

O processamento morfológico de sufixos derivacionais e modificadores: uma análise de -eiro e -inho

Camila Witt Ulrich¹

Luiz Carlos da Silva Schwandt²

RESUMO

Este artigo trata do processamento morfológico de sufixos rotulados como derivacionais — aqueles capazes de alterar características morfosintáticas da base e, portanto, núcleos da estrutura complexa (ex. -eiro) — e de sufixos modificadores, representados pelos avaliativos — que mantêm características morfosintáticas da base (ex. -inho), conforme Villalva (1994, 2000). Partindo do pressuposto de que há decomposição de unidades morfológicamente complexas em unidades menores no processamento de dados linguísticos no português (GARCIA, 2009; PINTO, 2017), nosso objetivo é verificar se as duas classificações apresentadas para os sufixos demandam diferentes custos de processamento. Para isso, desenvolvemos três experimentos: uma Tarefa de Associação de Palavras (TAP), uma Tarefa de Decisão Lexical (TDL) e uma Tarefa de Decisão Lexical com *Priming* (TDLP), as quais são compostas por estímulos formados pelos sufixos -eiro e -inho. De modo geral, os participantes da TAP demonstraram ter conhecimento dos itens linguísticos testados e atribuíram a eles, predominantemente, associações de cunho morfológico ou semântico. Em ambas as tarefas envolvendo decisão lexical (TDL e TDLP), estímulos formados por -inho apresentaram menores índices de acurácia e tempos de reação relativamente maiores, mas sem significância estatística em comparação aos formados por -eiro. Esse resultado parece sugerir, considerando-se o recorte experimental assumido, que a categoria funcional do afixo não exerce papel no processamento das palavras em análise.

Palavras-chave: *Processamento morfológico; Sufixos derivacionais; Sufixos modificadores.*

ABSTRACT

This paper deals with the morphological processing of Brazilian Portuguese affixes,

1 Professora substituta no Instituto Federal Sul-rio-grandense — Sapucaia do Sul.

2 Professor Titular do Departamento de Linguística, Filologia e Teoria Literária da Universidade Federal do Rio Grande do Sul, bolsista de produtividade em pesquisa CNPq, processo 310921/2018-0.

in particular derivational suffixes — those capable of altering morphosyntactic characteristics of the base and, therefore, heads of the complex structure (ex. -eiro) — and modifier suffixes, represented by evaluative suffixes — those that maintain morphosyntactic characteristics of the base (ex. -inho), according to Villalva (1994, 2000). Assuming that the processing of morphologically complex Portuguese units may involve their decomposition into smaller units (GARCIA, 2009; PINTO, 2017), our objective is to verify whether the two categories presented for the suffixes have different processing costs. For this purpose, we developed three experiments: a Word Association Task (WAT), a Lexical Decision Task (LDT) and a Lexical Decision Task with Priming effect (LDTP), which are composed of stimuli with the suffixes -eiro and -inho. In general, the participants of WAT demonstrated to have knowledge about the linguistic items and attributed to them, predominantly, associations of a morphological or semantic nature. In both tasks involving lexical decision (LDT and LDTP), stimuli formed by -inho showed lower accuracy rates and relatively longer reaction times, but without statistical significance in comparison to those formed by -eiro. This result suggests, considering the assumed experimental outline, that the affix functional category does not play a role in the processing of the words under analysis.

Key words: *Morphological processing; Derivational suffixes; Modifier suffixes.*

Diferentes sufixos apresentam diferentes estatutos — seja no âmbito morfológico, pela função que assumem na estrutura complexa, seja no âmbito prosódico, por apresentarem ou não características de palavra do ponto de vista acentual (SCHWINDT, 2013a). Neste artigo, focamos nas diferenças morfológicas entre os sufixos *-eiro* e *-inho*, representantes de dois grupos distintos: os sufixos derivacionais e os sufixos modificadores (VILLALVA, 1994, 2000).

O grupo dos sufixos derivacionais compreende a grande maioria dos sufixos do português e é caracterizado por provocar alterações na base a que se anexa. Segundo Villalva (2000, p. 148), “os sufixos derivacionais têm diversas funções: determinam sempre o valor da categoria sintática da palavra em que ocorrem; das subcategorias morfológicas; das categorias morfossintáticas; e das categorias morfosssemânticas”. Em (1i), a partir do radical nominal (RN), a anexação do sufixo derivacional pode gerar nomes (N) ou verbos (V); em (1ii), é determinada a informação sobre a conjugação verbal; em (1iii), a forma resultante é masculina, apesar de o radical nominal ser feminino; em (1iv), o

radical com a propriedade [contável] faz parte da formação de um nome não contável, característica determinada pela anexação do sufixo derivacional.

(1) Sufixo derivacional (VILLALVA, 2000, p. 149)

Característica	Base
i. determina categoria sintática	simbol _{RN} – simbolismo _N / simbolizar _V
ii. determina subcategoria morfológica	dorm _{RV[3*conj]} – dormir _{V[1*conj]}
iii. determina categoria morfossintática	rac _{RN[+fem]} – racismo _[-fem]
iv. determina categoria morfossemântica	herói _{RN[+cont]} – heroísmo _[-cont]

Por determinarem as características da palavra complexa, os sufixos derivacionais são considerados predicadores transitivos, de acordo com Villalva (1994, 2000), em teoria que rotula como XM-barras. São, portanto, considerados os núcleos da estrutura complexa.

Por outro lado, há sufixos que não atuam como núcleo: é o caso dos sufixos avaliativos, que apresentam com a base uma relação de modificação³. Estes sufixos, segundo Scalise (1984) e Villalva (1994, 2000), (i) não alteram a categoria sintática da base; (ii) não alteram as propriedades morfossintáticas ou morfossemânticas da base; (iii) não alteram a estrutura argumental da base; (iv) alteram a interpretação semântica da base; (v) podem coocorrer em posições adjacentes; (vi) ocorrem à direita dos sufixos derivacionais; (vii) precedem os sufixos de flexão⁴.

(2) Sufixo avaliativo (VILLALVA, 2000, p. 150, 254-259)

Característica	Base
i. não altera categoria sintática	cas _N – casinha _{Ndim}
ii. não altera propriedades morfossintáticas ou morfossemânticas da base	cães _{N[-hum,+plu]} – cãezinhos _{N[-hum,+plu],dim}

³ A relação de modificação no âmbito dos sufixos é representada exclusivamente pelo grupo dos avaliativos. Contudo, segundo Villalva (2000, p. 256), “nem todos os sufixos que permitem interpretações diminutivas, aumentativas ou depreciativas são sufixos avaliativos”. A autora assume que constituintes que não obedecem a alguma destas propriedades sejam classificados como *pseudoavaliativos*.

⁴ O item (vii) diz respeito ao fato de, em uma sequência de sufixos de diferentes propriedades, os avaliativos ocorrem imediatamente à esquerda dos sufixos flexionais externos – os últimos da palavra. Isso não impede que sejam precedidos e seguidos por marcas flexionais (ex. *pasteizinhos*).

iii. não altera estrutura argumental	morder _[-- SN] – mordiscar _[-- SN]
iv. altera interpretação semântica	cedo – cedinho
v. coocorre em posições adjacentes	dedinhozinho
vi. ocorre à direita de sufixos derivacionais	certezinha
vii. ocorre à esquerda de sufixos flexionais	novinhos

As propriedades (i-iii) distinguem sufixos avaliativos de derivacionais, ao passo que as propriedades (iv-vii) diferenciam-nos de sufixos flexionais. Por terem como característica principal a manutenção de características da base, Villalva (1994, 2000) afirma que sufixos avaliativos exercem a função de modificação e não atuam como núcleo da estrutura complexa.

Apresentadas as propriedades dos sufixos derivacionais e dos modificadores avaliativos, na próxima seção comentaremos alguns pressupostos a respeito do processamento linguístico de estruturas morfológicas para, em seguida, avaliar se o processamento é mais custoso para algum dos grupos de sufixos apresentados.

2 Processamento morfológico de palavras complexas

176

Na literatura sobre processamento morfológico, há três modelos que pretendem dar conta da forma com que palavras são armazenadas e acessadas no léxico mental dos falantes: o modelo de computação plena, o modelo de listagem plena e o modelo de dupla rota (DOMÍNGUEZ et al., 2000).

No modelo de computação plena, assume-se que o léxico armazena partes menores do que a palavra. Esse modelo, apesar de estabelecer um limite de estocagem no léxico, é em geral interpretado como custoso para o componente computacional. No modelo de listagem plena, por outro lado, admite-se um acesso lexical direto — não há decomposição da palavra em partes menores. O custo desse modelo é a sobrecarga de itens armazenados no léxico dos falantes. Como conciliação entre as duas propostas, o modelo de dupla rota assume que as duas rotas podem ser utilizadas, e a rota vencedora depende de vários fatores linguísticos e/ou extralinguísticos, como alomorfa, frequência etc.

Para o português, estudos em morfologia experimental têm sugerido que há decomposição da palavra morfológica em partes menores (GARCIA, 2009; PINTO, 2017) — abordagem condizente com modelos de computação plena ou de dupla rota. Partindo do pressuposto de que unidades menores

do que a palavra são relevantes para o processamento linguístico, resolvemos investigar se há diferenças de processamento entre grupos de sufixos com estatutos distintos no léxico por meio de três paradigmas experimentais, um de medição *offline* e dois de medição *online*⁵.

Para os três testes elaborados, utilizamos 32 estímulos experimentais: 16 estímulos com o sufixo -eiro⁶ — sufixo derivacional que atua como núcleo do nome complexo — e 16 com o sufixo -inho — sufixo avaliativo que preserva as propriedades da base⁷.

Quadro 1 – Lista de estímulos da Tarefa de Associação de Palavras⁸

-inho				-eiro			
banquinho	livrinho	lixinho	barquinho	banqueiro	livreiro	lixeiro	barqueiro
portinha	pedrinha	casinha	cartinha	porteiro	pedreiro	caseiro	carteiro
conselhinho	cavalinho	sapatinho	tesourinho	conselheiro	cavaleiro	sapateiro	tesoureiro
madeirinha	fazendinha	pistolinha	cervejinha	madeireiro	fazendeiro	pistoleiro	cervejeiro

A amostra acima foi formada a partir de substantivos dissilábicos e trissilábicos (ex. *livro*, *cavalo*), de gênero único (ex. *o livro*, *a porta*), os quais aceitam formações tanto em -eiro quanto em -inho e formam substantivos complexos com significado predominantemente composicional (ex. *livreiro/livrinho*, *sapateiro/sapatinho*).

3 Tarefa de Associação de Palavras

A Tarefa de Associação de Palavras consiste em o informante ler um estímulo apresentado e informar a primeira palavra que lhe vem à cabeça

5 O método *offline* oferece pistas a respeito do processamento morfológico depois que ele já foi finalizado, por meio de respostas e reações dos participantes (ex. testes de aceitabilidade, testes de familiaridade). Já o método *online* permite avaliar reações dos participantes durante o processamento linguístico, ou seja, de modo simultâneo (ex. testes de decisão lexical, rastreamento ocular e neuroimagem), conforme Derwing e de Almeida (2005).

6 A escolha de -eiro é justificada pelo fato de este sufixo derivacional poder formar nomes a partir de nomes (ex. *jornal* – *jornaleiro*), o que permite comparar seu processamento ao de um avaliativo, sem que a classe da base ou da forma resultante atuem como variáveis intervenientes. Reconhecemos que este constituinte pode estar presente também em outros tipos de formações; essas, contudo, não fazem parte do escopo de nosso estudo experimental.

7 A partir de uma lista de dados do uso, iniciamos a busca por palavras formadas por -eiro que atendessem aos critérios formais estabelecidos. Filtramos os dados e excluímos (i) outras classes que não a de substantivo (ex. *financeiro*_{Adj}), (ii) bases não paroxítonas (ex. *hóspede* – *hospedeiro*), (iii) palavras derivadas de nomes próprios (ex. *Brasil* – *brasileiro*), (iv) alomorfa da base na palavra derivada (ex. *feijão* – *feijoeiro*) e (v) palavras com -eiro não agentivo (ex. *palmito* – *palmitoeiro*). Mantivemos aqueles que tivessem a possibilidade de apresentar um item com -inho na mesma família e, com base na lista de diminutivos, excluímos (i) bases que formam diminutivos intuitivamente menos familiares (ex. *justiça/justiceiro, justicinha*), (ii) bases com a consoante final similar ao início do sufixo diminutivo (ex. *mina/mineiro, *mininha*), (iii) formas complexas com significado predominantemente não composicional (ex. *vaquinha* ‘coleta de dinheiro’) e (iv) palavras em que a frequência da forma diminutiva é maior do que da forma com -eiro (ex. *tomateiro* < *tomatinho*).

8 Os dados aqui apresentados provêm de tese de doutorado de Ulrich (2021), realizada com bolsa de pesquisa CNPq, processo 141774/2017-7.

após a leitura por meio do preenchimento de uma caixa de diálogo aberta. O objetivo geral da tarefa é avaliar se o informante conhece o estímulo a ponto de estabelecer alguma associação de forma ou de sentido. Com os resultados obtidos, pretendemos validar a utilização dos estímulos selecionados para as tarefas *online*, podendo considerá-los suficientemente conhecidos pelos informantes.

3.1 Materiais e métodos

A fim de evitar que um mesmo informante lesse itens de uma mesma família lexical e, assim, tivesse sua resposta influenciada por um item já ativo no léxico, dividimos os 32 itens em duas listas, as quais são apresentadas no quadro abaixo.

Quadro 2 – Listas 1 e 2 da Tarefa de Associação de Palavras

Lista 1				Lista 2			
banqueiro	livrinho	lixeiro	barquinho	banquinho	livreiro	lixinho	barqueiro
porteiro	pedrinha	caseiro	cartinha	portinha	pedreiro	casinha	carteiro
conselheiro	cavalinho	sapateiro	tesourinho	conselhinho	cavaleiro	sapatinho	tesoureiro
madeireiro	fazendinha	pistoleiro	cervejinha	madeirinha	fazendeiro	pistolinha	cervejeiro

Em um primeiro momento, o participante era informado sobre os objetivos e critérios éticos da pesquisa e sobre a tarefa a ser executada. Logo após, respondia às questões sobre idade, grau de escolaridade e, em caso de ensino superior, curso de formação. Enfim, o questionário ancorado na plataforma Google Forms apresentava os estímulos formados por *-inho* e *-eiro* de forma alternada, sem itens distratores, e um estímulo era exibido a cada nova página, a fim de evitar que o participante comparasse sua resposta com as respostas já atribuídas a palavras anteriores.

Participaram da tarefa 100 pessoas, de 16 a 58 anos de idade (média = 31,37 / desvio-padrão = 11,47), sendo a maior parte concentrada na faixa de 20 a 40 anos. Em relação à escolaridade, 20 informantes possuem Ensino Médio completo, 50 possuem Ensino Superior em andamento ou completo e 30 possuem Pós-graduação em andamento ou concluída. Cada uma das listas foi respondida, portanto, por 50 informantes distintos.

3.2 Resultados

As respostas dadas pelos participantes foram classificadas em 4 categorias:

- a) relação morfológica — se havia a presença do radical (ex. *cavaleiro* – *cavalo*);
- b) relação semântica — se a palavra fazia parte do mesmo campo semântico (ex. *cavaleiro* – *armadura*);
- c) relação contextual/cultural (ex. *cavaleiro* – *Game of Thrones*);
- d) sem relação/desviante (ex. *cavaleiro* – *gentil*).

Considerando-se que o objetivo principal desta etapa é validar os estímulos para as tarefas de decisão lexical, estabelecemos o ponto de corte de 10% de respostas desviantes para que um item fosse excluído das etapas seguintes. Apesar de serem registradas algumas respostas equivocadas (ex. *cavaleiro* – *gentil*) para estímulos específicos, o único par de estímulos em que uma palavra teve 10% de respostas equivocadas foi *tesourinho/tesoureiro*. Foram recorrentes as respostas associando a forma derivada à palavra *tesoura*, e não a *tesouro*. Por esse motivo, esses itens foram substituídos por *escudinho/escudeiro* na etapa posterior⁹.

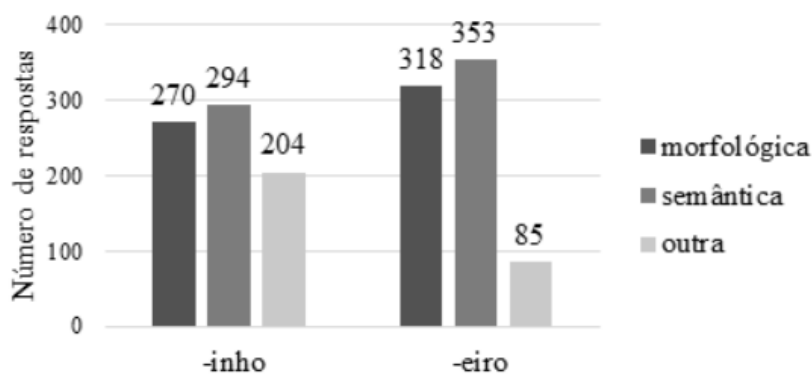
Os resultados gerais mostram que foram predominantes as respostas relacionadas a padrões morfológicos e semânticos, sem diferença significativa entre essas duas categorias principais ($X^2(323) = 327.33, p = 0.4$).

Tabela 1 – Média de tipos de associações na Tarefa de Associação de Palavras

morfológica	semântica	outra	sem relação
36,8%	40,5%	18,0%	4,5%

Se analisamos a distribuição de respostas por sufixo, vemos que a proporção entre respostas de cunho morfológico e de cunho semântico parece se manter nos dois grupos, e essa diferença não é estatisticamente significativa em nenhum deles (morfologia x semântica nos dados formados por *-inho*: $X^2(156) = 164, p = 0.31$ / morfologia x semântica nos dados formados por *-eiro*: $X^2(132) = 133.33, p = 0.45$).

Gráfico 1 – Tipo de associação e sufixo na Tarefa de Associação de Palavras



⁹ Apesar de os itens *escudinho* e *escudeiro* não terem sido testados, imaginamos que as formas complexas sejam facilmente identificáveis para os informantes, assim como a maior parte dos itens da TAP, já que não há outra forma similar na língua que possa causar desvio de interpretação, como acontece no par *tesoura* e *tesouro*.

A partir do gráfico acima, podemos afirmar também que palavras formadas pelo sufixo -inho permitem mais associações pautadas em elementos outros que não a forma da raiz ou o significado – seja ele composicional (ex. *pistolinha* – *arma*) ou não (ex. *pistolinha* – *brabinho*)¹⁰. A grande maioria dos dados classificados como *outra* está ligada a fatores culturais, como se observa na resposta *água* para o estímulo *pistolinha*, *Cinderela* para o estímulo *sapatinho* ou *Seu Lobato* para o estímulo *fazendeiro*.

Depois de validarmos o conhecimento dos falantes sobre os estímulos apresentados e realizarmos os ajustes necessários, passamos à Tarefa de Decisão Lexical.

4 Tarefa de Decisão Lexical

A Tarefa de Decisão Lexical exige o registro do julgamento do participante como resposta a uma pergunta feita sobre um estímulo linguístico. Esse registro, por meio de um botão ou tecla pressionada, por exemplo, permite que sejam computadas não só as respostas do participante sobre a pergunta, mas também o tempo que ele levou para tomar a decisão. O pressuposto básico desse tipo de investigação é de que estruturas menos complexas tenham menores tempos de reação, enquanto estruturas mais complexas gerem mais custos ao processamento lexical¹¹.

4.1 Materiais e Métodos

Na Tarefa de Decisão Lexical, medimos duas variáveis dependentes: (i) índice de acertos sobre ser ou não uma palavra conhecida pelos informantes e (ii) tempo de reação para a tarefa de decisão lexical (milissegundos)¹². A variável independente investigada é o tipo de estrutura morfológica, se formação avaliativa com o sufixo -inho ou derivação com o sufixo -eiro. Nosso objetivo específico é verificar se existe diferença de processamento entre grupos de afixos. Investigamos a hipótese de que palavras formadas por sufixos avaliativos, com função de modificação, apresentam menor tempo de resposta, dado que estes não alteram a categoria da base, não atuam

10 Apesar de haver uma relação morfológica pelo emprego do mesmo sufixo, classificamos como respostas de cunho morfológico apenas aquelas que mencionavam a raiz em questão. Na classificação semântica, registramos tanto análises composicionais – a grande maioria – como também algumas análises não composicionais, já que ambas permitem entender que o participante conhece o item em questão.

11 Por exemplo, Villalva e Pinto (2018) mostram que, dentre as palavras formadas por -oso, apresentam menor custo de processamento aquelas geradas regularmente (ex. *venenoso*), enquanto formas lexicalizadas — seja pela alomorfia da base (ex. *medroso*) ou do sufixo (ex. *luxuoso*) – apresentam maiores tempos de reação.

12 Os dados referentes ao tempo de reação apresentam uma distribuição não normal.

como núcleo da forma complexa e, conseqüentemente, podem facilitar o reconhecimento da base em comparação a formações com -eiro¹³.

Assim como na tarefa anterior, os 32 estímulos experimentais foram divididos em duas listas. Além dos itens de teste, cada lista contou com 48 itens distratores (16 palavras e 32 pseudopalavras, 75% da amostra) que respeitavam a mesma estrutura fonológica formal (posição do acento, vogal terminal e número de sílabas).

Concordando com o objetivo do estudo e os aspectos éticos da coleta de dados, os informantes responderam a um pequeno questionário social, a fim de controlarmos a influência de variáveis extralingüísticas. Em seguida, a tarefa consistia em ler o estímulo apresentado e responder à questão “você conhece essa palavra?”. Antes do estímulo, uma cruz ‘+’ servia como ponto de fixação do olhar do participante por 500 ms; em seguida, surgia o estímulo, em letras minúsculas, fonte Arial, tamanho 40, o qual permanecia por, no máximo, 5000 ms, ou até que o participante tomasse sua decisão. Ainda na etapa de instruções, os participantes passavam por uma rodada de treino com 6 itens não relacionados à amostra testada.

Participaram desta etapa 80 informantes¹⁴ (40 informantes por lista) entre 18 e 58 anos de idade (média = 31,76, desvio-padrão = 9,07 – o que reflete uma concentração de informantes na faixa de 20 a 40 anos). Em relação ao gênero, 31 (38,8%) respondem pelo gênero feminino e 49 (61,2%) pelo gênero masculino. Em relação à escolaridade, a maior parte dos informantes possui pós-graduação em andamento ou concluída (56,2%); 38,7% afirmam ter Educação Superior parcial ou completa e 5% cursaram até o Nível Médio.

O experimento foi desenvolvido e ancorado na plataforma Psytoolkit, versão 3.2.0 (STOET, 2010, 2017), disponível em <https://www.psytoolkit.org/>. Esta plataforma permite a participação de forma remota; por isso, cada participante participou de sua casa, sem a presença física dos pesquisadores. Para a análise estatística descritiva e inferencial destes dados, utilizamos o software R 4.0.3, interface RStudio versão 1.3.1093 (R CORE TEAM, 2020)¹⁵.

13 Contudo, ressaltamos o fato de que esses elementos possuem maior carga de informação acentual, pelo fato de se configurarem como palavras prosódicas (ULRICH; SCHWINDT, 2018) — fato que pode vir a inverter os resultados esperados.

14 O experimento de decisão lexical foi realizado em torno de 2 meses após a tarefa de associação. Pelo fato de a divulgação ser aberta e a participação ser anônima, não temos garantia de que as respostas foram dadas pelos mesmos ou por novos participantes.

15 Utilizamos o método remoto de coleta de dados em função da pandemia de COVID19. Pedimos aos informantes que comessem o teste apenas quando tivessem o tempo hábil de 7 minutos para finalizá-lo e que, durante a realização da tarefa, não tivessem distrações ao redor. Além disso, coletamos dados apenas de informantes destros, a fim de que a mão dominante não fosse uma variável interveniente nos índices de acurácia ou nos tempos de reação.

4.2 Resultados

Em relação à variável **índice de acertos**, o conjunto de 1280 dados experimentais — palavras complexas formadas por -eiro e por -inho — apresenta média geral de 99,0% de acertos.

A variável independente analisada nesta tarefa — **estrutura morfológica**, que divide os itens experimentais em dois tipos de formações morfológicas — mostra que os estímulos com -eiro apresentam maiores índices de acerto (99,2%) do que estímulos formados por -inho (98,9%).

Gráfico 2 – Índice de acertos (%) e estrutura morfológica na Tarefa de Decisão Lexical



A diferença, contudo, é bastante baixa. O teste chi-quadrado entre os índices de acurácia não mostra diferença significativa ($X^2(1) = 0.084122$, $p = 0.77$). Nossos resultados iniciais parecem refletir que os falantes, de fato, conhecem totalmente as palavras apresentadas e, por isso, não tiveram problemas na discriminação dos itens. As taxas de acerto foram bastante elevadas e não houve diferença significativa entre as condições.

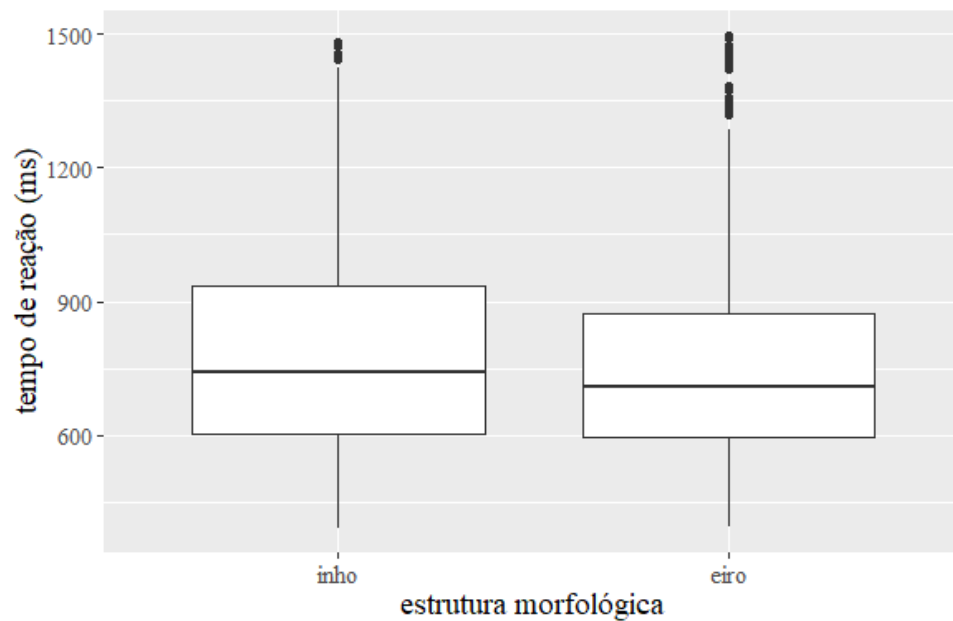
182

Para a variável **tempo de reação**, depois de retirados os *outliers*, o grupo dos itens experimentais (1202 dados: 599 com -inho e 603 com -eiro) apresenta mediana de 721 ms e 312 ms de intervalo interquartil (IQR)^{16,17}. Desmembrando esses valores por **estrutura morfológica**, os itens formados por -inho apresentam mediana de 740 ms/IQR 332 ms, enquanto os itens formados por -eiro apresentam mediana de 707 ms/ IQR 278 ms.

16 A amostra de dados tanto anterior quanto posterior à exclusão dos *outliers* apresenta uma distribuição não normal. Por isso, reportamos os valores de mediana e de intervalo interquartil (IQR, do inglês *interquartile range*).

17 Os resultados para a variável tempo de reação levam em conta apenas respostas corretas e não atípicas. Excluimos as respostas incorretas (12/1280), visto que elas podem representar equívocos de processamento das palavras analisadas, já que a resposta não reflete nenhum nível do processamento. Em seguida, excluimos os itens acima de 1,5 intervalo interquartil acima do terceiro quartil da amostra dos dados experimentais (66/1280).

Gráfico 3 – Tempo de reação (ms) e estrutura morfológica na Tarefa de Decisão Lexical



Apesar de os valores para -inho serem mais elevados, o teste Wilcoxon não revela significância para o contraste entre as duas condições ($W = 168829$, $p = 0.05048$). Levando-se em conta apenas a significância estatística, esse resultado sugere não haver diferença no processamento entre palavras com os dois sufixos analisados. Assim, formas presas seriam processadas igualmente, independentemente da função que possuem na palavra. Contudo, tal resultado precisa ser confrontado com análises de outros dos sufixos em questão a fim de confirmar a tendência encontrada.

5 Tarefa de Decisão Lexical com efeito de *Priming*

183

Para a última tarefa, adicionamos a testagem do efeito de *priming*¹⁸. Esse efeito consiste na pré-ativação de um estímulo (alvo) a partir da apresentação prévia de outro estímulo relacionado (*prime*). Nossa intenção é verificar se a apresentação da palavra base (ex. *cavalo*) causa diferença no tempo de reação nas diferentes condições morfológicas (ex. *cavalo* / *cavalinho* / *cavaleiro*).

5.1 Materiais e Métodos

Como na etapa anterior, medimos duas variáveis dependentes: (i) índice de acertos e (ii) tempo de reação (milissegundos). As variáveis independentes analisadas neste experimento são duas: (i) realidade da palavra — se palavra

18 Segundo França (2005, p. 75), o teste de *priming* serve para “desvendar aspectos da arquitetura do léxico mental, por exemplo, os critérios de agrupamento”.

ou pseudopalavra¹⁹ — e (ii) relação morfológica entre *prime* e *alvo* — se identidade (ex. *cavalo* – *cavalo*), modificação pelo avaliativo -inho (ex. *cavalo* – *cavalinho*) ou derivação por -eiro (ex. *cavalo* – *cavaleiro*). Perseguimos as hipóteses de que palavras serão processadas mais rapidamente do que pseudopalavras e de que a condição de identidade é a condição com menor tempo de reação, seguida da condição com -inho e, por fim, da condição com -eiro. Essas hipóteses partem da ideia geral de que formas reconhecidas ou facilmente identificáveis pelo falante serão processadas mais rapidamente.

A combinação de variáveis nos leva a seis condições distintas: (i) palavra com relação de identidade (ex. *lixo* – *lixo*); (ii) palavra com sufixo avaliativo (ex. *lixo* – *lixinho*); (iii) palavra com sufixo derivacional (ex. *lixo* – *lixeiro*); (iv) não palavra com relação de identidade (ex. *bido* – *bido*); (v) não palavra com sufixo avaliativo (ex. *bido* – *bidinho*); (vi) não palavra com sufixo derivacional (ex. *bido* – *bideiro*).

A tarefa conta com 16 palavras reais simples e 32 palavras reais complexas — as mesmas utilizadas na tarefa anterior. No caso das pseudopalavras, geramos automaticamente pelo script Word Generator (GARCIA, 2014) estruturas simples que obedecessem a todos os critérios estruturais estipulados para as palavras reais: equivalência em número de sílabas (2/3 na forma de base) e em relação à vogal terminal (o/a). A essas estruturas criadas acrescentamos os sufixos -inho e -eiro. Assim, temos 16 estímulos em cada uma das seis condições, totalizando 96 pares experimentais distintos.

Os pares de itens experimentais foram divididos em três listas, a fim de que cada informante só tivesse contato com um único tipo de relação morfológica para cada palavra base. Acrescentamos, ainda, 64 pares de itens distratores por lista (equivalente a 66% das listas).

Quadro 3 – Exemplo de distribuição dos itens experimentais na Tarefa de Decisão Lexical com *Priming*

Lista 1	Lista 2	Lista 3
banco - banco	banco - banquinho	banco - banqueiro
barco - barqueiro	barco - barco	barco - barquinho
lixo - lixinho	lixo - lixeiro	lixo - lixo
zirpo - zirpo	zirpo - zirpinho	zirpo - zirpeiro
jebro - jebreiro	jebro - jebro	jebro - jebriinho
bido - bidinho	bido - bideiro	bido - bido
...

19 Apesar do desejo de analisar pseudopalavras em ambas as tarefas de decisão lexical, na tarefa anterior elas foram utilizadas apenas como itens distratores devido à dúvida quanto à eficácia e ao bom uso da plataforma virtual por informantes de perfis tão distintos em uma abordagem completamente remota. A fim de coletarmos os dados durante a pandemia de COVID19, optamos por fazer um teste mais curto e garantir a atenção e a confiança dos participantes, mesmo sem a presença da pesquisadora. Depois de validado uso da plataforma, optamos por adicionar pseudopalavras à nova tarefa.

Assim como no experimento prévio, o informante era instruído a ler o estímulo na tela e informar se conhecia a palavra apresentada ou não. Para isso, deveria clicar na tecla “P” para estímulos conhecidos, usando sua mão dominante, ou na tecla “Q” para estímulos desconhecidos.

Na sequência temporal do experimento, surgia um ponto de fixação em formato de cruz ‘+’ no centro da tela por 500 ms para fixação do olhar do participante; logo após, surgiam o *prime* (base) com tempo de exposição de 50 ms e o alvo (forma idêntica ou formas complexas), que ficava na tela até a decisão do participante ou por, no máximo, 5 segundos²⁰. O *prime* e o alvo eram exibidos em letras minúsculas, fonte Arial, tamanho 40, também no centro da tela.

Participaram desta etapa 80 informantes²¹ entre 20 e 82 anos de idade (média = 32,82, desvio-padrão = 11,24). Em relação ao gênero, o conjunto de participantes conta um total de 43 mulheres (53,7%) e 37 homens (46,3%). Para a variável escolaridade, 5 informantes apresentam nível fundamental/médio/técnico (6,25%), 29 estão cursando ou já concluíram algum curso superior (36,25%) e 46 estão cursando ou já concluíram algum curso de pós-graduação (57,5%).

O experimento foi ancorado no software Psytoolkit versão 3.3.0 (STOET, 2010, 2017) e realizado de forma remota pelos participantes. Para a análise estatística descritiva e inferencial destes dados, utilizamos o software R 4.0.3, interface RStudio versão 1.3.1093 (R CORE TEAM, 2020).

5.2 Resultados

O grupo de dados experimentais (2560 dados divididos entre as 6 condições apresentadas acima) apresenta uma média geral de 99,2% de acertos.

Em relação à **realidade** dos itens experimentais, os resultados para a variável **índice de acertos** revelam que pseudopalavras (99,6%) apresentam maiores taxas de acerto do que palavras reais (98,8%).

20 Dada a grande variação existente nas escolhas metodológicas em estudos de *priming* morfológico, optamos por tomar como base as escolhas de Pinto (2017). A única diferença entre esta análise e a nossa diz respeito no limite de tempo estipulado para a exposição do alvo — escolha essa justificada por se tratar de uma coleta de dados a distância.

21 O experimento que inclui o efeito de *priming* foi realizado em torno de 2 meses após o experimento anterior. Pelo fato de a divulgação ser aberta e a participação ser anônima, não temos garantia de que as respostas foram dadas pelos mesmos ou por novos participantes. Contudo, caso os participantes sejam os mesmos, devido ao tempo de diferença entre as duas análises, imaginamos que não haja diferença nos índices de acurácia ou tempos de reação e que o conhecimento da plataforma possa representar uma vantagem nesse sentido.

Gráfico 4 – Índice de acertos (%) e realidade do estímulo na Tarefa de Decisão Lexical com *Priming*



O teste estatístico revela diferença significativa entre as duas condições ($X^2(1) = 4.8013$, $p = 0.02$). Esse resultado nos sugere que os falantes não hesitaram em dizer que não conhecem os itens inventados. Quanto às palavras reais, alguns itens específicos podem ter gerado dúvidas.

A análise da variável **relação morfológica** mostra uma escala decrescente da condição de identidade em um polo (99,8%) e a condição de avaliação em outro (98,6%) — mas, ainda assim, com índices de acertos bastante elevados.

Tabela 2 – Índice de acertos (%) e tipo de relação morfológica na Tarefa de Decisão Lexical com *Priming*

Relação morfológica	identidade	-inho	-eiro
estímulos	99,8%	98,6%	99,2%

O teste estatístico que avalia se há diferenças entre as três relações morfológicas mostra que nem todas elas apresentam o mesmo padrão quanto ao índice de acertos ($X^2(2) = 7.1811$, $p = 0.02758$). A diferença está na comparação entre a condição morfológica de identidade e a formação por -inho ($X^2(1) = 5.8123$, $p = 0.015$). Não há diferença significativa entre identidade e derivação ($X^2(1) = 1.7778$, $p = 0.18$) ou entre as duas condições complexas ($X^2(1) = 0.85158$, $p = 0.35$).

Cruzamos, então, as duas variáveis que são foco da análise: **realidade** do estímulo e **relação morfológica** entre *prime* e alvo.

Tabela 3 – Índice de acertos (%), realidade do estímulo e tipo de relação morfológica na Tarefa de Decisão Lexical com *Priming*

Realidade/Rel. morfológica	identidade	--inho	--eiro
palavra	99,5%	97,6%	99,0%
pseudopalavra	100%	99,5%	99,2%

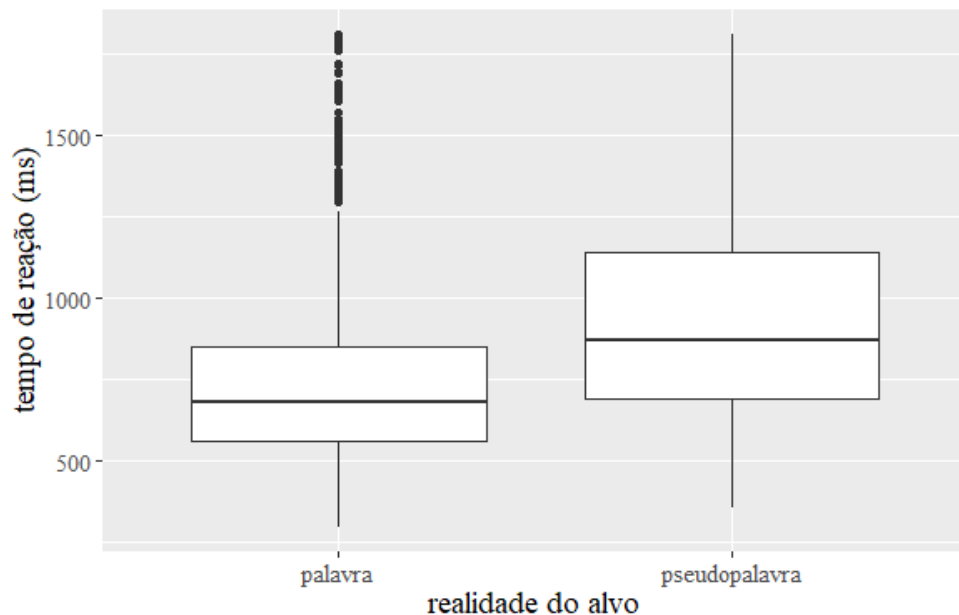
Como esperado, o teste mostra diferença significativa entre as seis condições ($X^2(5) = 17.113$, $p = 0.004$). Comparando as três relações morfológicas entre palavras reais, há significância estatística apenas entre palavras com relação de identidade e palavras com relação de modificação por -inho ($X^2(1) = 4.125$, $p = 0.04$). Dentre as pseudopalavras, nenhum dos grupos se opõe aos demais.

Em um extremo do contínuo de acurácia, temos pseudopalavras na condição de identidade, que apresentaram 100% de acertos. No outro extremo, temos formações com a anexação de -inho em palavras reais (ex. *lixinho*, *fazendinha*). Os altos índices de acurácia para a condição de identidade já eram esperados; um resultado, contudo, que merece destaque é a inversão de resultados para o sufixo -inho a depender da realidade do estímulo. Os dados sugerem que os falantes tendem a julgar mais corretamente itens diminutivos em palavras inventadas; palavras reais, por outro lado, podem sofrer efeitos externos aos aqui controlados.

Passamos agora à análise da variação **tempo de reação** (ms). Excluídos os itens com valores desviantes, os dados experimentais correspondem a 2385 itens e apresentam 756 ms de mediana e amplitude interquartil de 405 ms²².

Em relação à **realidade** do par experimental, vemos que palavras (mediana 680 ms / IQR 294 ms) são processadas mais rapidamente do que pseudopalavras (mediana 869 ms / IQR 454 ms).

Gráfico 5 – Tempo de reação (ms) e realidade do estímulo na Tarefa de Decisão Lexical com *Priming*



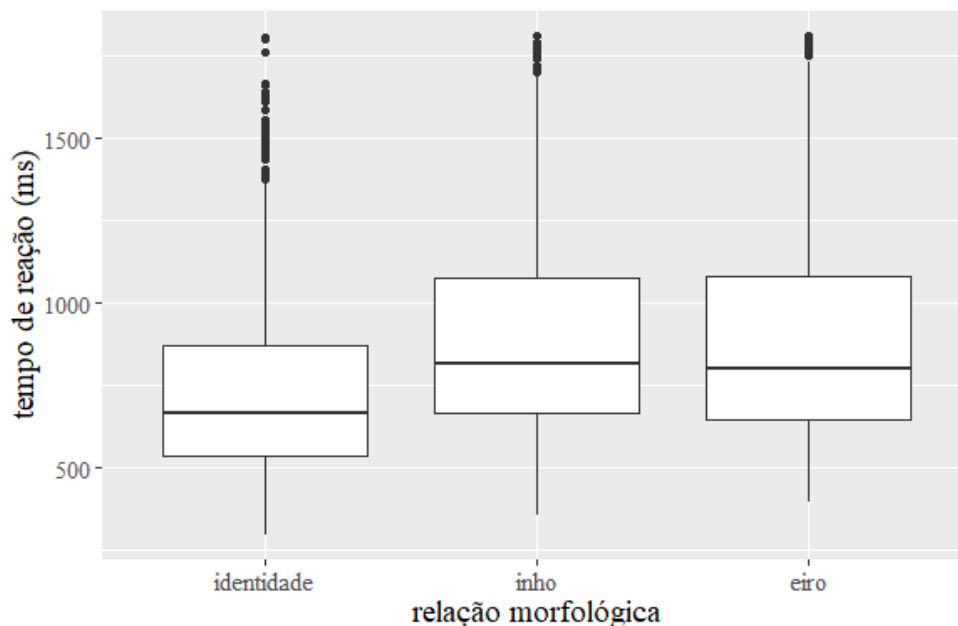
O teste de Wilcoxon mostra que há significância estatística na comparação dos dois grupos ($W = 438500$, $p < 2.2e-16$).

A variável tipo de **relação morfológica** mostra, conforme esperado, que a condição de identidade — em que o *prime* e o alvo são iguais e, portanto, sem mudança na tela (ex. *cavalo / cavalo*) — é mais rápida do que as demais. Nosso interesse maior, contudo, está na comparação entre as duas condições de palavras complexas. Descritivamente, vemos valores bastante próximos, com um mínimo aumento do tempo de reação em palavras terminadas por

22 Assim como na etapa anterior, a análise leva em conta apenas as respostas corretas e os dados não desviantes (< 250 ms ou $> 1,5$ IQR acima do terceiro quartil – neste caso, 1814 ms).

-inho (mediana 814 ms / IQR 413 ms) em relação a -eiro (mediana 800 ms / IQR 436 ms).

Gráfico 6 – Tempo de reação (ms) e relação morfológica na Tarefa de Decisão Lexical com *Priming*



O teste de Kruskal-Wallis mostra que há diferença entre as condições ($X^2(2) = 164.2, p < 2.2e-16$). O teste *post hoc* de Dunn mostra que a diferença está na comparação entre a condição de identidade e as demais ($p = 1.28e-26$ em comparação à derivação por -eiro; $p = 3.06e-29$ em comparação à modificação por -inho). Não há diferença significativa quando comparamos as duas relações morfológicamente complexas ($p = 5.23e-1$)²³.

188

Em conjunto, as variáveis **realidade** do estímulo e tipo de **relação morfológica** revelam padrões próximos aos encontrados na análise do índice de acertos.

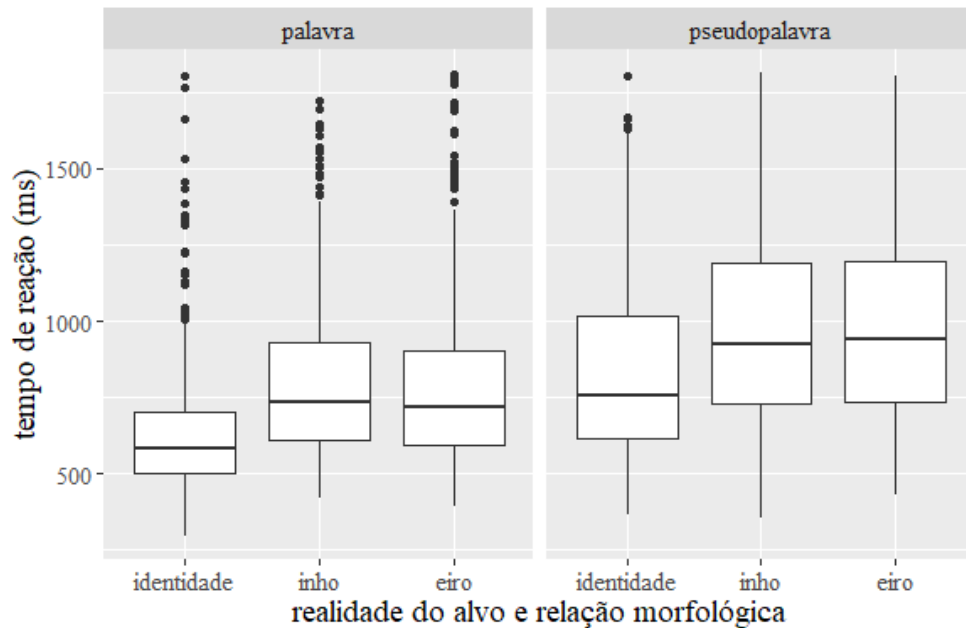
Tabela 4 – Tempo de reação (medianas e IQRs em ms), realidade do estímulo e tipo de relação morfológica na Tarefa de Decisão Lexical com *Priming*

Condições	identidade	-inho	-eiro
palavra	586 ms / 198 ms	736 ms / 320 ms	717 ms / 311 ms
pseudopalavra	755 ms / 400 ms	924 ms / 463 ms	942 ms / 460 ms

Os mesmos dados são apresentados no gráfico abaixo, para melhor visualização.

²³ Para o teste *post hoc*, apresentamos o valor estatístico de p ajustado.

Gráfico 7 – Tempo de reação (ms), realidade do estímulo e tipo de relação morfológica na Tarefa de Decisão Lexical com *Priming*



O teste de Kruskal-Wallis mostra que há diferença entre condições ($X^2(5) = 436.1, p < 2.2e-16$). O teste *post hoc* de Dunn aponta vários pares com diferença significativa:

(i) entre palavras reais, há diferença significativa entre as condições de identidade e modificação por -inho ($p = 1.99e-20$) e entre as condições de identidade e derivação por -eiro ($p = 3.37e-17$), e não há diferença entre as condições morfológicamente complexas ($p = 7.41e-1$);

(ii) entre pseudopalavras, da mesma forma, há diferença significativa entre as condições de identidade e modificação ($p = 6.20e-11$) e entre as condições de identidade e derivação ($p = 4.27e-11$), e não há diferença entre as condições morfológicamente complexas ($p = 9.71e-1$);

(iii) entre as diferentes realidades de estímulos, só *não* há significância entre palavras reais com o sufixo -inho e pseudopalavras em condição de identidade ($p = 7.57e-2$) e entre palavras reais com o sufixo -eiro e pseudopalavras em condição de identidade ($p = 4.51e-1$) — ou seja, entre as duas condições mais lentas de palavras reais e a condição mais rápida de pseudopalavras.

Os resultados confirmam a hipótese de que o processamento de pseudopalavras é mais custoso do que o processamento de palavras reais; dentre os fatores de cada uma das realidades, não há diferenças entre os dois sufixos investigados²⁴. Apesar disso, a análise descritiva mostra um cruzamento interessante para as condições complexas: palavras reais demoram mais a serem processadas na presença do sufixo -inho, enquanto palavras inventadas

24 Se comparados a dados sem relação morfológica, hipotetizamos que nossos resultados apresentariam facilitação devido à pré-ativação pela base (*prime*). Contudo, pelo fato de analisarmos, além da condição de identidade, dois grupos de constituições próximas (apesar de funções distintas), não podemos fazer nenhuma menção ao efeito de *priming* nestas condições complexas.

demoram mais a serem processadas na presença do sufixo -eiro. O resultado para pseudopalavras está de acordo com a hipótese investigada, já que a não influência do sufixo sobre as propriedades da base pode vir a tornar seu conhecimento mais facilitado. Além disso, tal resultado pode ter relação com efeitos externos aos investigados aqui. Uma análise de um conjunto maior de dados pode permitir conclusão mais contundente sobre manutenção ou não desse padrão.

6 Considerações finais

Neste artigo discutimos se há diferença de processamento entre dois tipos de sufixos no PB: sufixos derivacionais e sufixos modificadores. Para isso, uma sequência de tarefas *offline* e *online* foi proposta.

Após selecionar estímulos formados por -eiro e por -inho a partir de bases comuns na língua, vimos, por meio de uma Tarefa de Associação de Palavras, que esses itens parecem ser conhecidos e familiares para os falantes do português brasileiro. Passamos, então, para a análise do processamento online desses itens. Na Tarefa de Decisão Lexical, vimos que os dois tipos de sufixos parecem não apresentar diferenças de processamento para os informantes. Desenvolvemos, posteriormente, uma Tarefa de Decisão Lexical com efeito de *Priming*, a fim de verificarmos se a apresentação prévia da base causaria facilitação em uma das condições complexas. Analisamos palavras e pseudopalavras em três condições de relações morfológicas — identidade, derivação por -eiro ou modificação por -inho — e vimos que palavras formadas por -inho apresentam os menores índices de acerto e os maiores tempos de reação, apesar de a variável tempo não ter se mostrado significativa no teste.

Embora alguns valores apontem para maior tempo de reação em palavras formadas por -inho, os resultados inferenciais sugerem, de modo geral, que não há diferença entre as condições. Unicamente a partir destes dados, não podemos julgar se há alguma distinção entre grupos com funções distintas na palavra ou se os resultados são provenientes de alguma escolha metodológica pré-estabelecida. Por esse motivo, a continuidade da pesquisa enseja expandir a investigação das mesmas variáveis para outros sufixos da língua.

Referências

DERWING, B. L.; DE ALMEIDA, R. G. Métodos experimentais em Linguística. In: MAIA, M.; FINGER, I. (ed.). *Processamento da Linguagem*. Pelotas, Educat. 2005, p. 401-442.

DOMÍNGUEZ, A.; CUETOS, F.; SEGUI, J. Morphological processing in word recognition: a review with particular reference to Spanish data. *Psicológica* 21, 2000, p. 375-401.

FRANÇA, A. I. O léxico mental em ação: muitas tarefas em poucos milissegundos. *Linguística*, Rio de Janeiro, v. 1, n. 1, 2005, p. 45-80.

GARCIA, D. C. *Elementos estruturais no acesso lexical: o reconhecimento de Palavras multimorfêmicas no Português brasileiro*. Dissertação de mestrado. Rio de Janeiro: UFRJ/Faculdade de Letras. 2009.

GARCIA, G. D. *Word Generator: an R script for generating pseudo-random words*. GitHub repository disponível em https://github.com/guilhermegarcia/r/blob/master/word_generator.md, 2014.

PINTO, C. A. G. *O Papel da Estrutura Morfológica nos Processos de Leitura de Palavras*. Tese de Doutorado. Universidade de Lisboa, 2017.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2020.

SCALISE, S. *Generative Morphology*, Dordrecht: Foris, 1984.

SCHWINDT, L. C. Palavra fonológica e derivação em português brasileiro: considerações para a arquitetura da gramática. In: BISOL, L.; COLLISCHONN, G. (org.) *Fonologia: teorias e perspectivas*. 1ª ed. Porto Alegre: EDIPUCRS, p. 15 - 28, 2013a.

STOET, G. PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 2010, p. 1096-1104.

STOET, G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 2017, p. 24-31.

ULRICH, C. W. *O estatuto morfológico e prosódico da sufixação avaliativa em português brasileiro*. Tese de Doutorado. Universidade Federal do Rio Grande do Sul, 2021.

ULRICH, C. W.; SCHWINDT, L. C. O status morfoprosódico dos sufixos -inho/-zinho, -mente e -íssimo no português brasileiro. *Revista Delta*, v. 34, n. 2, 2018. Disponível em <https://revistas.pucsp.br/index.php/delta/article/view/39419/26676> (acesso em 10 de jun. de 2020).

VILLALVA, A. *Estruturas Morfológicas: Unidades e Hierarquias nas Palavras do Português*. Tese de doutoramento. Universidade de Lisboa, 1994. 399f.

VILLALVA, A. *Estruturas Morfológicas: Unidades e Hierarquias nas Palavras do Português*. Lisboa: FCT e FCG, 2000. 399 f.

VILLALVA, A.; PINTO, C. Complexidade morfológica e custos de processamento lexical. *Revista Alfa*, vol. 62, n. 1, 2018. Disponível em https://www.scielo.br/scielo.php?pid=S1981-57942018000100151&script=sci_arttext (acesso em 10 de jun. de 2020).