oxítona	1.291 (65%)	664 (33,4%)	31 (1,6%)	1.986 (19,1%)
	Paroxítona	913 (45,4%)	1.073 (53,3%)	27 (1,3%)	2.013 (19,4%)
	Proparoxítona	256 (20,5%)	845 (67,7%)	147 (11,8%)	1.248 (12%)
	Total	4.569 (44%)	5.436 (52,3%)	388 (3,7%)	10.393 (100%)

Fonte: Benevides (2022, p. 153).

Tabela 5 - Proporção de tonicidade de produção por tonicidade da palavra base e por nível de frequência.

Nível de Frequência	Tonicidade de Produção = Tonicidade da Palavra Base		
	Oxítona	Paroxítona	Proparoxítona
G1 - Similar de alta frequência	91,6%	60,5%	21,5%
Demais grupos (G2, G3 e G4)	60,7%	53,4%	10,5%

Fonte: Benevides (2022, p. 206).

Observe, a partir da Tabela 5, que há significativamente mais atribuições oxítonas ($p < 0,0001$) quando a palavra base é oxítona no grupo similar do que nos demais grupos - o mesmo vale para as demais tonicidades. Tal fato traz evidências de que a frequência da palavra base é um fator relevante na atribuição de tonicidade nesses vocábulos, podendo atuar, como pressuposto pelos Modelos baseados no Uso, como um mediador no acesso lexical. Isto é, a alta frequência das palavras faz com que elas tenham conexões mais robustas e, conseqüentemente, estejam mais disponíveis para serem acessadas pelos falantes, permitindo que os seus padrões sejam mais facilmente reproduzidos em novos vocábulos.

Vamos, agora, observar os resultados quanto à sílaba final. A Tabela 6 apresenta a frequência e a proporção de tonicidade de produção de acordo com a sílaba final da pseudopalavra, se CV ou CVC. Esta análise se faz necessária a fim de

avaliarmos se a forte correlação entre sílaba final e padrão acentual, herdada do latim, se mostra ainda um atrator importante da língua no que se refere à atribuição acentual.

Tabela 6 - Frequência e proporção de tonicidade de produção por sílaba final da pseudopalavra.

		Tonicidade de Produção			Total
		Oxítona	Paroxítona	Proparoxítona	
Sílaba Final da Palavra	CV-CV-CV	570 (10,6%)	4.518 (84,4%)	267 (5%)	5.355 (51,5%)
	CV-CV-CVC	3.999 (79,4%)	918 (18,2%)	121 (2,4%)	5.038 (48,5%)
	Total	4.569 (44%)	5.436 (52,3%)	388 (3,7%)	10.393 (100%)

Fonte: Benevides (2022, p. 151).

Note que há um maior número de atribuições de acento oxítono quando a pseudopalavra termina com CVC: 79,4% em contraste com 10,6% ($p < 0,0001$). De modo semelhante, há mais atribuições paroxítonas quando a pseudopalavra termina com CV (84,4%). Para as proparoxítonas, o percentual é maior com CV (5%) em detrimento de CVC (2,4%). Este resultado é fundamental de ser comentado, visto que a literatura linguística, de modo geral, prediz que o acento proparoxítono não emerge quando a palavra termina em sílaba ramificada. Neste caso, boa parte deste resultado (2,4%) refere-se à emergência do acento proparoxítono em pseudopalavras que tinham como palavra base a palavra real *Júpiter*, explicitando, dessa forma, que a similaridade fonológica mostrou-se um fator mais relevante do que a baixa frequência de exemplares na língua com esse padrão (CVC final com tonicidade proparoxítona).

Observamos, até aqui, que as três variáveis isoladas são relevantes para a atribuição do acento em português brasileiro. Vamos, então, avaliar a interação entre elas. A Tabela 7 apresenta os dados referente à frequência e à proporção de tonicidade de produção por grupo experimental, o qual correlaciona a variável similaridade com frequência lexical.

Tabela 7 - Frequência e proporção de tonicidade de produção por grupo experimental.

Grupo Experimental	Tonicidade de Produção/Tonicidade da Palavra Alvo		
	Oxítona/ Oxítona	Paroxítona/ Paroxítona	Proparoxítona/ Proparoxítona
Grupo 1 Similar de alta frequência	91,6%	60,5%	21,5%
Grupo 2 Dissimilar de alta frequência	46,7%	53,2%	3,7%

Grupo 3 Similar de baixa frequência	79,9%	55,2%	23%
Grupo 4 Dissimilar de baixa frequência	55,4%	51,7%	4,9%

Fonte: Benevides (2022, p. 148).

Note que a correlação entre os fatores só reforça os resultados individuais de cada uma das variáveis: nos grupos similares (G1 e G3), há um maior percentual de atribuições oxítonas e proparoxítonas quando a palavra base é, respectivamente, oxítônica e proparoxítona; o percentual dos grupos dissimilares (G2 e G4) é muito semelhante, em torno de 50%, reforçando a nossa (re)análise de que esses grupos, na verdade, constituem o mesmo grupo, já que o pressuposto da formação desses grupos é de que os falantes não conseguem recuperar informações da palavra base. Ao unirmos esses grupos, temos 51,1% de produções oxítonas e 52,5% de proparoxítonas, percentuais que podem indicar a aleatoriedade. Esta afirmação só se mostra equivocada quando observamos os resultados dos grupos experimentais também por sílaba final, conforme exposto na Tabela 8.

Ao expandirmos o olhar, verificamos que, na verdade, essa atribuição, aparentemente aleatória, é motivada pela composição da sílaba final: quando ramificada, há um favorecimento do acento oxítono em todos os grupos e, quando não ramificada, há um favorecimento do acento proparoxítono também em todos os grupos. O percentual em torno de 50% pode decorrer, na verdade, de termos vocábulos com as duas estruturas de sílaba final, CV e CVC. O fator central desencadeador do acento nos grupos dissimilares, portanto, parece ser a sílaba final. Note, entretanto, que o percentual de atribuições de acento oxítono com CVC e proparoxítono com CV cai nos grupos similares: em média, 74,3% de oxítonas com CVC e 76,7% de proparoxítonas com CV, ao passo que os grupos dissimilares apresentam, em média, 84,5% de oxítonas com CVC e 88,6% de proparoxítonas com CV. A explicação para tal queda só pode ser o fator que distingue esses grupos, a similaridade fonológica. Há, para esses casos, mais de um fator que interage: similaridade fonológica, frequência lexical e sílaba final.

Tabela 8 - Frequência e proporção de tonicidade de produção por grupo experimental e por sílaba final.

Grupo Experimental	Sílabas Final da Palavra	Tonicidade de Produção		
		Oxítônica	Paroxítona	Proparoxítona
G1 - Similar Frequente	CV-CV-CV	45 (4,3%)	873 (83,4%)	129 (12,3%)
	CV-CV-CVC	950 (78,1%)	259 (21,3%)	7 (0,6%)
G2 - Dissimilar Frequente	CV-CV-CV	120 (7%)	1.549 (90,6%)	40 (2,3%)

	CV-CV-CVC	994 (84,7%)	173 (14,7%)	7 (0,6%)
G3 - Similar de Baixa Frequência	CV-CV-CV	228 (24,9%)	641 (70%)	47 (5,1%)
	CV-CV-CVC	897 (70,4%)	295 (23,1%)	83 (6,5%)
G4 - Dissimilar de Baixa Frequência	CV-CV-CV	177 (10,5%)	1.455 (86,5%)	51 (3%)
	CV-CV-CVC	1.158 (84,3%)	191 (13,9%)	24 (1,7%)

Fonte: Benevides (2022, p. 155).

É possível constatar, dessa maneira, que todos esses fatores contribuem, em alguma medida, para a atribuição do acento em pseudopalavras. Ao nosso ver, esses, possivelmente, não são os únicos. É preciso mais estudos para mapear e avaliar o papel que outras variáveis podem desempenhar em um fenômeno linguístico tão complexo como o acento lexical.

4 CONSIDERAÇÕES FINAIS

Diante das inúmeras pesquisas que buscam determinar um único fator como motivador para a atribuição do acento, este estudo, de caráter experimental, tem como premissa a atuação de múltiplos fatores, à luz dos Modelos baseados no Uso. Os fatores investigados foram: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final. Os resultados experimentais evidenciam a importância de todos eles na determinação do local em que o acento incide.

A relevância de mais de um fator só reforça a hipótese de análise desta pesquisa de que os padrões fonológicos, incluindo, neste caso, os padrões acentuais, são abstraídos e categorizados na gramática fonológica dos falantes. Por estarem disponíveis nas representações mentais (ou seja, no seu léxico mental), podem ser, e são, replicados em vocábulos novos da língua - neste caso, em pseudopalavras, a partir de processos analógicos. Nesta pesquisa, o mecanismo desencadeador desses processos foi a similaridade fonológica entre os vocábulos, sendo que a frequência de ocorrência do item lexical mostra-se um potencializador da analogia. Os dois fatores interagem com a sílaba final, que é um *template* extremamente relevante e ativo na língua, sendo o principal determinante, por exemplo, do local em que o acento incide quando não há similaridade fonológica ou outros fatores interagindo.

É possível constatar, dessa forma, que os falantes acessam informações segmentais e as consideram no momento de atribuir tonicidade a um vocábulo, demonstrando que os níveis segmentais e métricos estão relacionados nas representações mentais (não devendo ser concebidas de forma independente). Em outros termos, os falantes fazem processos analógicos segmentais e acentuais, o que indica que o acento está associado à palavra, isto é, é lexical.

Em síntese, os resultados desta pesquisa estão em consonância com a hipótese de Cantoni (2013) de que o acento primário em português brasileiro traz resquícios de

uma herança latina que associou sílaba final com padrão acentual, por isso, a sua maior influência em pseudopalavras dissimilares; e, principalmente, a defesa de que o acento é lexical, visto que ele está associado a um vocábulo que tem seu padrão abstraído e replicado em palavras fonologicamente semelhantes.

REFERÊNCIAS

- Benevides AL. O acento primário em pseudopalavras: uma abordagem experimental [dissertação]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2017.
- Benevides AL. O acento em pseudopalavras [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2022.
- Benevides AL, Guide BF. Corpus ABG. *Texto Livre: Linguagem e Tecnologia*. 2017;10 (1):139-163.
- Bisol L. O acento e o pé métrico. *Letras de Hoje*. 1994;29(4):25-36.
- Bybee J. *Phonology and language use*. Cambridge: Cambridge University Press; 2001.
- Bybee J. From usage to grammar: the mind's response to repetition. *Scholarly journals online*; 2006. p. 711-732.
- Bybee J. *Language, usage and cognition*. Cambridge: Cambridge University Press; 2010.
- Câmara Jr. JM. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Simões; 1953.
- Câmara Jr. JM. *Estrutura da língua portuguesa*. 34.^a ed. Petrópolis: Editora Vozes; 2001 [1970].
- Cantoni MM. O acento no português brasileiro: uma abordagem experimental [tese]. Belo Horizonte: Estudos Linguísticos da Faculdade de Letras, Universidade Federal de Minas Gerais; 2013.
- Coltheart M. Modeling reading: the dual-route approach. In: Snowling MJ, Hulme C, editores. *The science of reading: a handbook*. Oxford: Blackwell; 2005. p. 6-23.
- Coltheart M, Rastle K. Serial processing in reading aloud: evidence for dual-route models of reading. *Journal of Experimental Psychology*. 1994;20(6):1197-1211.
- Cristóvão-Silva T. Intervocalic palatal glides in Brazilian Portuguese. *Revista de Estudos Linguísticos*. 1995;2(4):5-20.
- Kemmer S, Barlow M. Introduction: a usage-based conception of language. In: Barlow M, Kemmer S. *Usage based models of language*. Califórnia: CSLI Publications; 1999. p. vii-xxviii.
- Langacker R. A dynamic usage-based model. In: Barlow M., Kemmer S. *Usage-based models of language*. Califórnia: CSLI Publications; 1999. p. 1-63.
- Langacker R. *Cognitive grammar: a basic introduction*. Nova Iorque: Oxford University Press; 2008.
- Langacker R. *Essentials of cognitive grammar*. Nova Iorque: Oxford University Press; 2013.
- Lee SH. *Morfologia e fonologia lexical do português do Brasil* [tese]. Campinas: Instituto de Estudos da Linguagem, Universidade Estadual de Campinas; 1995.
- Lee SH. O acento primário no português: uma análise unificada na teoria da otimalidade. In: Araújo GA, organizador. *O acento em português: abordagens fonológicas*. São Paulo: Parábola Editorial; 2007. p. 121-144.
- Magalhães JS. *O plano multidimensional do acento na teoria da otimalidade* [tese]. Porto Alegre: Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul; 2004.
- Massini-Cagliari G. *Do poético ao linguístico no ritmo dos trovadores: três momentos da história do acento*. São Paulo: Cultura Acadêmica; 1999.
- Mateus MHM. O acento em palavra em português: uma nova proposta. *Boletim de Filologia*. 1983;27:211-229.

Pierrehumbert JB. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee J, Hopper P, editores. Frequency effects and the emergence of linguistic structure. Amsterdam: John Benjamins; 2001. p. 137-157.

Protopapas A, et al. Lexical and default stress assignment in reading Greek. Journal of Research in Reading. 2006;29(4):418-432.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [programa de computador]. Vienna, Austria. 2015. Disponível em: <http://www.R-project.org>.

Rastle K, Coltheart M. Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. Journal of Memory and Language. 2000;42:342-364.

Wetzels WL. Consoantes palatais como geminadas fonológicas no Português Brasileiro. Revista de Estudos Linguísticos. 2000;9(2):5-15.

Wetzels WL. Primary word stress in Brazilian Portuguese and the weight parameter. Journal of Portuguese Linguistics. 2007;5(2):9-58.

APÊNDICE - Pseudopalavras por grupo experimental com suas respectivas frequências

Grupo 1 - Pseudopalavras similares de alta frequência

CV-CV-CV			
Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Paraná	77.547	tarana, patana, parafa
			perana, parena, parane
	Canadá	25.505	fanada, cadada, canaga
			conada, canida, canadu
	Comitê	33.263	jomite, copite, comife
			camite, comote, comitu
Paroxítona	Cidade	525.295	zidade, cicade, cidafe
			cedade, cidide, cidadu
	Tabela	144.478	babela, tafela, tabeda
			tobela, tabala, tibela
	Figura	372.894	digura, fidura, figuba
			fagura, figora, figuri
Proparoxítona	Número	842.189	lumero, nupero, numeco
			nimero, numaro, namero
	Século	219.623	neculo, sevulo, secuno
			siculo, saculo, secule
	Método	317.435	setodo, mepodo, metopo

			mutodo, metedo, metode
--	--	--	------------------------

CV-CV-CVC

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Capital	197.974	lapital, cabital, capifal
			copital, caputal, capitol
	Diretor	140.646	siretor, difetor, diredor
			daretor, dirotor, diretur
	Natural	114.030	datural, nabural, natupal
			nitural, nataral, naturol
Paroxítona	Possível	300.554	bocivel, pobivel, pocibel
			picivel, pocevel, pocivol
	Caráter	115.284	darater, cacater, caraper
			curater, caruter, carator
	Difícil	104.458	zifícil, dipícil, difíbil
			dafícil, difucil, dificol

Grupo 2 - Pseudopalavras dissimilares de alta frequência

CV-CV-CV

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Paraná	77.547	lapaca, dabaza, zavaba
			poruna, piruno, poranu
	Comitê	33.263	zofire, zoteme, dovife
			camota, camate, comato
	Canadá	25.505	fapama, lacafa, bavaga
			conedo, conado, cundo
Paroxítona	Medida	380.157	tefifa, lenipa, tepiva
			midode, madda, medoda
	Passado	211.473	narafo, nagafu, dalapo
			piceda, pucido, pacido

	Futuro	137.205	pucuto, nubuvo, dumufo futira, fetara, fitare
Proparoxítona	Título	183.459	sibumo, mivipo, lirufu titale, tetola, tatolo
	Código	210.472	nofilo, zofisu, bocito cadegu, coduga, cudoga
	Década	180.750	tefaza, sedafa, revafa dacodo, decode, ducoda

CV-CV-CVC

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Federal	471.606	pevebal, befipal, mefebal faduril, fodorol, fidarel
	Militar	115.102	pitidar, dipinar, sipivar mulotar, malutor, molitur
	Popular	92.011	mogutar, lonutar, zoricar pipalor, papilor, pepalir
Paroxítona	Possível	300.554	gonitel, lomipel, rovinel pacevol, pecival, puceval
	Caráter	115.284	fanaler, zamafer, napaver corutar, carotir, coruter
	Difícil	104.458	mizidil, fidimil, fepibil dufical, dafocol, dafical

Grupo 3 - Pseudopalavras similares de baixa frequência

CV-CV-CV

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Paletó	1.591	taletto, pabeto, palevo pileto, paluto, poletto
	Javali	690	bavali, japali, javari

			juvali, javoli, javala
	Cabaré	196	nabare, camare, cabate cobare, cabure, cabara
Paroxítona	Safári	932	nafari, savari, safapi sefari, safuri, safara
	Pateta	417	lateta, pafeta, pateda puteta, patuta, patetu
	Sacola	311	vacola, sanola, sacoda sucola, sacala, sacolo
Proparoxítona	Pétala	1.928	tetala, pefala, petafa patala, petula, petale
	Cócega	261	docega, cobega, coceca cacega, cucega, cocegu
	Tísica	245	disica, tirica, tisiba tasica, tisaca, tisoca

CV-CV-CVC

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Secador	1.543	becador, segador, secapor secader, secodor, secadar
	Chafariz	606	nafariz, xabariz, xafadiz xofariz, xafuriz, xafaraz
	Recital	143	fecital, relital, recival rucital, racital, recitol
Paroxítona	Cadáver	191	ladaver, cazaver, cadaber codaver, cadiver, cadavar
	Mutável	933	zutavel, muvavel, mutazel mitavel, mutevel, mutaval
	Lavável	204	davavel, labavel, lavatel

			lavavil, lavevel, lavaval
Proparoxítona	Júpiter	1.904	tupiter, jutiter, jupiner
			japiter, jupiter, jupitor

Grupo 4 - Pseudopalavras dissimilares de baixa frequência

CV-CV-CV

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Cabaré	1.020	lafape, barane, vatabe
			cabori, cobiro, cubere
	Sucuri	348	gumuli, govubi, vunuti
			socarú, sicori, sicora
	Camelô	184	bareno, navedo, lacepo
			cumile, comalo, camola
Paroxítona	Mazela	188	davena, fabeta, capeca
			mazole, muzola, mezilo
	Farofa	842	lamofa, dacuga, tagota
			firafe, farefo, firufo
	Jaleco	260	narebo, pamezo, tafeno
			juluca, jalica, jeluci
Proparoxítona	Sádica	250	gafila, tapina, lafipa
			soduca, sedaco, sedoca
	Pêssego	197	lemifo, demefo, zedeno
			pucega, pacigo, piciga
	Súbito	189	bufino, nufilo, zutino
			subata, sebeta, sobite

CV-CV-CVC

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Funeral	1.975	mupeval, lufipal, puvebal
			fanirel, fenarel, finoral

	Comedor	232	podecor, votepor, motefor
			cumodor, cameder, motefor
	Liminar	193	sidigar, pifitar, ritecar
			lomenar, lemunor, lomenir
Paroxítona	Cadáver	191	lamater, nafapir, dapamer
			codivar, cudivor, cadovur
	Mutável	933	sulabel, ruvateľ, vupadel
mitoval, mitavol, motevel			
	Lavável	204	salanel, fadapel, fapatil
			luvivol, livoval, levavol
Proparoxítona	Júpiter	1.904	luniver, tuzemer, gudiler
			japutar, jepatar, jopater