



Universidade Federal do Rio de Janeiro  
Faculdade de Letras

HÁ UMA FRONTEIRA ENTRE A FLEXÃO E A DERIVAÇÃO NO PROCESSAMENTO?  
O CASO DO PARTICÍPIO PASSADO DO PORTUGUÊS BRASILEIRO

Leonardo Cabral

Rio de Janeiro  
2022



Universidade Federal do Rio de Janeiro  
Faculdade de Letras

HÁ UMA FRONTEIRA ENTRE A FLEXÃO E A DERIVAÇÃO NO PROCESSAMENTO?  
O CASO DO PARTICÍPIO PASSADO DO PORTUGUÊS BRASILEIRO

Leonardo Cabral

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários para a obtenção do título de Mestre em Linguística

Orientadora: Profa. Doutora Marije Soto

Rio de Janeiro  
Março de 2022

## CIP - Catalogação na Publicação

CL581h Cabral, Leonardo  
Há uma fronteira entre a flexão e a derivação no processamento? O caso do participípio passado do português brasileiro / Leonardo Cabral. -- Rio de Janeiro, 2022.  
111 f.

Orientadora: Marije Soto.  
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Faculdade de Letras, Programa de Pós Graduação em Linguística, 2022.

1. Processamento lexical. 2. Morfologia. 3. Psicolinguística. 4. Participípio passado. I. Soto, Marije, orient. II. Título.

HÁ UMA FRONTEIRA ENTRE A FLEXÃO E A DERIVAÇÃO NO PROCESSAMENTO?  
O CASO DO PARTICÍPIO PASSADO DO PORTUGUÊS BRASILEIRO

Leonardo Cabral

Orientadora: Professora Doutora Marije Soto

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários para a obtenção do título de Mestre em Linguística.

Examinada por:

---

Presidente, Profa. Doutora Marije Soto

---

Profa. Doutora Anieli Improta França – UFRJ

---

Prof. Doutor Gustavo Estivalet – UFPA

---

Profa. Doutora Juliana Novo Gomes – UFRJ, Suplente

---

Prof. Doutora Katia Abreu – UERJ, Suplente

Rio de Janeiro

Março de 2022

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior – Brasil (CAPES), no Programa de Excelência Acadêmica (PROEX) – código de financiamento 001.

## DEDICATÓRIA

Dedico esta dissertação a duas pessoas: minha mãe, Audrey, a luz da minha vida; e a meu irmão, Serginho, quem desejo que cresça sabendo o valor do conhecimento.

## AGRADECIMENTOS

Tantos para agradecer, que só posso pensar na sorte que tenho por ter tantas pessoas boas próximas a mim.

Em primeiro lugar, agradeço a minha mãe, Audrey, que me apoia em tudo que faço e que preciso, sendo fundamental para minha vida, e a meu pai, Luciano, cujo auxílio me viabilizou toda minha carreira acadêmica.

Em segundo lugar, agradeço a minha família, por tudo em que me ajudam, e a meu amirmão, Diogo, cujo laço comigo é mais forte que qualquer laço de sangue.

Em terceiro lugar, agradeço a Daniela Feijó e Marcelle Castro. Não sabem o quanto foram (e são) importantes!

Declaro minha gratidão também a todos os professores que me trouxeram até aqui, desde o maternal até a universidade. Devo citar por nome os meus professores da UFRuralRJ que, na graduação, fizeram toda a diferença para minha formação acadêmico-profissional. No curso de História, prof. Marcos Caldas, o maior – e, provavelmente, mais humilde – erudito desse país (e minha inspiração). No curso de Letras, professores Claudia Barbieri, João Carlos Lopes, Angela Bravin, Simone Batista, Elisa Abrantes, Gerson Rodrigues, Mikaela Roberto, Marli Pereira, Alexander Luz, Luiza Oliveira, Maria do Rosario Roxo, Roza Palomanes, Simone Batista, Wagner Costa, Maria das Graças Salgado, Anderson Gomes, Marcos Pasche e Roberto Bozzetti, todos importantíssimos para minha formação humana, linguística e profissional.

Não poderia deixar de dedicar um agradecimento especial a duas professoras. Prof. Claudia Rebello, minha ex-orientadora que nunca deixou de me orientar, sempre me dando forças e me incentivando. Prof. Adriana Lessa, que despertou o linguista dentro de mim e terminou de me entregar aos estudos da Linguística.

Agradeço – demais! – a todos que vim a conhecer na pós-graduação em Linguística da UFRJ. Professora Aniela, muito obrigado por ter me recebido tão bem nas reuniões do laboratório e sempre ter disposição para trocar ideias e responder minhas perguntas. Meus colegas Julia Cataldo, Matheus Gomes, Wellington Almeida, Ana Carolina, vocês têm noção do quanto vocês enriqueceram minha experiência nessa pós?! Demais colegas do grupo Neuróticas e do Lab ACESIN – Emily, Isadora, Diane e tantos outros –, cada reunião, cada conversa com vocês é algo de positivo a mais que vocês contribuem para minha vida acadêmica; não poderia ter conseguido colegas de laboratório melhores! E não poderia deixar de dar meu

obrigado aos incríveis professores com quem tive a grande oportunidade de aprender nessa pós: Aléria Lage, Alessandro Boechat, Marcus Maia, Ana Calindro, Gean Damulakis, Maria Cecília Mollica, Marcelo Melo e Márcio Leitão.

E, enfim, dedico este último parágrafo a minha orientadora, Marije Soto. Nem toda reunião foi fácil, esses anos foram difíceis para todos nós, mas obrigado por me dar algo que esperar de bom toda semana, quando precisamos nos encontrar para debater a pesquisa. Espero poder retribuir toda a sua gentileza e toda a dedicação que você direciona ao seu trabalho.



## RESUMO

### HÁ UMA FRONTEIRA ENTRE A FLEXÃO E A DERIVAÇÃO NO PROCESSAMENTO? O CASO DO PARTICÍPIO PASSADO DO PORTUGUÊS BRASILEIRO

Leonardo Cabral

Resumo da Dissertação de Mestrado submetida ao Programa de Pós-graduação em Linguística, Faculdade de Letras, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Linguística.

Reporta-se dois experimentos que objetivaram investigar como o particípio passado de primeira conjugação do português brasileiro é processado. O primeiro experimento, de tarefa de decisão lexical com *priming* encoberto, investigou a hipótese de que particípios com leitura preferencialmente nominal seriam processados pela forma inteira quando sua frequência superficial fosse alta, enquanto os preferencialmente nominais de baixa frequência e preferencialmente verbais seriam segmentados em morfemas. Leminen *et al.* (2013) sugerem que palavras derivadas de alta frequência são armazenadas em forma inteira, enquanto palavras flexionadas sempre seriam segmentadas. A derivação equivaleria aqui aos casos de particípio nominal, enquanto os particípios verbais seriam flexões. Dois efeitos foram encontrados: facilitação do processamento por *priming* morfológico independente de categoria da palavra; e facilitação por alta frequência sem interação com a categoria. Os resultados indicam refutação da hipótese, mas apontou-se uma hipótese alternativa: o experimento, por lidar com palavras isoladas, pode ter ativado uma entrada lexical verbal. Uma entrada lexical referente a um particípio adjetivo em forma inteira poderia não ter sido ativada, sua ativação podendo ser dependente de contexto sintático. Para testar esta hipótese, elaborou-se um experimento de leitura automonitorada em que se inseriu os particípios, antes classificados como preferencialmente nominais, em sintagmas nominais de sentenças (particípio adjetivo) ou em perífrases verbais (particípio verbal). Não houve efeito de frequência ou de categoria. Os resultados dos experimentos sugerem que o particípio passado de primeira conjugação é sempre segmentado e refutaram a hipótese alternativa. Contribui-se, assim, para o debate ainda pouco explorado sobre o processamento de palavras morfológicamente ambíguas, entre flexão e derivação.

Palavras-chave: processamento lexical; morfologia; particípio passado; psicolinguística.

## ABSTRACT

### HÁ UMA FRONTEIRA ENTRE A FLEXÃO E A DERIVAÇÃO NO PROCESSAMENTO? O CASO DO PARTICÍPIO PASSADO DO PORTUGUÊS BRASILEIRO

Leonardo Cabral

*Abstract* da Dissertação de Mestrado submetida ao Programa de Pós-graduação em Linguística, Faculdade de Letras, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Linguística.

Two experiments are reported which aimed to investigate how the past participle of first conjugation from Brazilian Portuguese is processed. The first experiment, a lexical decision task with masked priming paradigm, investigated the hypothesis that past participles with a preferentially nominal reading would be processed via full form when their surface frequency were high, while the preferentially nominal of low frequency and the preferentially verbal would be decomposed into morphemes. Leminen *et al.* (2013) suggested that high frequency derived words are stored in full form, while inflected words would always be decomposed. The derivation would be equivalent to the nominal participles, and the verbal participles would be inflections. Two effects were found: processing facilitation by morphological priming independently of word category; and facilitation by high frequency without interaction with category. The results indicate the hypothesis' refutation, but an alternative hypothesis was formulated: the experiment, for dealing with isolated words, may have activated a verbal lexical entry. A lexical entry referring to an adjective participle in whole form may not have been activated, its activation being dependent of syntactic context. To test this hypothesis, a self-paced reading experiment was prepared, in which the participles previously classified as preferentially nominal were inserted into noun phrases of sentences (adjective participle) or in verbal periphrasis (verbal participle). There was no effect of frequency or category. The results of the experiments suggest that the past participle of first conjugation is always decomposed and refuted the alternative hypothesis. Thus, the research contributed to the debate about the processing of morphologically ambiguous words, between inflection and derivation.

Keywords: lexical processing; morphology; past participle; psycholinguistic.

# Sumário

1 INTRODUÇÃO	1
2 FLEXÃO, DERIVAÇÃO E O CASO DO PARTICÍPIO PASSADO	8
2.1 Flexão <i>versus</i> derivação	8
2.1.1 Algumas definições em gramáticas normativas	8
2.1.2 Definições da morfologia estruturalista	9
2.1.3 Gramática Gerativa	9
2.1.4 Gramática de Construções	12
2.1.5 Conclusões a partir do exposto até este ponto	13
2.2 O que dizer do particípio passado no português brasileiro?	14
3 PROCESSAMENTO MORFOLÓGICO: MODELOS, SUAS EVIDÊNCIAS E VARIÁVEIS QUE O AFETAM	22
3.1 Modelos não-decomposicionais	23
3.1.1 Propostas conexionistas	23
3.1.2 Naive discriminative reader	24
3.2 Modelos decomposicionais	24
3.2.1 Decomposição precoce automática	25
3.2.2 Modelos de dupla rota de Pinker (1991) e Schreuder e Baayen (1995)	27
3.2.4 Decomposição tardia: modelo supraleixal	28
3.3 Processamento de morfologia e modelos de representação	29
3.4 O que dizem os achados da literatura sobre os modelos?	31
3.4.1 Psicolinguística	31
3.4.2 Evidências de estudos com EEG	36
3.4.3 Evidências de estudos com MEG	39
3.4.4 Evidência de estudos com fMRI	40
3.4.5 Conclusões sobre os modelos de processamento a partir da literatura	42
3.5 Variáveis que afetam o processamento lexical e morfológico	42
4 EXPERIMENTO 1	47
4.1 Metodologia	47
4.1.1 Hipótese e expectativas	47
4.1.2 Desenho experimental e materiais	48
4.1.3 Participantes	53
4.1.4 Procedimentos	54
4.2 Análise estatística	55
4.3 Resultados	56
4.3.1 Acurácia	56

4.3.2 Tempos de resposta	56
4.4 Discussão	60
5 EXPERIMENTO 2	64
5.1 Metodologia	64
5.1.1 Hipótese e expectativas	64
5.1.2 Desenho experimental e materiais	65
5.1.3 Participantes	67
5.1.4 Procedimentos	67
5.2 Análise estatística	68
5.3 Resultados	70
5.3.1 Taxa de acurácia	70
5.3.2 Tempo de leitura	70
5.4 Discussão	74
6 CONCLUSÃO	79
REFERÊNCIAS	83
<b>APÊNDICES</b>	<b>91</b>

## LISTA DE FIGURAS

<b>Figura</b>	<b>Descrição</b>	<b>Página</b>
1	Figura adaptada de Halle (1973).	11
2	Árvore adaptada de De belder, Fauster e Lampitelli (2009).	12
3	Árvore adaptada de Medeiros (2008, p. 185) de participípios de estado alvo.	18
4	Árvore adaptada de Medeiros (2008, p. 185) de participípios de estado resultante.	18
5	Figura adaptada de Leminen <i>et al.</i> (2013).	39
6	Proposta de análise de Matchin <i>et al.</i> (2019, p. 119) das condições controladas em seu experimento.	42
7	<i>Screenshot</i> de parte do pré-teste, realizado via Google Formulários.	51
8	Figura ilustrativa da etapa principal do experimento.	55
9	Figura ilustrativa do teste de leitura automonitorada seguido da tarefa de compreensão.	68

## LISTA DE GRÁFICOS

<b>Gráfico</b>	<b>Descrição</b>	<b>Página</b>
1	Resultado do agrupamento em clusters a partir da análise de <i>kmeans</i> .	52
2	Comparação de efeito de <i>priming</i> em palavras e em pseudopalavras.	57
3	Distinção entre médias de Tempo de Resposta (em milissegundos) de categorias.	58
4	TRs médios, em milissegundos, por tipo de <i>prime</i> (morfologicamente relacionado ou não-relacionado), frequência (alta ou baixa) e categoria (tendência nominal ou tendência verbal).	59
5	TRs médios, em milissegundos, apenas de palavras reais por frequência (alta ou baixa) e tipo de <i>prime</i> (não relacionado ou com relação morfológica).	60
6	Médias de tempo de leitura (eixo y) do segmento crítico ( <i>segm_crit</i> ) e do primeiro segmento pós-crítico ( <i>segm_pos1</i> ) em milissegundos no subconjunto de dados de sentenças com participípio adjetivo, cujas tarefas de compreensão foram respondidas corretamente.	73
7	Médias de tempo de leitura (eixo y) do segmento crítico ( <i>segm_crit</i> ) e do primeiro segmento pós-crítico ( <i>segm_pos1</i> ) em milissegundos no subconjunto de dados de sentenças com participípio verbal, cujas tarefas de compreensão foram respondidas corretamente.	74

## LISTA DE TABELAS

<b>Tabela</b>	<b>Descrição</b>	<b>Página</b>
1	Distribuição em listas dos estímulos por tipo de <i>prime</i> no formato de quadrado latino do Experimento 1.	49
2	Exemplos de estímulos por condição e tipo de <i>prime</i> no Experimento 1.	50
3	Exemplos de estímulos por condição do segmento crítico e por classe da palavra do Experimento 2.	66
4	Comparações aninhadas entre modelo nulo e modelo mais complexo de efeitos mistos para TLs para o segmento crítico.	71
5	Médias de tempo leitura de sentenças cujas tarefas foram respondidas corretamente, com remoção de <i>outliers</i> .	72

# 1 INTRODUÇÃO

Esta dissertação se debruça sobre uma questão muito investigada, mas ainda longe de ter uma resposta definitiva: como nosso cérebro organiza o conhecimento sobre palavras e as processa palavras? Uma pergunta que reflete nosso recorte é a seguinte: como nosso cérebro processa palavras de morfologia complexa no particípio passado? Para responder a segunda pergunta, e adicionar evidências que apontem caminhos para complementar a primeira, dois experimentos foram criados para aferir se há, em primeiro lugar, processamento de informações morfológicas ou, ainda, morfo-ortográficas e, em segundo lugar, se há diferenças entre o processamento de particípio passado quando este exibe propriedades verbais, o que caracteriza um caso de flexão, e quando ocorre com propriedades nominais, representando um caso de derivação.

Toma-se as seguintes sentenças como exemplo:

- (1) O garoto tinha *irritado* o pai.
- (2) O pai *irritado* brigou com o garoto.

Como será mostrado no capítulo 3, não há dúvidas do caráter verbal da palavra em destaque em (1), "irritado", que é um caso de particípio passado de primeira conjugação (-AR). Por se tratar da conjugação de um verbo, considera-se este um caso indubitável de flexão verbal. Mas em (2), a palavra "irritado" faz parte de um sintagma nominal e atua como modificadora do nome, "pai". Neste exemplo, caso presume-se que a palavra base para a formação de irritado é o verbo "irritar", então há uma alteração de classe gramatical da palavra de verbo para adjetivo, já que a palavra exibe propriedades adjetivais e nominais (além de atuar como modificadora, também pode ser flexionada em gênero e número), fato este que é abordado por Foltran e Crisóstimo (2005). Por esses motivos, pode-se considerar o caso em (2) como uma derivação, já que a este processo se associa a alteração de classe de palavras, o que será debatido no capítulo 2.

As perguntas feitas e investigações conduzidas são parte de um campo de estudo ou, ainda, um campo de estudo e seu subcampo: os estudos sobre o léxico mental e o processamento de palavras, respectivamente. É desafiador definir o que seria o "léxico mental", já que diferentes correntes da linguística o tratam de formas diferentes. Intuitivamente, a definição

mais ampla seria que o léxico mental é um repositório de palavras na memória. Mas a Morfologia Distribuída, por exemplo, divide o que seria esse repositório em três listas separadas (MARANTZ, 1997) e as “palavras” são formadas via computação, já que suas informações estariam armazenadas na memória em constituintes. Para outras teorias, como a de Bybee (1995), esse repositório se refere a redes em que há conexões semânticas entre formas fonológicas. Toma-se provisoriamente, então, a seguinte definição de Jarema e Libben (2008):

“O léxico mental é o sistema cognitivo que constitui a capacidade para atividade lexical consciente e inconsciente.”

Esta definição é abrangente e parece representar o cerne do que se espera em termos de capacidade para uso de palavras, a atividade consciente enquanto a capacidade do falante de intuir sobre o léxico e seu repertório, e o inconsciente enquanto os processos cognitivos de atividade lexical. Mas ela é pouco informativa quando se questiona sobre a organização do léxico, já que se pode considerá-lo como uma rede de relações entre formas fonológicas e significados, conforme a proposta conexionista, ou como um repositório de raízes categoricamente neutras, traços formais e definições com associações a formas fonológicas, como propõe a Morfologia Distribuída. Como uma ou outra corrente pode assumir ou negar a existência de certos tipos de informação, ou mesmo de níveis hierárquicos inteiros (por exemplo, um nível estritamente morfológico), talvez a seguinte definição de Emmorey e Fromkin (1988) seja mais adequada:

“O léxico mental é aquele componente da gramática que contém toda a informação - fonológica, morfológica, semântica e sintática - que falantes sabem sobre palavras individuais e/ou morfemas”.

Mas aí esbarramos em uma outra barreira, ao inquirir-se se o léxico mental é parte da gramática mental. Apesar de ser mais informativa e abordar aspectos do léxico que são retratados em várias correntes, esta definição pode ser problemática para estudos que busquem um ponto de partida a partir do qual uma investigação mais incontroversa sobre como o léxico se organiza tendo por base pressupostos oriundos de estudos psicolinguísticos, no lugar daqueles de base mais teórica.



Seja ele um componente em si ou um conjunto de componentes que emergem como a capacidade de reconhecer e utilizar palavras, o léxico mental suscita muitas perguntas sobre sua representação na memória. Destaca-se, enfim, que o léxico mental e suas definições se constituem como metáforas, uma vez que sua organização e contraparte neurobiológica ainda estão sob estudo.

Se de fato há como diferenciar os casos de participio passado de (1) e (2) em termos de classe da palavra, o que isso diria sobre o léxico mental? Se de fato o participio em (1) é uma derivação e em (2) é uma flexão verbal, isso implicaria processos de formação diferentes no âmbito do léxico mental e/ou da gramática mental, com implicações para como a linguagem é processada?

A natureza do léxico mental pode ser investigada tanto através de estudos mais teóricos quanto experimentais, sendo a disciplina de “processamento de palavras” parte dos últimos. Esta disciplina se debruça sobre como o córtex cerebral reconhece palavras e as utiliza, o que pode ser um estudo em si ou mesmo uma forma de intuir sobre o léxico mental (GARCIA, 2015), o qual é mais abstrato. O processamento de palavras ou processamento lexical se enquadra também em uma área de estudo mais ampla, a de processamento da linguagem.

A área que se especializa em investigar conhecimentos linguísticos mais abstratos assim como as vias que esses conhecimentos são processados pelo cérebro é a psicolinguística. Esta área utiliza experimentos *online* e *off-line* para discutir a partir de seus resultados o “como” do processamento de linguagem e as representações que o subjazem (MAIA, 2015). Associada a tal área, a neurociência da linguagem aplica técnicas próprias, como o eletroencefalograma (EEG) e o magnetoencefalograma (MEG), para especificar o “como” do processamento, investigando o “quando” e o “onde” dos potenciais estágios envolvidos no processamento da linguagem (RODDEN; STEMMER, 2008). Tais técnicas da neurociência serão retratadas em maiores detalhes no capítulo 2.

Uma das questões que perpassa tanto questões relativas ao léxico mental quanto ao processamento de palavras é aquela debatida por modelos de processamento lexical e a que nos dirigimos neste trabalho, sobre se o processamento utiliza conhecimento abstrato, formal, na forma de representação linguística sobre morfemas quando processa palavras. O modelo brevemente citado anteriormente de decomposição automática precoce diz que sim. Os modelos de dupla rota também mencionados diriam “sim, mas depende”. Modelos conexionistas como o de Bybee (1995) responderiam que não, já que a morfologia seria uma relação semântica entre palavras mediada pela fonologia. Os resultados dos experimentos desta pesquisa apontam que,

sim, há processamento com base em informações morfológicas formais, mais especificamente morfo-ortográficas.

Alguns estudos psico-neurolinguísticos apontam para uma decomposição automática das palavras (*e.g.*, PYLKKANEN *et al.*, 2004 e MORRIS; STOCKALL, 2012 em inglês, FRANÇA *et al.* 2008 em português) e, conseqüentemente, um léxico mental que representa as palavras de morfologia complexa em constituintes. Outros apontam variáveis que podem influir em como o cérebro processa palavras derivadas e flexionadas e como ele representa essas palavras, como a frequência de ocorrência da palavra (*e.g.*, BAAYEN; DIJKSTRA; SCHREUDER, 1997 em holandês, ESTIVALET, 2020 em português).

Mas uma variável pouco explorada como preditora para modelos de dupla rota é o processo de formação da palavra. Um estudo em finlandês que será mais bem detalhado no capítulo 4 sugere que palavras derivadas de alta frequência são armazenadas em forma inteira na memória, enquanto as derivadas de baixa frequência e flexionadas independentemente da frequência seriam segmentadas em constituintes (LEMNEN *et al.*, 2013). Parte do argumento é a alta produtividade das flexões em finlandês comparada com a produtividade das derivações, o que tornaria custoso para a memória armazenar em forma inteira todas as palavras flexionadas de alta frequência. Uma investigação sobre a variável “processo de formação” é necessária em português brasileiro que exhibe propriedades semelhantes ao apontado no finlandês e inquiriu-se se a mesma variável teria impacto em como processamos palavras de morfologia complexa em nossa língua.

O que instigou ainda mais a curiosidade neste trabalho foi o que aconteceria com uma morfologia que pode ser ambígua, a do particípio passado. Esta morfologia possui propriedades nominais quando ocorre como adjetivo (FOLTRAN; CRISÓSTIMO, 2005) e verbais quando ocorre em tempo composto (MEDEIROS, 2008). O que ocorreria, então, no processamento de uma morfologia que aparenta estar na fronteira entre flexão e derivação, já que quando ocorre como adjetivo representa uma derivação, mas quando ocorre como verbo representa uma flexão? Encontrar-se-ia um processamento dividido de acordo com a variável “processo de formação”, assumindo-se que há, de fato, uma divisão dos casos de particípio passado em nominal e verbal?

Assim, o objetivo geral deste trabalho é observar se o processamento de palavras de morfológica complexa do particípio passado é diferente quando este ocorre como resultado de uma derivação (sua ocorrência nominal, por exemplo “rapaz ocupado”) e quando ocorre como uma flexão (sua ocorrência verbal, “tinha ocupado”). Como na teoria linguística a flexão e a

derivação podem ser tratadas como operações distintas ou semelhantes na essência (por exemplo, na MD considera-se que ambas são sintáticas), um caso como o do particípio passado que é ambíguo pode ajudar a observar a existência dessa separação ou, de fato, uma semelhança entre as operações. Qualquer resultado que apontasse para a possibilidade de separação das operações no processamento seria relevante, já que a morfologia do particípio passado, por motivos a serem expostos posteriormente, está enviesada para a existência em um possível *continuum* ou fronteira por conta de suas propriedades. Dentre os objetivos específicos estão: investigar a organização do léxico mental de falantes nativos do português brasileiro; prover dados que demonstrem a ocorrência de um determinado modelo de processamento morfológico (decomposição precoce automática ou dupla rota); diferenciar o processamento de particípio passado de primeira conjugação nominal do verbal. A hipótese geral deste trabalho é que o particípio nominal se distingue no processamento do particípio verbal, ao passo que casos de palavras no particípio nominal de alta frequência podem ser armazenados em forma inteira, enquanto demais condições do particípio são sempre segmentadas. Há duas hipóteses alternativas. A primeira prevê que uma mesma palavra no particípio de alta frequência (e.g., “convidado”) possui uma entrada lexical na memória em que é nominal e está armazenada em forma inteira, enquanto sua forma verbal (e.g., quando ocorre em tempo composto, “tinha convidado”) está segmentada em morfemas. O que ativaria uma ou outra entrada lexical seria o contexto sintático. A segunda hipótese alternativa é de que as palavras no particípio passado de primeira conjugação são sempre segmentadas, independentemente de suas condições de frequência e categoria, o que dialogaria com o modelo de decomposição precoce automática e com modelos mentais como a Morfologia Distribuída, a qual prevê que o léxico mental se organiza em morfemas computados ao longo da derivação sintática.

A relevância deste trabalho justifica-se pela necessidade de aprofundar-se o conhecimento sobre a organização do léxico mental e como o cérebro processa palavras, especificamente, e a linguagem, de forma geral. Em nível micro, apesar do crescimento de pesquisas psicolinguísticas e de neurociência da linguagem no português brasileiro, grande parte dos estudos ainda são de outras línguas, o que justifica a necessidade de mais estudos sobre português brasileiro. Em nível macro, conforme será mostrado no capítulo 3, as evidências sobre processamento linguístico e, também, sobre processamento lexical e morfológico se desencontram quanto aos modelos que apoiam. Logo, dados sobre o português brasileiro podem contribuir, também, para a compreensão da linguagem e de seu processamento em nível global.

Os experimentos realizados foram desenhados com base nos estudos previamente citados e nas questões formuladas. Controlou-se a frequência superficial e categoria dos participípios para observar o impacto destas variáveis no reconhecimento visual de palavras no participípio passado de primeira conjugação. Outras variáveis foram mantidas constantes para que qualquer resultado não pudesse ser atribuído a possíveis *confounds*, ao invés dos modelos investigados, como o tamanho das palavras e o segmento final da palavra (terminadas em “-o”, comum ao gênero gramatical masculino que ocorre na forma nominal do participípio passado e à desinência “-ado” que corresponde à mesma morfologia em verbos de tempo composto).

O primeiro experimento utilizou o paradigma de *priming* encoberto em tarefa de decisão lexical para aferir a resposta às duas perguntas. Como será explicitado no capítulo 4, o *priming* encoberto é um paradigma conhecido por evidenciar processamento de natureza morfo-ortográfica (i.e., informações ortográficas, formais que são associadas à morfemas sem associação com o significado destes), já que facilita o acesso à raiz de palavras morfologicamente relacionadas (RASTLE; DAVIS; NEW, 2004). Haveria, então, três possibilidades: haveria facilitação em todas as condições, o que evidenciaria um modelo de processamento morfológico que prevê decomposição automática precoce de palavras lidas (TAFT; FORSTER, 1975); haveria facilitação em todas as condições exceto em uma, quando o participípio seria um caso nominal de alta frequência, o que levaria à conclusão de que este caso de participípio acarreta em armazenamento da palavra em forma inteira (i.e., sem que haja informação sobre os constituintes da palavra na memória) e proveria evidências a modelos de dupla rota (PINKER, 1991; SCHREUDER; BAAYEN, 1995); por fim, uma última alternativa seria a ausência de efeitos facilitadores em todas as condições, o que dialogaria com modelos que não preveem existência de informações morfológicas (e.g., BYBEE, 1995). Como será mostrado no capítulo 5, os resultados podem ser interpretados como favoráveis ao modelo de decomposição automática, mas não de forma definitiva, já que uma outra explicação não pode ser atestada neste tipo de experimento: a existência de mais de uma entrada na memória para palavras com a mesma forma ortográfica (e.g., “convidado” como adjetivo teria uma entrada na memória, “convidado” como verbo teria outra), cujo acesso se daria por restrições sintáticas.

Para investigar a explicação alternativa, foi criado um experimento de leitura automonitorada, em que o contexto sentencial restringe a leitura do participípio como adjetivo (e.g., “o rapaz convidado”) ou como verbo (e.g., “tinha convidado”). Os resultados, no entanto, não foram conclusivos e não permitiram aprofundar a resposta à pergunta feita, mesmo que

para refutar a hipótese inicial. Isto indica que mais dados são necessários ou mesmo um experimento com maior refinamento metodológico.

A dissertação foi dividida em 6 capítulos, sendo este o primeiro. O capítulo 2 explora as definições de flexão e derivação na teoria linguística assim como o tratamento dado ao particípio passado, com base nos dois processos de formação de palavras. O capítulo 3 mostra os achados da literatura que embasam diferentes modelos de processamento morfológico e as variáveis que afetam este processamento. Os capítulos 4 e 5 explicitam a metodologia, resultados e discussão referentes aos experimentos 1 e 2, cada qual em um dos respectivos capítulos. Enfim, o capítulo 6 apresenta a conclusão deste trabalho.

## 2 FLEXÃO, DERIVAÇÃO E O CASO DO PARTICÍPIO PASSADO

Apesar de ser intuitivamente claro o que é derivação e o que é flexão, cabe realizar uma revisão de como diferentes esferas dos estudos de linguagem tratam destes fenômenos. Inicia-se a revisão com o tratamento dado pela Gramática Tradicional brasileira. A seguir, tratar-se-á de como a morfologia estruturalista realiza esta distinção. Posteriormente, considerações serão realizadas no tocante ao tratamento dado a esses processos pela gramática gerativa e pela Morfologia Distribuída. Por fim, olhar-se-á para contribuições da Gramática de Construções.

É importante que fique claro como os conceitos de derivação e flexão foram tratados no correr desta pesquisa, já que estão no centro do debate. Também no centro, está o recorte desta pesquisa, o particípio passado, sobre o qual será feita uma discussão no fim deste capítulo.

### 2.1 Flexão *versus* derivação

#### 2.1.1 Algumas definições em gramáticas normativas

Cunha e Cintra (2014) não dão uma definição para derivação. Os autores adotam no lugar disso uma definição para a expressão "formação de palavras" como a criação de unidades novas com base em morfemas. A este processo, os autores associam a derivação e a composição, cuja diferença para estes linguistas reside na preservação de sentido na derivação, o que não necessariamente ocorreria na composição.

A derivação recebe sua tradicional divisão em prefixal e sufixal. Cunha e Cintra explicam que esta última, foco deste trabalho, é responsável pela formação de substantivos, verbos, adjetivos e advérbios. Esta característica da morfologia derivacional é central para as pesquisas sobre morfologia e é destacada também em outras correntes, já que ela é a principal diferenciação entre a derivação e a flexão. Um exemplo de derivação sufixal do livro de Cunha e Cintra é a palavra "crueldade", com o sufixo "-dade" classificado como sufixo nominal.

Sobre a flexão, há ainda menos debate direcionado a esse ponto em Cunha e Cintra. Eles apenas atribuem à flexão variações nos nomes ou nos verbos. Nos nomes, as variações são de número, gênero e grau. Nos verbos as variações são de número, pessoa, modo, tempo, aspecto e voz.

Bechara (2019) direciona ainda menos atenção à diferenciação entre derivação e flexão. O autor aponta a diferença entre os processos com foco em exemplos de como a flexão expressa concordância e possui alternância vocálica, enquanto a derivação decorre de acréscimos de sufixos lexicais ou derivacionais.

### 2.1.2 Definições da morfologia estruturalista

Aqui recorre-se ao trabalho de Mattoso Camara Jr (2015) acerca da morfologia estruturalista no português brasileiro. Para este autor, a derivação é responsável pela formação de novos vocábulos. Caso leve-se em conta o que dizem Cunha e Cintra (2014) sobre a derivação sufixal, não há algo de diferente. Mas ao dirigir-se à flexão, Mattoso Camara Jr chama atenção para uma característica que retorna em outras discussões sobre morfologia flexional: a obrigatoriedade e sistematização.

Camara Jr (2015) argui que a derivação não possui sistematicidade já que nem todas as palavras derivadas possuem um outro caso correlato com o mesmo tipo de afixo. Um exemplo dado pelo autor é o que ocorre com "cantar" e "cantarolar", derivação que não possui formações semelhantes com outros verbos como "falar". Ainda, é dado o exemplo da morfologia de grau, aqui associada à derivação, já que sua ocorrência não depende de uma necessidade obrigatória.

O oposto ocorre com a flexão, ao menos em português brasileiro, segundo Camara Jr. As palavras flexionadas decorrerem de flexões obrigatórias exigidas, por exemplo, por efeitos de concordância de gênero ou número. Com relação a sua sistematicidade, também se chama atenção para a existência paradigmática que a flexão possui, em que o espaço para variação seria muito pequeno. Por exemplo, enquanto verbos como “destruir” e “relacionar” são transformados em substantivos por sufixos distintos (“destrui-ção” e “relacionamento”), a conjugação de verbos de primeira conjugação no presente do indicativo na terceira pessoa seria sempre pela desinência “-(a)m”, como em “canta-m” e “dança-m”, já que obedeceriam um paradigma sistemático com pouca variação.

### 2.1.3 Gramática Gerativa

A distinção entre derivação e flexão toma proporções maiores nos estudos de morfologia gerativa. Adiciona-se ao debate o lugar da morfologia no espectro linguístico (em qual módulo está a morfologia? Há um componente próprio? Derivação e flexão ocupam espaços diferentes?) e o pertencimento dos morfemas na estrutura sintática (*i.e.*, certos morfemas podem

funcionar como cabeças em estruturas? Possuem ou atribuem propriedades de traços? Ou sequer estão na estrutura sintática?). Não se deixa de lado, no entanto, o debate sobre características já retratadas no estruturalismo, com relação à sistematicidade e obrigatoriedade.

Há três hipóteses principais acerca da morfologia na gramática gerativa: a lexicalista forte, a lexicalista fraca e a sintática forte. Estas hipóteses giram em torno das questões apresentadas no parágrafo anterior. Para a hipótese lexicalista forte, ainda que a flexão tenha relação maior com a sintaxe, tanto esta quanto a derivação possuem seu próprio domínio, o do léxico; assim, as palavras seriam geradas no léxico, de onde o processo de derivação sintática as recuperaria (HALLE, 1973). Assumindo a sensibilidade da flexão à sintaxe, a hipótese lexicalista fraca pressupõe a existência de um léxico onde ocorrem processos de formação da palavra por derivação, enquanto a flexão ocorreria no componente sintático, já que esta última seria relevante para a sintaxe, enquanto a primeira, não (ANDERSON, 1982; ARONOFF, 1976, 2008). Enfim, a hipótese sintática forte está relacionada à corrente conhecida como Morfologia Distribuída, segundo a qual os processos de formação de palavra remetem a processos sintáticos (sobre a flexão, HALLE; MARANTZ, 1993; sobre a derivação, ARAD, 2003), sendo o léxico dividido em três listas que alimentam a derivação sintática em momentos diferentes: léxico estrito, onde se encontram os feixes de traços a serem inseridos na estrutura sintática; o vocabulário, onde se encontram as informações fonológicas correspondentes aos traços ou feixes de traços e que serão inseridos após a derivação sintática; e a enciclopédia, que contém as informações de significado (MARANTZ, 1997).

O lexicalismo surge a partir do artigo de Chomsky (1970) que serviu como um marco dos estudos de morfologia, "Remarks on Nominalization". Nele, Chomsky rebate a visão de que toda a formação de palavras é oriunda da nominalização de verbos a partir de regras transformacionais. Isto dá origem às considerações feitas por Halle (1973) que propõe um componente morfológico que opera dentro do léxico mental para gerar palavras, tanto via derivação quanto flexão, motivo pelo qual a proposta inicial de Halle é considerada como uma hipótese lexicalista forte, que é retomada por Chomsky (1995) no Programa Minimalista. Halle aponta idiosincrasias mesmo em alguns casos de flexão com exemplos do Russo, apesar de propriedades idiosincráticas serem tipicamente atribuídas à derivação, e Halle assim o faz como parte de seu argumento em favor de um "filtro" que impede a geração de formas não existentes na língua pelas regras de formação de palavras (*e.g.*, o filtro/*filter*, como presente na Figura 1, impediria a geração de uma palavra como "*arrivation*" para o dicionário de palavras reais, sendo a forma correta no inglês *arrival*, "chegada").



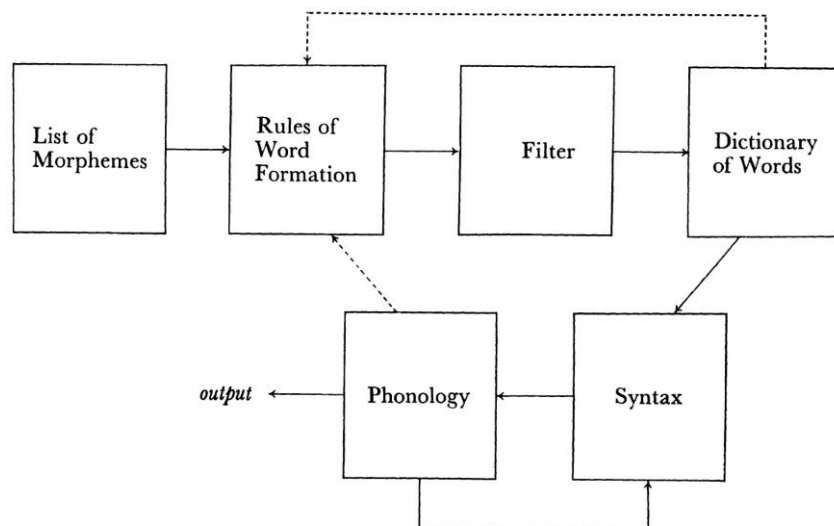


Figura 1 – Figura adaptada de Halle (1973). Proposta de organização do léxico, que separa os componentes lexicais do componente sintático.

Mesmo no bojo da Morfologia Distribuída, inclusa no rol da hipótese sintática forte em que toda morfologia faz parte da sintaxe, distinções são feitas entre derivação e flexão. Arad (2003) se debruça sobre os casos de derivação e mostra que eles também fazem parte da sintaxe. O que Arad (2003) demonstra com exemplos de hebraico e inglês é que alterações da semântica da raiz ocorrem em decorrência da primeira concatenação entre raiz e primeiro morfema categorizador. A partir desta primeira concatenação, segundo a autora, não é mais possível realizar alterações semânticas na raiz e, se há mais concatenações de morfemas com a palavra formada pela raiz, a cabeça referente a tal morfema só tem acesso às informações semânticas da cabeça que está abaixo, e não à raiz. Ainda assim, isso não significa um fim da diferença entre os processos, o que se pode observar a partir do tratamento que a autora dá aos seus dados no hebraico como sendo explicitamente casos derivacionais que se tratam de formação de palavras, uma expressão que alguns autores usam para se referir apenas aos processos de derivação e composição.

Um trabalho que exhibe uma lógica semelhante é o estudo de De Belder, Fauster e Lampitelli (2009) que aborda o caso do diminutivo no italiano. Para os autores, o morfema diminutivo que entra abaixo da primeira camada lexical altera a semântica palavra (não com uma desconexão completa da semântica original), como acontece com a palavra "*panino*" (sanduíche) e associam a este processo não-composicional a derivação. Desse modo, o que caracterizaria a flexão ou a derivação é seu domínio estrutural, segundo os autores, constituindo uma derivação apenas o que estivesse abaixo da primeira camada lexical "Lex". Na Figura 2,

por exemplo, o traço diminutivo de “*panino*” estaria sob o nódulo “Lex<sup>o</sup>”, já que neste caso há uma alteração semântica da raiz, segundo os autores; caso se tratasse de um diminutivo flexional, composicional, o nódulo que o abrigaria seria o “Size<sup>o</sup>”.

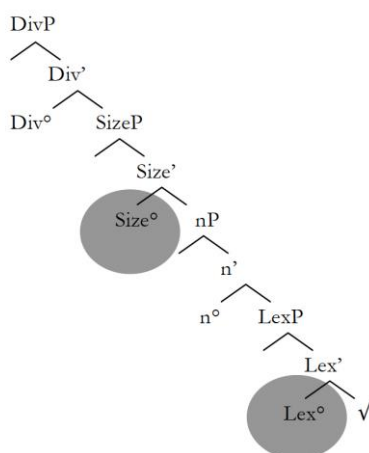


Figura 2 - Árvore adaptada de De belder, Fauster e Lampitelli (2009). Camada destacada acima flexional, e a mais abaixo derivacional.

O enfoque no lexicalismo forte e na Morfologia Distribuída nos parágrafos anteriores tem uma motivação. Ambas as teorias colocam a derivação e a flexão dentro de uma mesma competência, só que na lexical para a primeira corrente, e na sintática para a segunda. Ainda assim, conforme explicitado anteriormente, não deixam de lado distinções entre os processos de flexão e derivação.

#### 2.1.4 Gramática de Construções

Booij (2006) argumenta que a diferença entre flexão e derivação é primariamente funcional e retoma a noção de que lexemas (*i.e.*, entradas que se referem à forma inteira de uma palavra) podem ser criados pela derivação a partir de um lexema base enquanto a flexão altera formas de um mesmo lexema. O autor, contudo, debate os critérios formais que permeiam os processos de formação de palavra.

Não há dissonância sobre o que Booij fala sobre flexão como sendo obrigatória. Mas em outros critérios, como de relevância para a sintaxe, o autor apresenta pontos de vista que se afastam dos mencionados anteriormente. Em primeiro lugar, Booij dá um exemplo de como a escolha pela forma do plural de *book* na sentença "*John read these books*" não é governada pela

sintaxe, mas sim uma escolha do falante para expressar uma dada informação. Em contrapartida, Booij diz que a derivação também não é sempre sem relevância para a sintaxe, pois o fato de que a derivação altera uma categoria sintática já é, em si, relevante para a sintaxe. Um exemplo dado de relevância da derivação para sintaxe é da sentença "*John whitened the walls*", em que a derivação de um verbo causativo gera um verbo transitivo o qual requer um objeto direto.

No que tange a mudança de categoria sintática, atribuída comumente à derivação, Booij também traz um exemplo em que a flexão não mantém uma neutralidade categorial. O autor menciona o caso do infinitivo, em que um verbo pode se combinar com determinantes e adjetivos (o que é característico de substantivos), e ainda manter seu aspecto verbal. A sentença que Booij utiliza para ilustrar seu ponto é do holandês, "*Het je moeder cadeautjes geven moet afgelopen zijn*" ("Essa coisa de dar presentes à sua mãe deveria parar"). Um ponto interessante é que Cunha e Cintra (2014) e Bechara (2019) classificam o infinitivo no português como um afixo derivacional que gera verbos a partir de radicais, ao contrário do aqui expresso por Booij, que considera o infinitivo como um caso de flexão, o que exprime a ideia de que as divisões entre derivação e flexão podem não ser tão claras em alguns pontos.

Outras complexidades são trazidas por Booij. O autor exemplifica como flexões, consideradas mais paradigmáticas, podem não o ser sempre, com exemplos do latim (ex.: um mesmo morfema “-i” expressa número e gênero na palavra “*hort-i*”/jardim). A produtividade alta, associada também à flexão, não é sempre o caso também segundo o autor (ex.: em inglês, há o plural formado pelo fonema /z/ mas há também conjuntos fechados irregulares do tipo *phenomenon/phenomena*, fenômeno/fenômenos). Essas complexificações mostram que não é simples a tarefa de dividir processos derivacionais de flexionais com base nesses critérios, aos quais comumente recorrem-se.

Ainda assim, o próprio autor aponta que algumas destas complexidades são pontuais, em apenas algumas línguas. Além disso, Booij defende que certas afirmações como a de Bybee (1985 *apud* Booij, 2006) sobre a proximidade da derivação à raiz, dada a sua relevância para semântica da palavra, não são contestadas, demonstrando que há um reconhecimento de distinção, ainda que gradual, dos processos de derivação e flexão apesar das complexidades indicadas.

### 2.1.5 Conclusões a partir do exposto até este ponto

A partir do exposto, é possível estabelecer algumas delimitações aos conceitos de derivação e flexão a serem adotados neste trabalho. Primeiramente, ainda que haja casos em que há uma contrariedade referente aos critérios atribuídos a cada processo, a derivação é fortemente relacionada às alterações de ordem semântica e de categoria sintática. Já a flexão, exceto por um apontamento pontual de Booij (2006), está ligada à expressão de informações fundamentalmente gramaticais que não impactam a categoria ou a semântica da palavra.

O enfoque desta dissertação não é uma análise teórica, formal, abstrata dos processos de formação de palavras. O enfoque aqui adotado é no que os dados comportamentais de participantes dizem a respeito de seu processamento linguístico. Apesar disso, a análise teórica suscita critérios formais e funcionais úteis para estudos de natureza comportamental que buscam investigar como a linguagem é processada já que, em algum nível, este processamento pressupõe uma forma de conhecimento. Por exemplo, se há processamento diferenciado para palavras derivadas e para palavras flexionadas, presume-se que há representação linguística que distingue tais processos de formação da palavra, além de uma representação em constituintes no léxico mental (*i.e.*, em raízes e afixos) que justifique, também, o processamento decomposicional de palavras de morfologia complexa, em comparação com palavras monomorfêmicas que seriam processadas pelas suas formas inteiras.

A subseção a seguir dirigir-se-á ao particípio passado e as perguntas que ele suscita. Caso adote-se a noção de que há claramente uma fronteira entre o que é derivação e o que é flexão, um exemplo de caso de morfologia como o particípio passado é o tipo de fenômeno que suscita algumas dúvidas sobre a firmeza desta divisão.

## 2.2 O que dizer do particípio passado no português brasileiro?

Inicia-se esta subseção com uma descrição do que é o particípio passado. Esta morfologia é tida como uma flexão verbal com aspectos nominais por gramáticas normativas (CUNHA; CINTRA, 2014; BECHARA, 2019). Ela ocorre quando, na primeira conjugação, acrescenta-se a vogal temática "-a" e a desinência "-do" à palavras da primeira conjugação (ex.: apagar, *apagado*) ou a vogal temática "-i" e a desinência "-do" quando a palavra pertence à segunda conjugação (ex.: mexer, *mexido*) e à terceira conjugação (ex.: fingir, *fingido*). A estas formas atribui-se a qualidade de regulares. Quanto às irregulares, estas apresentam formas ditas mais curtas que o equivalente particípio regular quando esta é uma opção, como no caso do par

entregar-*entregue*, também chamados de participípios abundantes, e formas irregulares como dizer-*dito* em que não há na língua uma forma regular alternativa (como "dizido\*").

Bechara (2019) e Cunha e Cintra (2014) atribuem o participípio passado, dentre outros casos, à perífrases verbais de tempo composto, como em:

(1) Ela já tinha **chegado**. (exemplo próprio)

O recorte deste trabalho é o participípio passado regular de primeira conjugação, e os motivos para tal serão explicitados no próximo capítulo. Apresenta-se abaixo uma revisão do que a literatura fala sobre o participípio passado regular e, ao final desta subseção, algumas observações serão deixadas quanto ao participípio irregular. O foco é apresentar discussões que elucidem sua alocação a um processo flexional apenas (o que não envolve nenhuma controvérsia) ou também a um processo derivacional quando gera nomes (substantivos/adjetivos), o que por sua vez já pode ser considerado controverso.

Não há muita informação advinda de gramáticas normativas sobre o participípio passado neste tocante. Ele é colocado como uma forma verbal que também possui propriedades nominais, por exemplo, por Bechara (2019), e é apontado como parte de perífrases verbais em casos de tempo composto, como mencionado anteriormente, e à casos de voz passiva com verbos de cópula (ser, estar), como no caso a seguir:

(2) Ela foi **beijada**. (apenas a palavra foi em destaque foi utilizada pelo autor)

É interessante que a palavra no participípio possua concordância de gênero, uma vez que este é um atributo de classes nominais. Quando ocorre como adjetivo, o participípio também é apontado como uma oração adjetiva reduzida, como no seguinte exemplo de Bechara:

(3) “Os anais ensanguentados da humanidade estão cheios de facínoras, *empuxados* (= que foram empuxados) ao crime pela ingratidão injuriosa de mulheres muito amadas, e perversíssimas”. (BECHARA, 2019, p. 542)

Cunha e Cintra (2014) tampouco caracterizam o participípio como um possível caso de derivação. Esta morfologia não faz parte dos sufixos nominais que geram substantivos ou

adjetivos e são atribuídas, igualmente, a tempo composto, voz passiva e orações reduzidas adjetivas ou adverbiais.

Já Perini (2010) faz uma divisão clara de participípios em verbais e nominais. Para o autor, o participípio nominal é sempre um adjetivo já que considera que ele não seja uma oração reduzida, considerando-o sempre um sintagma adjetival. Ademais, existe uma irregularidade semântica no participípio segundo o autor que não há, por exemplo, no infinitivo e no gerúndio, como o que ocorre na palavra “intrometido” que atribui o papel de agente ao sujeito, como aquele que se intromete, enquanto uma palavra como “batido” não atribui papel agentivo. Por outro lado, o participípio verbal é aquele cujas características são retratadas também pelos outros autores em sua ocorrência como parte de tempo composto.

A pergunta que já se faz aqui é se não há a possibilidade de o participípio ser um adjetivo lexicalizado, sem que se presuma uma estrutura sintática inerente a ele, paralela, ou que se origine de derivação morfológica. Ou, se caso exista uma estrutura sintática inerente, se é possível uma lexicalização com ela preservada. O que chamarei aqui de lexicalização, seguindo Leminen *et al* (2013), é um processo em que uma palavra torna-se um item lexical, no sentido de que esse é acessado pelo falante na sua forma inteira, ou derivada com morfema adjetival direto, sem pressupor que o falante faça uma análise da estrutura interna como derivada de um participípio verbal.

Se levarmos em conta estudos de base gerativa, da Morfologia Distribuída, pode ser que não. Por exemplo, Rodrigues e Nevins (2014) realizaram um estudo de não-palavras no participípio passado (e.g., pseudo-verbo *mebulir*, na perífrase “tinha *mebulido/mebulo*”) que investigou a hipótese de Lobato (1999) que dizia que os casos de participípios regulares de maior comprimento (por exemplo, “expulsado” *versus* “expulso”, este último sendo irregular) poderiam ter relação com uma maior estrutura sintática subjacente. Os achados de Rodrigues e Nevins parecem favorecer esta hipótese, ao mostrar que participantes em um experimento tipo *wug* demonstravam uma preferência pela forma “mais curta” ou “mais longa” a depender da quantidade de argumentos na estrutura argumental (menor no primeiro tipo de forma, c.f. RODRIGUES; NEVINS, 2014). Este resultado pode ser problemático para teorias que prevejam uma “lexicalização” da palavra formada a partir do participípio, de forma que sua forma inteira perca a estrutura interna de constituintes. No entanto, o uso de pseudopalavras cujo propósito foi anular possíveis efeitos de frequência impede que demais considerações possam ser feitas com base na frequência, que pode ser indicadora de lexicalização em correntes fora da teoria gerativa.

Uma análise também com base na MD de Medeiros (2008) reconhece a qualidade de adjetivo do particípio em certos casos. Mas apesar de a noção de que um particípio pode alterar a classe de uma palavra, de verbo para adjetivo, a proposta de Medeiros (2008) também enquadra a estrutura sintática do particípio adjetival em um processo de derivação sintática que, também, pressupõe informações segmentadas e que não necessariamente correspondem a um caso de derivação morfológica, já que na MD toda a morfologia é parte da sintaxe. Aquilo que o autor denomina de particípio adjetivo está na estrutura da passiva adjetiva.

No exemplo abaixo dado por Medeiros, o autor afirma que o verbo “ser” na sentença é parte da estrutura morfológica, ocupando o lugar de flexão, já que ele não indica um estado:

(4) A bola foi **chutada** pelo zagueiro. (MEDEIROS, 2008, p. 157)

A sentença (4) possui, inclusive, um agente, o que mostra que a palavra destacada não está relacionada a um estado, e por isso a configuração em (4) se caracteriza como um particípio verbal em uma estrutura passiva eventiva (pois denota um evento). Medeiros diferencia (4) de (5), uma passiva estativa:

(5) A porta está aberta. (MEDEIROS, 2008, p. 158)

Em (5) o autor argumenta que um sintagma preposicional que indique o agente não é aceito pela estrutura sintática, já que um estado não possui um agente. Este caso é o que o autor denomina como passiva adjetiva, e diferencia o particípio verbal do que ocorre na passiva adjetiva. Ademais, na voz passiva, o particípio indica um evento, ao contrário do que ocorre na adjetiva. O exemplo (6) mostra um caso mais próximo do que será investigado nesta pesquisa quanto a adjetivo, em que o particípio ocorre em um sintagma nominal:

(6) **A maçã mordida** apodreceu na geladeira. (MEDEIROS, 2008, p. 161).

Em uma leitura baseada em Lieber (1980 *apud* MEDEIROS, 2008), Medeiros considera que uma palavra como “mordida”, neste contexto, possui um afixo (morfema adjetivo) nulo que faz de tal palavra um caso de particípio adjetivo. Como a proposta de Lieber não esclarece a diferença entre particípios verbais e adjetivos por não haver uma marca morfológica clara que o faça, o autor aponta as soluções apresentadas por outras propostas baseadas no inglês e no

alemão (cf. MEDEIROS, 2008) a fim de que se chegue a uma estrutura e características que façam uma divisão clara do ponto de vista sintático e semântico.

Medeiros chega à divisão de Kratzer (2000) que divide as passivas estativas em estado alvo e estado resultante para explicar a diferença semântica. O estado alvo se refere ao estado decorrente de um evento mas que está sujeito a alteração, enquanto o estado resultante decorre do evento e é imutável. O exemplo dado é de Parsons (1990 *apud* MEDEIROS, 2008) em quem Kratzer se inspirou: se uma bola é atirada para cima de um telhado, ela estar em cima do telhado é um estado alvo, já que ela pode descer, enquanto o estado de ter sido atirada não pode deixar de existir, sendo por isso um estado resultante.

Esta divisão é adotada pelo autor para dar conta de diferenças semânticas entre passivas adjetivas e as aplica ao português brasileiro, propondo as seguintes estruturas para cada tipo de passiva adjetiva:

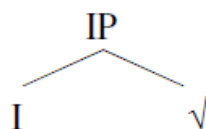


Figura 3 – Árvore adaptada de Medeiros (2008, p. 185) de participípios de estado alvo.

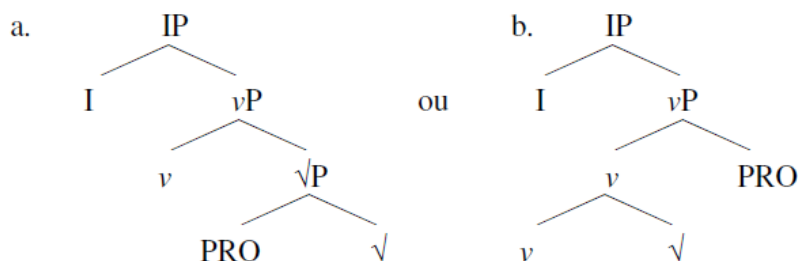


Figura 4 – Árvore adaptada de Medeiros (2008, p. 185) de participípios de estado resultante.

O caso da Fig. 3 expressa a estrutura sintática de passivas de estado alvo onde, segundo o autor, há a possibilidade de alteração não composicional (*i.e.*, idiossincrática) da raiz, já que o traço estativo é concatenado diretamente à raiz sob o terminal sintático “PRO”, que indica a presença de um argumento externo nulo. Já na Fig. 4, o nó onde o traço de estado



(correspondente ao particípio) é adicionado e se concatena a um nó terminal vezinho (nó que verbaliza uma raiz), o que possibilita que haja uma passiva adjetiva estativa gerada a partir de um evento. O traço de estado, então, é adicionado ao nó terminal PRO que, por sua vez, se concatena ao vezinho. Nota-se que em nenhuma destas estruturas há um núcleo de Voz o que, segundo Medeiros (2008), caracteriza os particípios adjetivos.

Essas considerações mostram o que para o autor representa as estruturas sintáticas do particípio adjetivo enquanto passiva adjetiva. Uma contribuição para este trabalho é a demonstração pelo autor de que há propriedades adjetivais no particípio que o distinguem de formas absolutamente verbais (ausência de núcleo de Voz), ainda que ele ocupe um IP, e não um AP.

No recorte deste experimento, no entanto, lidamos com particípios como em (7):

(7) O rapaz irritado chegou.

Casos como este não são analisados de forma detalhada por Medeiros (2008). Parece haver uma lacuna na literatura que analise palavras no particípio passado que ocorrem dentro de sintagmas nominais como modificadores de um substantivo, já que uma busca por respostas em outras leituras não foi muito frutífera (ainda que não completamente, como será visto). A discussão sobre o particípio adjetival recebe mais atenção quando, por exemplo, em contexto sentencial com verbos de cópula, como no caso abaixo:

(8) O rapaz está irritado.

Poder-se-ia argumentar que em (7) o que ocorre é uma oração reduzida, apagada da sentença fala através de algum mecanismo no curso da derivação sintática, de forma que (7) seria:

(7b) O rapaz [que está] irritado [por alguém] chegou.

Foltran e Crisóstimo (2005) apresentam um estudo sobre os adjetivos participiais em sintagmas nominais como (7) sem considerar orações implícitas como em (7b). As autoras mostram que, neste contexto sintagmático, os particípios exibem propriedades comuns a demais

adjetivos. O aprofundamento da discussão se dá na análise separada de verbos biargumentais e dos monoargumentais que dão origem aos adjetivos participiais.

As passivas adjetivas do tipo discutidas em Medeiros (2008) são tratadas por Foltran e Crisóstimo dentro de sintagmas nominais como aquelas que são biargumentais (*i.e.*, que selecionam dois argumentos). Os exemplos dados pela autora são:

- (9) a. A televisão transmitiu as notícias via satélite.  
b. As notícias *transmitidas* via satélite chocaram a população.  
(FOLTRAN, 2005, p. 133)

As autoras explicitam que esta classe de verbos é a única passível de ocorrer como passivas e que, por este motivo, tais participios podem ser usados como adjuntos em sintagmas nominais.

Há, no entanto, a classe dos monoargumentais, que selecionam apenas um argumento e que também são utilizados em sintagmas nominais como adjetivos participiais. A diferença reside na menor regularidade que esta classe de verbos exibe neste tocante, já que alguns verbos monoargumentais podem ser usados no participio como adjetivos em sintagmas nominais, enquanto outros não. O exemplo em (9) representa um caso em que há possibilidade de ocorrência (com verbo biargumental), enquanto o exemplo em (10) (como verbo monoargumental) expressa um caso em que não é permitido, segundo a autora:

- (10) a. Maria devolveu o envelope.  
b. O envelope *devolvido* estava rasgado. (FOLTRAN, 2005, p. 134)
- (11) a. O aluno bocejou na aula.  
b. \*O aluno *bocejado* saiu da aula.<sup>1</sup> (FOLTRAN, 2005, p. 134)

Sobre os verbos monoargumentais que não podem ocorrer como adjetivos, as autoras pontuam que há uma subdivisão dos verbos monoargumentais em inacusativos (aqueles cujo

---

<sup>1</sup> Indaga-se se o argumento da autora é completamente válido quanto à gramaticalidade de bocejado como adjetivo participial em SN, se levamos em consideração que talvez a agramaticalidade resida na forma como bocejado parece selecionar aluno como argumento interno, enquanto talvez uma construção como "o bocejo bocejado" possa ser plausível. O mesmo se aplica a outros exemplos de caso impossível de verbo monoargumental como adjetivo participial dados pela autora.

sujeito origina-se como argumento interno, como no *cair* no exemplo “as folhas caíram”) e intransitivos (não selecionam argumentos internos, seus sujeitos são agentes e possuem todas as características de sujeitos, como *bocejar* no exemplo 11a), apenas os primeiros sendo passíveis de aparecer em ambiente sintático como participios adjetivos. Foltran e Crisóstimo (2005) pontuam que o participio adjetivo costuma modificar seu argumento interno, não externo, motivo pelo qual frases como (11b) são agramaticais.

Algumas outras propriedades são apontadas por Foltran e Crisóstimo que aproximam os participios (que podem ocorrer em sintagmas nominais como adjuntos) de adjetivos. São elas: poderem ser modificados por modificadores de grau, como *muito*; terem uma forma superlativa absoluta sintética, como em “A perna está *inchadíssima*”; poderem aparecer em superlativas relativas, como em “As meninas são *as menos cansadas* da turma”; e, enfim, poderem ser coordenados, como em “Como essa empresa está confusa e *decaída*”.

Toda essa discussão permeia uma questão importante para este trabalho: se os participios podem atuar como verdadeiros adjetivos, e não apenas como passivas adjetivas ou pequenas orações ou orações reduzidas, então fala-se de uma mudança de classe de palavras que eram originalmente verbos. E se isto é verdade, então o participio passado, em certos contextos, pode ser considerado um caso de derivação, e não de flexão, já que dentre as propriedades de derivações citadas na seção anterior estão a alteração da classe da palavra, o que está condizente com a análise de Foltran e Crisóstimo (2005), e a alteração semântica da raiz, o que condiz com o que Medeiros (2008) aponta como passivas adjetivas de estado alvo. Essa distinção é importante, já que as previsões de algumas propostas de modelos de processamento lexical a serem discutidas nos próximos capítulos são distintas se uma palavra é originada de um processo derivacional ou se é gerada por um processo flexional. E o que queremos investigar aqui é, justamente, se o participio adjetival, que chamaremos genericamente de nominal, como fizeram outros autores, se comporta como um caso de derivação nos experimentos realizados e se há, de fato, uma separação no processamento observável nos dados comportamentais entre participio verbal e nominal e, conseqüentemente, entre derivação e flexão.

O próximo capítulo tratará dos principais modelos de processamento lexical. Também haverá menção às evidências na literatura que embasam tais propostas e quais variáveis afetam o processamento de palavras de morfologia complexa.

### 3 PROCESSAMENTO MORFOLÓGICO: MODELOS, SUAS EVIDÊNCIAS E VARIÁVEIS QUE O AFETAM

Os modelos de processamento lexical morfológico se referem a propostas que detalham as etapas envolvidas na conversão de um sinal físico ao acesso lexical (aqui entendido como junção da forma com o sentido da palavra), no caso da compreensão, ou o inverso no caso da produção. Estes modelos podem ser caracterizados como não-decomposicionais ou como decomposicionais, apenas estes últimos levando em consideração a existência de processamento de estrutura interna da palavra.

Destaca-se aqui a preferência de uso do termo "processamento lexical" para se referir a estudos acerca do tema. Em uma discussão terminológica, Taft (2001) problematiza como o termo "acesso lexical" passou a designar os estudos de processamento de palavras. Tal uso é problemático pois, como demonstrado pelo autor, o modelo teórico por trás de um dado estudo pode alterar o que se entende por acesso lexical. Por exemplo, em um estudo interessado na recuperação da forma fonológica de uma palavra, poder-se-ia determinar como acesso lexical o momento em que um sinal físico (neste caso, acústico) é convertido em fonemas e a palavra, reconhecida quanto a sua forma. No entanto, Taft (2001) cita casos como o de estudo de palavras homônimas em que se considera acesso lexical apenas o momento de pareamento entre forma fonológica e significado da entrada lexical correspondente a uma das formas homônimas. Zauner *et al.* (2014), por exemplo, definem como momento do "acesso lexical" o pareamento entre uma informação visual e uma forma ortográfica de uma palavra, o que antecede o acesso ao significado dela. Já França *et al.* (2008) consideram como acesso lexical, em se tratando de palavras de morfologia complexa, quando a raiz é concatenada com o primeiro morfema categorizador e o seu significado pareado. Tais definições, inclusive, afetam o que se define como processos pré ou pós lexicais. Por isso, adota-se aqui uma definição de acesso lexical que serve ao propósito deste trabalho, que é o resultado do curso do *processamento lexical* em que uma dada forma fonológica ou ortográfica é pareada com seu significado, enquanto utiliza-se a expressão "processamento lexical" como a que designa os estudos sobre as etapas e variáveis envolvidas no processamento de palavras.

A seguir, estão duas seções dedicadas a destrinchar os modelos relativos a cada uma destas caracterizações. Uma terceira seção tratará da relação entre modelos de léxico mental e

de representação de morfologia com os estudos de processamento. Ao final deste capítulo, uma conexão entre o tema desta dissertação e os modelos de processamento relevantes a ela será realizada.

### 3.1 Modelos não-decomposicionais

#### 3.1.1 Propostas conexionistas

As propostas conexionistas não necessariamente presumem a existência de uma morfologia na mente do falante ou, em outras palavras, de regras que as gerem. A morfologia é uma relação entre palavras que emerge do uso, da frequência e da semelhança de propriedades fonológicas.

O modelo de Bybee (1995), por exemplo, pressupõe a formação de redes que compartilham propriedades fonológicas e sintático-semânticas cujas produtividades aumentam conforme a frequência das formas derivadas/flexionadas. O autor contrapõe seu modelo ao de dupla rota de Pinker (1991), por exemplo, que prevê que palavras irregulares do passado no inglês seriam listadas na memória, enquanto as regulares seriam formadas por regras. Bybee ressalta que mesmo algumas palavras classificadas como "irregulares" possuem um padrão de alternância vocálica (ex.: *ring* e *rang*, *sing* e *sang*) que são percebidas por o que ele chama de "rede". O que determinaria a produtividade de certos tipos de morfologia seria a frequência de tipo, enquanto número de ocorrências de um padrão morfológico regular em uma língua (contrastado com a frequência de *token*, que é quantas vezes uma mesma palavra ocorre em uma língua), fator este que é central em seu modelo, segundo o autor, ao contrário de outros modelos que não o consideram. Apesar de elaborar este modelo para explicar a morfologia no léxico mental, Bybee também afirma que não há regras morfológicas na memória; ou seja, cada constituinte que forma uma palavra não possuiria uma entrada própria, já que o que de fato caracterizaria o nível morfológico é uma relação de sobreposição entre fonologia e semântica. Sua argumentação tem por base a discussão de simulações de modelos conexionistas sobre a língua hauça (língua falada na África oriental) e uma análise da produtividade de tipos de morfologia plural no alemão.

Uma outra proposta conexionista que vai mais adiante em retirar a significância da morfologia no léxico mental é a teoria de convergência de Seidenberg e Gonnerman (2000). Assim como Bybee, os autores associam a morfologia (regular, ou como chamam, *quasiregular*) à sobreposição de ortografia, fonologia e significado. Contudo, diferentemente

de Bybee, os autores negam mais abertamente tanto a existência de um nível morfológico quanto dos constituintes, sendo o que se chama de morfologia apenas o peso de conexões entre palavras que possuem semelhanças fonológicas e semânticas.

### 3.1.2 Naive discriminative reader

Uma outra proposta não-decomposicional é a abordagem de *naive discriminative learning* ("aprendizagem discriminativa ingênua"). Baayen *et al.* (2011) diferencia seu modelo aos conexionistas, ao acrescentar mais fatores à natureza da "morfologia".

Segundo Baayen *et al.* (2011), a aprendizagem de morfologia no reconhecimento visual ocorre por experiência de uso e análises probabilísticas. O surgimento de um nível morfológico ocorre pelas associações que surgem entre formas que se sobrepõem e o significado. O que Baayen *et al.* acrescentam é a afirmação de que toda a aprendizagem ocorre sem que haja parâmetros especificamente linguísticos, mas sim uma análise que tem por base parâmetros de aprendizagem e conhecimento gerais. Um outro argumento a favor de seu modelo, baseado no sucesso com sua modelagem computacional de paradigmas flexionais e derivacionais nas línguas servia e inglês, de acordo com os autores, é que ele possui poder explicativo sobre o funcionamento de morfologia em quaisquer línguas, independentemente da produtividade da morfologia nelas. Também conforme os autores, a leitura de palavras isoladas é possível por pistas contextuais e o efeito que prediz a facilidade desse processamento é a entropia relativa (*i.e.*, a diferença entre uma variante flexional enquanto um "exemplar" e a probabilidade de distribuição de protótipos dentro de um paradigma flexional).

Ressalta-se que este modelo decorre de modelagem computacional. O *naive discriminative reader* é uma modelagem computacional que utilizou dados inseridos de sérvio e de língua inglesa para analisar o poder do modelo de predizer efeitos observados em experimentos passados de tempo de resposta à tarefa de decisão lexical, no caso do sérvio, e de interações entre força de conexão lexical de significado e forma e paradigmas flexionais, no caso do inglês. Baayen *et al.* informam por fim que seu modelo é explicativo apenas para estágios iniciais de leitura de palavras e não possui previsões quanto a casos de palavras ambíguas cuja desambiguação requer contexto sentencial.

### 3.2 Modelos decompositivos

Foi mencionado que os modelos anteriores a este se tratavam de propostas não-decomposicionais, ou seja, que não preveem nem a existência de representações de morfemas na memória nem o papel destes no processamento de forma decomposicional. No entanto, eles falham em explicar os dados oriundos de experimentações em que efeitos especificamente morfológicos, formais, desassociados de fonologia/ortografia ou semântica são detectados. Esta subseção trata dos modelos surgidos de tais achados que, apesar de diferirem quanto aos estágios envolvidos no curso do processamento lexical morfológico, têm em comum a asserção de representações morfológicas no léxico mental.

### 3.2.1 Decomposição precoce automática

Este modelo se refere a uma proposta em que as palavras de morfologia complexa são decompostas em morfemas de forma automática no início do processamento. Trata-se de um modelo que prevê uma segmentação que antecede o acesso à semântica da palavra, com base em informações morfo-ortográficas (no caso de palavras lidas) ou morfo-fonológicas (no caso de palavras ouvidas). Os estágios são de (1) segmentação da palavra em morfemas, (2) acesso às informações semânticas e funcionais dos morfemas e (3) combinação dos constituintes.

Seus principais proponentes foram Taft e Forster (1975). Os autores reportaram três experimentos comparando o tempo de resposta em tarefa de decisão lexical de palavras e não-palavras. O recorte dos autores foi de palavras prefixadas e as raízes (reais ou pseudo) que se obtém delas.

O primeiro experimento comparou não-palavras que eram radicais reais de palavras prefixadas (ex.: *juvenate*/*juvenescer* de *rejuvenate*/*rejuvenescer*) com pseudo-radicais (ex.: *bel*/*belar* de *rebel*/*rebelar*) e mostrou que os radicais reais eram respondidos mais rapidamente na tarefa. Esclarece-se que *juvenate* não figura isoladamente como uma palavra no inglês; no entanto, há uma semântica transparente ao considerar-se uma semântica desse radical com o significado do prefixo (o ato de de *juvenate*/*juvenescer* de novo, como encontrado no dicionário segundo os autores), ao contrário do que acontece com *rebel*. A previsão é que se radicais como *juvenate* não existem como radicais no léxico mental, então o tempo de resposta a elas deveria ter sido maior, e não menor que não-palavras que foram classificadas como pseudoradicais, devido a uma semelhança daqueles com palavras reais do inglês, o que não ocorreu.

O segundo experimento comparou não-palavras radicais que são morfemas presos com palavras cuja ortografia é a mesma que o radical, mas que são morfemas livres (ex.: *vent*/*vir* de *prevent*/*prevenir*), que possui um significado como morfema livre e outro como preso. A

hipótese foi que a existência de uma entrada lexical associada ao morfema preso acarretaria uma demora maior de resposta à tarefa quando sua frequência é maior que a sua homógrafa que é uma palavra real. A previsão pôde ser então observada nos resultados do teste, estando correta. Uma outra comparação ocorreu com a condição em que o morfema livre é mais frequente que o preso (ex.: *card*/cartão vs. *discard*/descartar), o que não deveria acarretar diferenças no tempo de resposta comparado à morfemas presos ou livres que não possuem homógrafos, que foi o que ocorreu.

O terceiro e último experimento associou prefixos indevidos a não-palavras que eram ou radicais reais (ex.: *de* + *juvenate* = *dejuvenate*) ou pseudoradicais (ex.: *de* + *pertoire* = *depertoire*). Caso ocorra decomposição morfêmica, então a primeira condição demandaria maior tempo de resposta, já que houve segmentação de dois morfemas com entradas lexicais, enquanto o segundo caso teria latência menor pois o pseudoradical não possuiria uma entrada lexical própria, o que facilitaria a resposta à tarefa de decisão lexical. Conforme previsto, a primeira condição teve latência maior que a segunda.

O conjunto dos resultados deu suporte à proposta dos autores de que, quando lemos palavras de morfologia complexa - neste caso, derivadas por prefixação - há uma segmentação no início do processamento das mesmas. No entanto, um novo estudo foi realizado por Taft (2004) para averiguar a validade de seu modelo em decorrência do surgimento de propostas de dupla-rota (estas serão abordadas na próxima subseção), em que é possível que uma palavra de morfologia complexa não passe por decomposição em morfemas caso o efeito de frequência superficial (*i.e.*, da palavra inteira) seja maior que a frequência da base (*i.e.*, frequência do radical).

Neste estudo, Taft mostrou que mesmo quando há uma aparente ausência de efeito de frequência de base (o que contradiz um modelo decomposicional precoce automático) é possível que a explicação esteja em uma atribuição de peso diferenciado do participante ao estágio de combinação de morfemas. Isso ele concluiu ao mostrar que o efeito facilitador de alta frequência da base é revertido quando a frequência de base é relativamente maior que a frequência da forma de superfície. Isso se explicaria se em estágios posteriores de composição morfológica, a seleção da palavra candidata é relativamente mais custosa, já que o reconhecimento dependeria de um mapeamento para uma forma de superfície relativamente menos frequente. De acordo com o autor, o efeito reverso de frequência de base observado no seu estudo seria mais bem explicado por um modelo decomposicional, assim, provendo uma explicação alternativa ao modelo de dupla-rota, que pressupõe que uma ausência de efeito de



base indica processamento por forma inteira da palavra, mas Taft (2004) também indica que pode haver diferenças de língua para língua quando à via de processamento lexical, que favorece decomposição automática no holandês e dupla rota no inglês (Taft cita resultados de um outro estudo em inglês para chegar a esta afirmação, que se contrapõem aos dele).

Posteriormente, Stockall e Marantz (2006) estenderam a proposta de decomposição automática para palavras em inglês irregulares, que Pinker (1991) associou ao processamento por forma inteira, conforme será explicado na próxima subseção. Segundo os próprios autores, o modelo de Taft e Forster não se contrapõe àquele que apresentaram, mas Taft e Forster não abordaram a questão das palavras irregulares no passado, que foi alvo de muito debate em modelos de dupla rota.

### 3.2.2 Modelos de dupla rota de Pinker (1991) e Schreuder e Baayen (1995)

Pinker (1991) propôs a existência de duas formas de processamento lexical com base nos casos do passado simples do inglês. O autor argumenta que palavras regulares, cuja flexão ocorre pela adição de um afixo à raiz (ex.: *walk* + *ed* = *walked*), são armazenadas na memória em morfemas que são recuperados e combinados no curso do processamento, enquanto palavras irregulares, cuja flexão no passado não segue uma regra (ex.: *ran*, passado do verbo *run*), são armazenadas na memória e recuperadas por um mecanismo associativo.

Uma implementação neurológica do modelo de Pinker (1991) é proposta por Ullman *et al.* (1997) em uma comparação entre pessoas com afasia posterior (i.e., em porção posterior do córtex cerebral) ou doença de Alzheimer com pessoas com afasia anterior (i.e., em porção anterior do córtex) ou mal de Parkinson. Como explicado pelos autores, o primeiro grupo de pessoas tinha a capacidade preservada de gerar palavras em inglês no passado com *-ed*, palavras estas que exigem a aplicação de um sistema de regras, o que foi associado à preservação da porção frontal do córtex e da gânglia basal. Já o segundo grupo, com afasia anterior e mal de Parkinson, mostrou uma capacidade preservada de utilizar palavras consideradas irregulares no inglês (como no caso do par *dig/cavar-dug/cavou*) o que foi associado a uma região de memória declarativa, onde as palavras estariam listadas na memória, na porção temporal-parietal/medial-temporal do córtex. Assim, o modelo dual se apoiaria na memória procedural para a geração de palavras a partir de regras, na região anterior do córtex (lobo frontal e gânglia basal), enquanto a listagem de palavras em forma inteira na memória (neste caso, palavras irregulares no passado em inglês) seria alocada na memória declarativa, a qual, por sua vez, é associada à porção posterior do córtex (região temporal-parietal/medial-temporal). No entanto, este modelo é

contestado por achados que mostram facilitação por relação morfológica mesmo para palavras regulares, o que só deve ser possível se as palavras irregulares têm acesso a uma representação interna de morfemas (cf. MORRIS; STOCKALL, 2012 para uma contestação em inglês; SMOLKA *et al.*, 2007 para uma contestação em alemão).

Em contraposição, Schreuder e Baayen (1995) propõem um meta-modelo que dê conta das computações necessárias para o processamento morfológico em diferentes línguas. O modelo que apontaram como o melhor se refere a um modelo de dupla-rota em que desde o início do processamento, tanto a rota de processamento por forma inteira quanto por decomposição em morfemas são ativadas concomitantemente para acessar o significado de uma palavra de morfologia complexa. O que leva ao melhor desempenho de uma ou outra rota é a frequência superficial da palavra, não especificamente sua regularidade.

Logo percebe-se a diferença desta para a proposta de Pinker. A previsão é que mesmo palavras regulares cuja frequência superficial é alta também seriam processadas pela forma inteira, o que retira da regularidade do processo morfológico o critério para a via de processamento a ser tomada. Baayen, Dijkstra e Schreuder (1997) atestam sua previsão em um estudo comportamental que demonstrou que substantivos regulares no plural em holandês são processados por forma inteira (o tempo de resposta de plural chegando a ser igual ao da sua forma singular monomorfêmica, contrastado com formas regulares de baixa frequência superficial que exigiam maior tempo de processamento), o que contraria a previsão do modelo de Pinker (1991).

Schreuder e Baayen também comparam sua proposta com outros modelos. Mais especificamente, são mencionados os modelos de dupla-rota paralela, *augmented morphology model* e *morphological race model*. A discussão perpetrada pelos autores se embasa em justificativas hipotéticas que favorecem o seu modelo como mais viável e apontam que uma vantagem do seu modelo sob os demais é o fato de que há uma retroalimentação (*i.e.*, os potenciais níveis hierárquicos se comunicam uns com os outros, permitindo inclusive um retorno dessa informação) no curso do processamento lexical e comunicação entre as duas vias de processamento, o que facilitaria o acesso lexical quando comparado com modelos em que as duas vias se ativam sem uma comunicação entre si.

#### 3.2.4 Decomposição tardia: modelo supraleixal

Girardo e Grainger (2000) apresentaram um modelo em que o processamento da palavra inteira ocorre, inicialmente, pela forma inteira. Após a forma inteira, a decomposição em

morfemas se torna possível, mas apenas em palavras de semântica transparente (quando a combinação da raiz com os afixos resulta no significado da palavra que é previsível pelos morfemas). Por isso, o modelo é chamado de decomposição tardia ou, ainda, de supralexical, já que as informações morfêmicas estariam disponíveis em um nível acima do nível do item lexical.

A ideia por trás do modelo é que a existência de uma unidade morfêmica decorre de todas as palavras morfologicamente relacionadas que a possuem. A raiz estaria, então, hierarquicamente acima das formas inteiras das palavras, e por isso seu acesso ocorreria apenas posteriormente ao acesso à palavra inteira. Para comprovar sua hipótese, Giraudo e Grainger usaram a técnica de *priming* encoberto morfológico para exibir *primes* morfologicamente relacionados com a palavra alvo cujas frequências superficiais foram controladas, sendo alta ou baixa. Se a hipótese dos autores estivesse correta, apenas a condição em que o *prime* é de alta frequência deveria ser capaz de facilitar o processamento do alvo, já que o acesso rápido à raiz (posterior ao acesso à palavra inteira) em um paradigma no qual o *prime* é exibido tão rapidamente só poderia ser alcançado pelas palavras de alta frequência de ocorrência, cuja velocidade de processamento é maior.

O experimento, que utilizou a variável dependente de tempo de resposta à tarefa, obteve o efeito previsto pelos autores. Giraudo e Grainger argumentam que isto se contrapõe tanto a modelos de decomposição precoce automática (TAFT, 2004) como modelos de dupla rota de forma geral, pois: (1) não houve facilitação detectada quando a frequência superficial era baixa, o que contraria o modelo de Taft e; (2) a previsão dos modelos de dupla rota, segundo os autores, aponta que as palavras alvo de alta frequência tenderiam a não obter vantagem de relação morfológica com os *primes*, pois utilizariam a via de forma inteira da palavra, o que não foi o caso no estudo de Giraudo e Grainger.

### 3.3 Processamento de morfologia e modelos de representação

No capítulo anterior debateu-se os conceitos de mente e cérebro e a implicação para o que se define como representação linguística e processamento da linguagem. Embora a natureza da relação entre um grupo de conexões neuronais e um dado conhecimento ou capacidade esteja longe de estar esclarecida, algumas considerações devem ser feitas sobre a divisão entre processamento e representação.

Tome-se como exemplo inicial a Morfologia Distribuída (HALLE; MARANTZ, 1992). Segundo a MD, as computações de morfologia não ocupam um nível a parte da sintaxe, sendo

também sintáticas. Para a MD, as computações que estão envolvidas em uma palavra como "descriminalização" perpassam a inserção de cada morfema e raiz em seus respectivos nós. Apesar de não ser necessariamente uma teoria de processamento, pode-se concluir a partir desta teoria que palavras de morfologia complexa - ao menos, as que possuem semântica transparente - são sempre computadas a partir de seus morfemas, já que estes estão listados separadamente na memória. Mas se a MD está correta, como explicar os resultados de modelos de dupla rota como de Schreuder e Baayen (1995), por exemplo, que encontram efeitos de processamento por forma inteira da palavra?

Olhe-se por outro lado. Aronoff (2008) volta a defender seu posicionamento lexicalista, segundo o qual as palavras de morfologia complexa estão armazenadas na memória em sua forma inteira. Ele o faz através de uma análise teórica de duas línguas, uma em que, segundo o autor, não há processos morfológicos (língua de sinais beduína de Al-Sayid) e uma em que há uma rica morfologia, mas que é absolutamente formal (língua semítica), ou seja, sem relação com significado. No entanto, se Aronoff está correto, como explicar os efeitos encontrados em testes tanto anteriores (ex.: RASTLE; DAVIS; NEW, 2004; PYLKKANEN *et al.*, 2004) quanto posteriores (STOCKALL *et al.*, 2018; PYLKKANEN *et al.*, 2019.) à publicação de seu texto? Como pode haver processamento morfológico sem que haja representação linguística de estrutura interna das palavras?

As perguntas feitas carecem de uma resposta, e não é objetivo deste trabalho respondê-las aqui. O que se intenciona é apontar para o fato de que há um descompasso entre estudos fundamentalmente teóricos e aqueles que utilizam de dados comportamentais ou neurofisiológicos. Os primeiros conseguem lidar com um grau de granularidade largamente maior que os estudos baseados nos últimos, enquanto estes testam e propõem modelos tanto de processamento que se aproximem da realidade neurobiológica que os sustentam.

O foco adotado aqui é nas contribuições que a psicolinguística e a Neurociência da Linguagem proveem nos estudos de processamento morfológico através de seus achados e seus modelos. Já é uma tarefa herculeana conciliar os achados contrastantes de estudos psiconeurolinguísticos, ainda conciliar distintas teorias com estes achados também contrastantes escapa ao escopo desta dissertação. Assim, a opção por partir de achados da psiconeurolinguística é de ordem pragmática, e não por uma atribuição de maior ou menor validade a uma ou outra área. Ademais, a opção por tratar de modelos psiconeurolinguísticos com inspirações na teoria linguística representa uma tentativa, ainda assim, de uma

aproximação entre os saberes, cuja divisão se justifica epistemologicamente, mas que é importante buscar conciliar para uma visão mais completa do fenômeno da linguagem.

### 3.4 O que dizem os achados da literatura sobre os modelos?

Os estudos são diversos em mais de um sentido. Por um lado, há resultados oriundos tanto da psicolinguística, em testes *offline* e *online*, quanto resultados de técnicas de neuroimagem, que advém de técnicas como eletroencefalograma (EEG), magnetoencefalograma (MEG) e ressonância magnética funcional (fMRI, *functional magnetic resonance imaging*). Por outro, as evidências criam mais perguntas que respostas ao apontarem para direções opostas, favorecendo tanto modelo de decomposição precoce automática quanto de dupla rota.

Focar-se-á nas contribuições advindas de técnicas *online* e *offline* da psicolinguística e dos estudos de EEG e MEG. Apesar de o fMRI ser uma excelente técnica de mapeamento de ativação cortical por processos cognitivos, as demais técnicas de estudo permitem uma análise dos estágios envolvidos no curso temporal do processamento lexical, motivo pelo qual tiveram preferência de enfoque.

#### 3.4.1 Psicolinguística

Áreas correlatas, a psicolinguística e a neurociência da linguagem compartilham o objeto de estudo. O que as divide, na realidade são o tipo de método e técnica aplicados nos estudos sobre a linguagem e o escopo da investigação, já que, quanto a este último, a neurolinguística tem algumas preocupações adicionais possibilitadas por suas técnicas.

A psicolinguística lança mão de estudos de duas naturezas, *online* e *off-line*. As técnicas *online* são aquelas que captam o processamento enquanto ele ocorre, com dados diferentes. Já as técnicas *off-line* envolvem dados que são resultados do processamento; ou seja, representam o produto dos processos cognitivos.

Maia (2015) menciona duas técnicas *on-line* muito recorrentes nos estudos de processamento de linguagem: leitura automonitorada e rastreamento ocular. A leitura automonitorada, técnica que foi aplicada no Experimento 2 desta pesquisa e que será retomada no capítulo 5, se refere ao pressionamento de teclas que fazem surgir palavras ou blocos de palavras em uma tela. Desse modo, é possível aferir através do tempo de leitura de cada segmento questões envolvidas no processamento de sentenças, por exemplo, já que o programa

usado para esta técnica grava o intervalo entre a exibição da palavra/bloco de palavras e o momento em que se pressiona a tecla para que o segmento seguinte seja exibido. Já o rastreamento ocular se refere ao acompanhamento dos movimentos sacádicos dos olhos, a quantidade de fixações de olhar em partes da tela do computador e o tempo de duração das fixações. Esta técnica requer um aparelho próprio, chamado de rastreador ocular (ou *eye tracker*). O rastreamento ocular oferece diversos benefícios, já que é possível realizar uma correlação maior entre os processos realizados no cérebro e os movimentos dos olhos necessários para a captação das informações visuais requisitadas por tal processo. Esta técnica é aplicada no experimento de Justino e Mota (2019), cujo resultado é relevante para este trabalho.

As técnicas *off-line* podem ser aplicadas em conjunto com as *online* como dados que complementam as análises, ou mesmo como os únicos dados almejados para a análise. Relevante para esta pesquisa, já que foi utilizada no Experimento 1, a tarefa de decisão lexical é uma técnica *offline* que consiste de um teste em que o participante deve decidir se uma dada sequência de segmentos ortográficos/fonéticos é uma palavra na língua alvo ou não (GARCIA, 2015). Obtém-se dois dados analisáveis em tarefas dessa natureza: a acurácia, referente à porcentagem de acertos que um participante obteve nas tarefas; e o tempo de resposta à tarefa, em que se registra o tempo que um participante levou para finalizar a tarefa, medido normalmente a partir do momento em que um estímulo é exibido. É possível que os dados de acurácia das tarefas não sejam utilizados na análise e tenham função somente atencional, para que se mantenha a atenção dos participantes ou mesmo para que se detecte quais participantes não deram a devida atenção ao experimento, e por isso seus resultados sejam removidos da análise, aplicação estas que fizeram parte do Experimento 2, reportado no capítulo 6.

Há paradigmas que podem ser usados nos experimentos e que contribuem para um refinamento metodológico dos estudos. Nos estudos de processamento lexical, um paradigma frequentemente encontrado é o de *priming*, ou pré-ativação. Como explica Garcia (2015), o *priming* é um estímulo exibido anteriormente ao estímulo alvo, e há alguma relação entre o par que deve (1) provocar uma facilitação do processamento do alvo, (2) provocar uma inibição (*i.e.*, “dificultação”) do processamento do alvo ou (3) não exibir qualquer efeito sobre o alvo. O *priming* deve pré-ativar alguma informação que é compartilhada com o alvo e que se espera observar se provoca algum dos efeitos mencionados anteriormente, que pode ser de natureza formal (no sentido de forma ortográfica/fonético-fonológica), semântica ou mesmo morfológica (RASTLE; DAVIS, 2003). Quando a exibição do *prime* (o estímulo anterior ao

alvo) ocorre visualmente em tempo curto que não permite sua percepção consciente, fala-se de um paradigma de *priming* encoberto, associado a ativação de apenas etapas iniciais do reconhecimento de palavras (RASTLE; DAVIS; NEW, 2004; RASTLE; DAVIS, 2008) e encontrado em testes sobre processamento de morfologia. Estes paradigmas são adicionados aos experimentos para que se investigue o impacto deles nos dados que serão analisados, oriundos das técnicas mencionadas anteriormente.

De especial relevância para este experimento, o estudo seminal de Rastle, Davis e New (2004) utilizou um paradigma altamente produtivo nas investigações acerca de processamento de morfologia: o *priming* encoberto (ou *priming* mascarado, do inglês *masked priming*). O paradigma de *priming* encoberto, brevemente citado anteriormente, consiste na exibição de uma máscara, que pode ser de *hashtags*, após o qual o *prime* é exibido por períodos inferiores a 60ms. Com isso, o *prime* não pode ser percebido conscientemente. Os *hashtags* tem o objetivo de impedir que o *prime* tenha influência no processamento do estímulo alvo por traços deixados na memória episódica pela exibição visual do *prime*, motivo pelo qual a máscara de *hashtags* deve ser exibida por 500ms, seguida do *prime* por tempo extremamente curto<sup>2</sup> (cf. FORSTER; DAVIS, 1984). Ademais, a compressão do *prime* entre a máscara e a palavra alvo contribui para tornar a percepção do *prime* inconsciente, o que faz com que o alvo também atue como uma máscara. O *prime* também é exibido em letras minúsculas enquanto o alvo é exibido em letras maiúsculas para evitar um efeito por semelhança física. A expectativa dos autores é que apenas estágios iniciais envolvidos no processamento do *prime* sejam ativados, dado o período rápido de exibição (cf. RASTLE; DAVIS, 2008). No caso de palavras morfologicamente complexas, a facilitação esperada ocorreria pois a raiz compartilhada pelo *prime* e pelo alvo já teria sido ativada anteriormente à exibição do alvo, ou por uma semelhança ortográfica ou semântica entre os itens do par.

A tarefa utilizada por Rastle, Davis e New (2004) foi de decisão lexical. Com isso, os autores mostraram que a exibição de um *prime* encoberto que compartilha informações morfológicas com o alvo, seja de forma transparente ou opaca (*i.e.*, palavra que possui uma raiz da língua e um afixo da língua, mas cuja combinação não resulte no significado da palavra inteira, como acontece com *brother*/irmão, que possui uma “raiz” *broth*/caldo e um sufixo nominalizador “-er”), facilita o processamento do alvo, quando a condição em que os pares compartilham apenas forma ortográfica não resulta em facilitação significativa, o que reforça a

---

<sup>2</sup> Rastle e Davis (2003) apontam outras formas de apresentar a máscara, mas que podem gerar interferências adicionais no tempo de reação às palavras exibidas visualmente.

lógica de que o *priming* encoberto possui relação significativamente mais forte com efeitos de morfo-ortografia. A facilitação ocorrida nas condições transparente e opaca, segundo os autores, evidencia o modelo de processamento morfológico de Taft e Forster (1975) já que a facilitação ocorreu independentemente de haver processamento da semântica das palavras, o que deveria inibir, e não facilitar o processamento de palavras de semântica opaca caso a semântica fizesse parte do processamento do *prime* encoberto. Ainda, isso demonstrou que há uma segmentação das palavras de morfologia complexa que é automática e precoce. Efeitos semelhantes do paradigma de *priming* encoberto também foram encontrados em estudos comportamentais como de Kazanina *et al.* (2008), McCormick, Brysbaert e Rastle (2009) e Beyersmann, Castles e Coltheart (2013) e de EEG de Morris e Stockall (2012), por exemplo.

O estudo de Justino e Mota (2019) de rastreamento ocular com leitura automonitorada também tem diálogo direto com o recorte do presente trabalho. Os autores investigaram o processamento de palavras flexionadas nas três conjugações verbais do português: a primeira, terminada em -ar; a segunda, terminada em -er; e a terceira, que termina em -ir. As autoras verificaram a hipótese de que o tipo de conjugação e o tempo verbal conjugados são efeitos preditores do processamento lexical. Justino e Mota apontaram que a primeira conjugação (-ar) representa a classe de verbos com maior regularidade, ao contrário das demais, e pode ser, por isso, considerada *default*. Também argumentam que alguns tempos verbais, como o pretérito perfeito, são mais previsíveis que outros, como o presente do indicativo. Dessa forma, elas investigam qual modelo representa melhor o processamento lexical no português tendo por base como as variáveis de classe do verbo e tempo verbal podem afetar o processamento: se os modelos de mecanismo dual, como o de Pinker (1991), ou o mecanismo único de decomposição total, de Stockall e Marantz (2006). Segundo as autoras, como a classe “-ar” é a mais regular, em uma proposta de mecanismo dual, esta seria computada por um sistema de regras e segmentada, enquanto as classes “-er/-ir” seriam listadas na memória declarativa. Quanto ao tempo verbal, as autoras testaram o Pretérito Imperfeito do Indicativo, o Futuro do Presente do Indicativo (estes os mais imprevisíveis) e o Pretérito Imperfeito do Subjuntivo (o mais previsível). As medidas de primeira fixação, primeira passada de olhar e tempo de fixação confirmaram as previsões das autoras, com diferenças com significância estatística em que: primeiramente, a classe -AR refletia custo de processamento menor que as classes -ER/-IR, ainda que a vantagem de -ER sobre -IR não tenha sido explicitada; e em segundo lugar, o Pretérito Imperfeito do Subjuntivo apresentou maior facilidade de processamento que os tempo Futuro do Presente do Indicativo e Pretérito Imperfeito do Indicativo, ainda que uma gradação



entre estes dois últimos não tenha ficado clara. Porém, uma ressalva feita pelas autoras impediu que a variável tempo verbal pudesse ser determinada neste experimento como preditora das variáveis dependentes: as sentenças alvo, mantidas semelhantes nos três tempos verbais, tinham uma pequena diferença quando o verbo estava no Pretérito Imperfeito do Subjuntivo, que exigia a ocorrência da expressão "pediu que" antes do segmento crítico (o verbo conjugado no tempo investigado), enquanto a expressão que precedia o segmento críticos nos outros dois tempos verbais era "disse que". Este fato fez com que as autoras atribuíssem o processamento mais facilitado neste tempo verbal a uma restrição imposta pelo contexto sintático, e não ao tempo verbal em si. Ainda assim, tendo-se em conta os resultados que comparam as classes de conjugação, as autoras argumentam que os resultados condizem com uma proposta de mecanismo dual, que prevê a tentativa automática e inicial de aplicação de regras como bem sucedida apenas na classe mais produtiva e que explica as medidas de tempo de fixação, primeira fixação e primeira passagem menores para a classe de primeira conjugação.

O achado de Justino e Mota (2019) é importante pois esta pesquisa investiga o processamento de palavras no particípio passado de palavras regulares e da primeira conjugação. Se os resultados das autoras estiverem, de fato, em linha com uma previsão de que a primeira conjugação é *default* e, logo, sempre segmentada, espera-se um reflexo disso nos experimentos aqui realizados. A opção, inclusive, de trabalhar apenas com a classe de conjugação mais regular foi para colocar a teste de fogo a hipótese de que casos de particípio preferencialmente nominal (esta classificação será explicitada no próximo capítulo) podem ser processados por forma inteira a depender da variável frequência superficial, já que há estudos que sugerem que palavras derivadas de alta frequência são listadas na memória em sua forma inteira. Para se certificar de que o *prime* é segmentado, a fim de que se obtenha o efeito pretendido de *prime* morfológico, optamos por utilizar a mesma raiz da palavra alvo, mas conjugada no pretérito perfeito do indicativo, já que como constitui um caso indubitável de flexão (ao contrário do particípio que, como explicado, pode ser considerado nominal) deve ser sempre segmentado, se levamos em consideração a proposta de Leminen *et al.* (2013). Dessa forma, qualquer resultado comportamental que sugira armazenamento em forma inteira da palavra não poderia ser atribuído à maior ou menor regularidade e produtividade da classe de conjugação das palavras escolhidas.

Quanto a outros achados que evidenciam a ocorrência de modelos de processamento morfológico, sobre o de decomposição automática precoce, estão os estudos *offline* de Marslen-Wilson *et al.* (1994), Taft (2004) e Rastle, Davis e New (2004), Silva e Clahsen (2008),

Kazanina *et al.* (2008), McCormick, Brysbaert e Rastle (2009) e Beyersmann, Castles e Coltheart (2013), entre outros. Para além destes, a favor dos modelos de dupla rota, pode-se mencionar estudos de natureza *online*, de rastreamento ocular, de Maia, Lemle e França (2007), Garcia *et al.* (2009) e Medeiros *et al.* (2014) e os estudos comportamentais de Estivalet (2020) e Savinova e Malyutina (2021).

#### 3.4.2 Evidências de estudos com EEG

Os estudos da Neurociência da Linguagem, diferentemente da psicolinguística, registram atividade cortical concomitantemente a dados de tarefas, o que confere uma vantagem em termos de determinação do “quando” e do “onde” os processos cognitivos - e estágios a eles relacionados - ocorrem. Estes estudos utilizam principalmente três tecnologias: o EEG, o MEG e o fMRI.

O eletroencefalograma (EEG) é uma técnica que detecta e registra a atividade elétrica do córtex cerebral (RODDEN; STEMMER, 2008). Os neurônios se conectam via processos neuroquímicos que causam diferenças de carga elétrica dentro e fora do corpo neuronal. Isso causa uma passagem de sinais elétricos entre neurônios (sinapse), dos quais o EEG registra a oscilação dos potenciais elétricos que ocorrem no entorno dos corpos neuronais, resultante das sinapses, o que se chama “atividade pós-sináptica”. Como os sinais elétricos são pequenos, um amplificador é requerido. E como o cérebro está em constante atividade e, logo, emitindo sinais que não necessariamente estão relacionados ao teste, experimentos de EEG envolvem a exibição de número elevado de estímulos para que se extraia uma média geral. Além do mais, o registro de atividade é *time-locked*; ou seja, o registro do exato momento em que o estímulo é exibido fica gravado junto ao registro da atividade elétrica. O nome da técnica utilizada para o registro de picos de sinais elétricos que possam ser associados a processos cognitivos relativos a eventos de apresentação de estímulos sincronizados se chama extração de ERPs (*event related brain potentials*, potenciais cerebrais relacionados a eventos). Enfim, o EEG é reconhecido por sua alta precisão temporal, em milissegundos, ainda que não tenha precisão espacial, já que o escalpo distorce as fontes da atividade elétrica.

Cada ERP se caracteriza por um pico de atividade elétrica registrado pelos eletrodos. Os ERPs possuem uma de duas polaridades, negativa ou positiva. Quando respaldado pela literatura, os ERPs, ou componentes, recebem nomenclaturas que podem estar relacionadas à sua polaridade e ao momento aproximado em que ocorre após o início do registro dos dados.

Um componente importante para esta pesquisa é o MMN (*mismatch negativity*, negatividade de divergência). Ele ocorre quando uma sequência de estímulos sensoriais (acústicos ou visuais) repetidos – chamados de estímulo padrão – é interrompida por um sinal acústico que escapa a este padrão por um pequeno desvio, geralmente sensorial – o chamado estímulo desviante (ex. [ba] [ba] [ba] [pa]). O MMN é detectável cerca de 170ms após o momento em que o estímulo desviante se diferencia do estímulo padrão. A característica que o torna muito interessante é o seu caráter pré-atencional: mesmo que o participante realize uma tarefa (assistir a um filme mudo ou jogar um videogame) seu cérebro registra os estímulos acústicos que estão sendo emitidos. Isso faz com que o MMN potencialmente reflita processos cognitivos automatizados e inconscientes, ou pré-atencionais. Shtyrov e Pulvermuller (2006) associam o MMN, um componente que reflete desencontro de informações sensoriais, a uma sensibilidade à lexicalidade (ou seja, se uma sequência de fonemas é uma palavra ou não), ao processamento de informações semânticas de palavras e à análise sintática, em um caso específico de concordância de número (“*we come/nós vimos*” comparado com “*we comes\*/nós vem\**”).

O experimento de Leminen *et al.* (2013) em finlandês, que utilizou o componente MMN, deu origem a uma das questões que norteiam esta pesquisa. Leminen *et al.* compararam palavras em finlandês cujos afixos eram fonologicamente iguais (o afixo “-ja”), mas que quando associado a algumas raízes, seria um afixo flexional de plural (ex.: *laulu/canção, laulu-ja/canções*), e quando associado a outras (ex.: *laula/cantar, laula-ja/cantor*), seria um sufixo derivacional. A hipótese dos autores foi que palavras derivadas tenderiam a uma armazenagem em forma inteira quando fossem de alta frequência, já que a derivação é pouco produtiva, enquanto os casos de flexão seriam sempre representados em constituintes já que as flexões são muitas, e seria custoso em termos de memória se todas as palavras flexionadas de alta frequência fossem armazenadas em forma inteira. O efeito previsto foi obtido: palavras flexionadas de alta e baixa frequência não tiveram diferenças significantes no MMN, enquanto palavras derivadas de alta frequência tiveram pico de MMN maior que as de baixa frequência (ver Figura 5), o que os autores tomam como evidência de alto nível de ativação, um reflexo da sua pronta disposição a ser ativada.

Um caso parecido com o do finlandês, ainda que não seja semelhante exatamente, é o afixo “-ado/-ada”, da morfologia do participio passado. Apesar de não haver uma distinção entre a quais raízes pode se combinar quando ocorre como flexão verbal ou como afixo derivacional (ex.: “roubado”/“roubada”), esta morfologia produz tanto verbos quanto palavras adjetivas. O

achado de Leminen *et al.* (2013) é importante pois trata de uma distinção que não possui tanta clareza na literatura, que é a (potencial) diferença entre processos de formação da palavra no curso do processamento lexical. Ademais, como o MMN é pré-atencional, isso faz com que o resultado do estudo de Leminen *et al.* esteja menos sujeito a ser criticado por uma suposta influência da atenção, comportamento ou em decorrência desenvolvimento de uma estratégia pelos participantes, o que pode interferir tanto na gravação de dados neurofisiológicos de componentes que presumem atenção do participante quanto em dados *offline* de tarefas, como de decisão lexical.

A pergunta que surgiu foi o que aconteceria no PB em uma morfologia como o particípio passado, que provoca mudança de classe gramatical e possui concordância de gênero e de número (características da derivação) ou flexiona verbos (característica da flexão). Se determinadas palavras tivessem uma tendência a ocorrer nominalmente ou verbalmente e fossem controladas quanto à frequência, haveria uma replicação de Leminen *et al.* (2013)? Haveria facilitação por *priming* encoberto em todas as condições, o que evidenciaria o modelo de Taft e Forster (1975), ou os casos em que o particípio é derivacional e a palavra, de alta frequência, teriam ausência de efeito de facilitação, sugerindo um modelo de dupla rota, como o proposto no estudo de Leminen *et al.* (2013)? O peculiar do particípio passado é que, diferentemente do "-ja" do finlandês, ele se afixa às mesmas raízes, o que coloca ele mais ao centro de um *continuum*<sup>3</sup> entre derivação e flexão, o que instigou a formação desta pesquisa. Ao final deste capítulo, há um aprofundamento acerca do particípio passado.

---

<sup>3</sup> Conferir Bybee (1985) para uma proposta de léxico mental que determina a morfologia como um continuum gradado entre derivação e flexão.

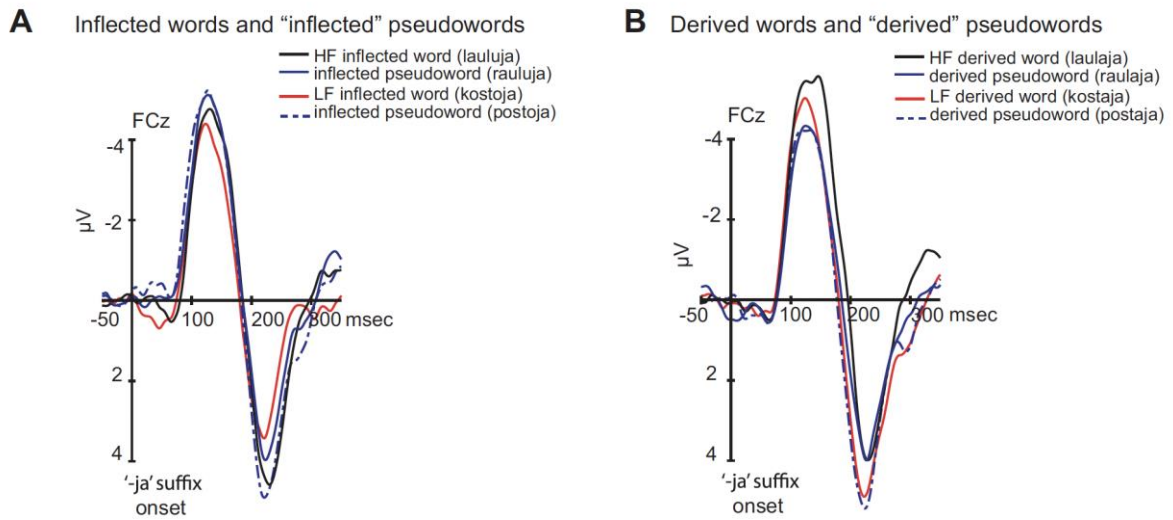


Figura 5 – Figura adaptada de Leminen *et al.* (2013). O pico negativo plotado para cima – denominado MMN – não é significativamente diferente na comparação entre palavras flexionadas de alta e baixa frequência (gráfico A), ao contrário da distinção de pico entre palavras derivadas de alta e baixa frequência, em que o pico da palavra derivada de alta frequência é significativamente maior que o de baixa frequência.

Sobre os modelos de processamento morfológico de forma geral, os achados obtidos através de experimentos com EEG também divergem quanto aos modelos que evidenciam. O modelo de decomposição inicial está evidenciado, por exemplo, em estudos de Smolka *et al.* (2007), França *et al.* (2008), Lavric, Rastle e Clapp (2010), Morris e Stockall (2012) e Coch, Bares e Landers (2012). Já dentre os que apontam para um modelo de dupla rota, estão os estudos de Münte *et al.* (1999), Leminen, Leminen e Krause (2010), Leminen *et al.* (2013) e MacGregor e Shtyrov (2013).

### 3.4.3 Evidências de estudos com MEG

Uma outra técnica produtiva nos estudos de neurociência da linguagem é o magnetoencefalograma, ou MEG. Esta técnica registra o campo magnético gerado pela atividade elétrica no córtex cerebral (GARCIA, 2013). Ao contrário do EEG, o MEG tem precisão espacial além da alta precisão temporal, ainda que a precisão espacial do MEG seja em centímetros, enquanto a técnica hemodinâmica de ressonância magnética funcional é precisa em milímetros. Da mesma forma que o EEG extrai ERPs, o MEG extrai ERFs (*event related magnetic fields*, campos magnéticos relacionados a eventos). O que torna o EEG mais atrativo

é o fato de que seu preço é significativamente menor que o do MEG, que custa milhões, o que torna o EEG mais acessível.

Da mesma forma que o EEG, há nomenclaturas que são atribuídas aos componentes de MEG. Dois exemplos relevantes para os estudos de processamento lexical são o M170 e o M350. Enquanto o “M” está relacionado ao tipo de dado neurofisiológico, que neste caso é magnético, o número adjacente se refere ao marco temporal aproximado de ocorrência dos componentes. O M170 mais recentemente foi associado à decomposição em morfemas de palavras de morfologia complexa, refletindo desta forma processamento morfológico (ZWEIG; PYLKKANEN, 2009), enquanto o M350 reflete o estágio de acesso lexical (PYLKKANEN *et al.* 2002; PYLKKANEN *et al.*, 2004; FIORENTINO; POEPEL, 2007).

Há evidências a favor de um modelo de decomposição automática e de dupla rota também no MEG. Referente ao modelo de decomposição automática, há evidências abundantes em estudos como de Pylkkanen *et al.* (2004), Stockall e Marantz (2006), Zweig e Pylkkanen (2009), Stockall *et al.* (2019) e Pylkkanen *et al.* (2020), enquanto evidências a favor de modelos de dupla rota podem ser encontradas em Fiorentino e Poeppel (2007)<sup>4</sup>, Whiting, Marslen-Wilson e Shtyrov (2013) e Bakker *et al.* (2013). Na subseção que destaca as variáveis envolvidas no processamento lexical, as contribuições de Pylkkanen *et al.* (2004) sobre a variável "frequência cumulativa da raiz" será abordada.

#### 3.4.4 Evidência de estudos com fMRI

O fMRI registra a atividade hemodinâmica do cérebro. Quando os neurônios realizam atividade, eles requerem oxigenação e o oxigênio chega a eles através de um fluxo sanguíneo. Neurônios com maior ativação fazem com que uma o fluxo hemodinâmico de sua região tenha uma atividade maior de trânsito de hemoglobinas oxigenadas e desoxigenadas, as quais exibem propriedades magnéticas diferentes, que o fMRI consegue detectar. Assim, o fMRI consegue obter alta precisão espacial, mas o fluxo sanguíneo tem um atraso de segundos ao passo que processos cognitivos ocorrem em milissegundos, o que faz com que esta técnica não tenha precisão temporal (SOTO, 2014). Para que a ativação de regiões seja associada a processos cognitivos, o fMRI pode contrastar regiões com maior e menor ativação ao longo de testes em

---

<sup>4</sup> Fiorentino e Poeppel (2007) defendem que seus resultados favorecem modelos que preconizam *full parsing*, mas apontam que seus resultados também são condizentes com algumas abordagens de dupla rota em que mesmo o reconhecimento via forma inteira ainda permite efeitos de constituintes.

blocos, cada bloco com estímulos que compartilham de uma condição, distinta da condição do bloco seguinte, ou ainda extrair resultados relacionados a eventos.

O fMRI torna viável, portanto, a localização mais precisa das regiões corticais que estão envolvidas com o processamento lexical. Um estudo de Lehtonen *et al.* (2006) em finlandês, uma língua com uma morfologia rica, mostrou ativação aumentada na Área de Brodmann (BA) 47, no giro frontal inferior esquerdo, e no sulco temporal superior posterior esquerdo (BA 22, 21 e 39) para palavras flexionadas, comparadas com palavras monomorfêmicas (*i.e.*, palavras com apenas um morfema, sem estrutura interna de constituintes). Os autores argumentam que isso demonstra que tais regiões estão envolvidas no processamento de flexão e que a flexão está associada a um nível de processamento sintático, com base nos achados anteriores de que o giro frontal inferior esquerdo é ativado em processos cognitivos de computação sintática. Um estudo em inglês de Vannest *et al.* (2005) mostrou ativação do giro frontal inferior esquerdo também para palavras derivadas.

Embora uma revisão extensa dos dados obtidos com fMRI sobre processamento morfológico não esteja dentro do escopo deste estudo, devido às suas limitações na caracterização do curso temporal, julgamos relevante tratar do estudo realizado por Matchin *et al.* (2019). Os autores testaram a ativação cortical associada a processamentos combinatorial sintático e combinatorial semântico. Em uma condição de sintagma nominal (ex.: *the frightened boy*, o garoto assustado), o particípio não atribuiria papel temático devido à sua estrutura sintática, segundo o autor, sendo a palavra no particípio apenas uma informação semântica a mais no sintagma (logo, não deveria provocar efeitos de combinatorialidade sintática. Já o sintagma verbal (ex.: *frightened the boy*, assustou o garoto) exibiria tanto propriedades sintáticas de seleção de argumento e atribuição de papel temático quanto de combinação semântica, motivo pelo qual os autores afirmam que haveria ativação cortical por combinatorialidade sintática. E uma condição controle de palavras exibidas em lista (e.g., *frightened, scrubbed, wounded*, respectivamente assustou, esfregou, feriu) foi adicionada, que não deveria mostrar nenhuma das propriedades supracitadas. As três condições e a proposta do autor de estrutura sintática a elas associadas estão apresentadas na Figura 6. Os resultados mostraram que a ativação de uma das regiões de interesse, o giro angular no hemisfério esquerdo, não mostrou diferença entre o sintagma nominal e o sintagma verbal, apenas entre estas duas condições e a condição de lista, o que, segundo Matchin *et al.*, demonstra que a região processou informações eventivas apenas, que adiciona informação semântica ao nome do sintagma. Já o sulco temporal posterior e o giro frontal inferior, ambos no hemisfério esquerdo, mostraram-se

sensíveis à combinação sintática, já que houve diferença significativa na ativação entre as condições, maior para os sintagmas verbais que os nominais. Tal resultado dialoga com a hipótese deste texto, já que, apesar da possibilidade de argumentar-se que palavras no particípio em sintagmas nominais serem, na realidade, orações reduzidas ou orações pequenas em uma estrutura não aparente em que a palavra no particípio mantém-se como um verbo e exibe propriedades de seleção argumental, a diferença exibida na ativação cortical mostra que é possível que palavras no particípio, ao menos em inglês, atuem de fato como adjetivos, e não como verbos aparentemente adjetivizados.

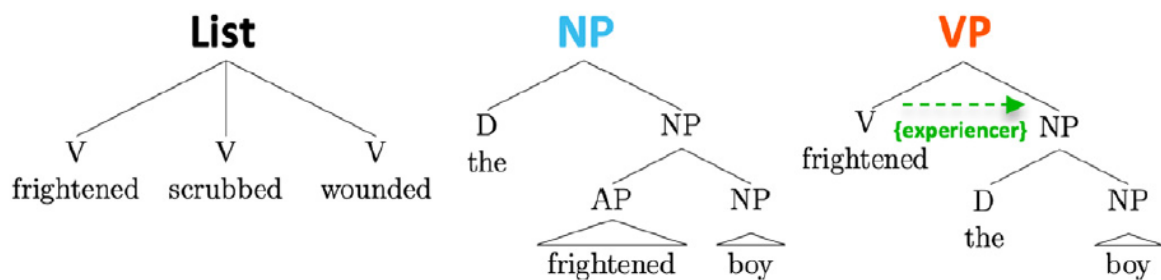


Figura 6 - proposta de análise de Matchin *et al.* (2019, p. 119) das condições controladas em seu experimento. Na esquerda, palavras em lista; no meio, sintagma nominal em que o particípio é um adjetivo que modifica o nome; na direita, um sintagma verbal em que o particípio encabeça a oração e atribui papel temático ao NP.

### 3.4.5 Conclusões sobre os modelos de processamento a partir da literatura

Os achados na literatura divergem de formas difíceis de conciliar. Chama-se atenção para a constatação de que, em algum nível, seja como opção de via ou como via única de processamento, há efeitos de morfologia no processamento lexical de palavras de morfologia complexa, efeitos estes que estão relacionados à morfo-ortografia/morfologia, e não à relações semânticas entre palavras morfologicamente relacionadas.

### 3.5 Variáveis que afetam o processamento lexical e morfológico

Os últimos anos viram uma série de estudos psiconeurolinguísticos testarem quais variáveis afetam o processamento lexical. Algumas variáveis, como será explanado, afetam o processamento lexical de palavras simples ou de morfologia complexa (i.e., casos de palavras derivadas, flexionadas ou compostas). Outras, segundo seus propositores, podem afetar a via



pela qual uma palavra de morfologia complexa é processada: se por sua forma inteira, ou pela segmentação em constituintes.

*Tamanho da palavra* - Baayen (2005) e Baddeley, Thomson e Buchanan (1975) mostraram que o tamanho da palavra impacta o seu tempo de processamento. Baayen o demonstrou em contexto de leitura de palavras através de mineração de dados de experimentos comportamentais e constataram duas variáveis com efeito sobre a latência dos TRs: tamanho da palavra (medido em quantidade de grafemas) e densidade da vizinhança ortográfica (grupo de palavras com formas ortográficas que compartilham segmentos)<sup>5</sup>. Já Baddeley Thomson e Buchanan (1975) colocaram seu ponto em relação ao tamanho da forma fonológica das palavras, já que a maior dificuldade de articulação incorreria em, também, maiores problemas na recuperação da palavra na memória: quanto maior a forma, maior seria em média o tempo de processamento.

*Probabilidade fonotática e vizinhança fonológica* - Além da constatação de Baddeley Thomson e Buchanan (1975) em termos de fonologia, Pytkkanen, Stringfellow e Marantz (2002) testaram a probabilidade fonotática de uma palavra como uma variável que impacta de processamento. Ela se refere à alta ou baixa probabilidade de uma sequência de sons dentro de uma palavra ocorrer na língua, e está conectada à alta ou baixa densidade de vizinhança fonológica (grupo de palavras com formas fonológicas que compartilham segmentos). O estudo apontou que há um efeito facilitatório das palavras com alta probabilidade fonotática e alta densidade de vizinhança no estágio referente à ativação de palavras na memória que estejam compatíveis com o estímulo exibido, mas que no estágio em que o cérebro deve escolher qual forma de uma vizinhança de palavras similares corresponde ao estímulo exibido, a alta densidade de vizinhança incorre em inibição, ou seja, uma "dificultação" do processo. Isso explicaria resultados neurofisiológicos e comportamentais aparentemente incoerentes, já que flagram momentos distintos no processamento, sendo que a seleção do candidato impacta diretamente o tempo de resposta que reflete o final da tarefa em que o participante indica o reconhecimento do item lexical.

*Vizinhança ortográfica* - Estivalet (2020) reportou um efeito, já discutido anteriormente na literatura, das variáveis frequência lexical e densidade de vizinhança ortográfica, esta última voltada para palavras exibidas visualmente. Enquanto muitos estudos são realizados em línguas como inglês e alemão, um diferencial de Estivalet é que seu teste é sobre português brasileiro.

---

<sup>5</sup> Coltheart *et al.* (1977) definem vizinhança ortográfica como palavras que podem pela substituição de uma letra de uma palavra. Um exemplo dado é da palavra em inglês "sand"/areia, que tem 11 vizinhos, já que há 11 palavras que compartilham três das quatro letras com elas.

O autor reporta efeitos tanto de frequência lexical quanto de densidade de vizinhança ortográfica no TR a tarefa de decisão lexical: a alta frequência e a alta densidade de vizinhança se mostram facilitadoras no curso do processamento. Ressalta-se que a natureza da tarefa pode não envolver estágios posteriores à ativação de candidatos na memória, o que explica a alta densidade ser tomada como facilitadora neste tipo de tarefa, mas inibidora em um estudo online de MEG como o de Pylkkanen, Stringfellow e Marantz (2002), já que este último consegue discernir os estágios do processo de reconhecimento de palavras para apontar em que ponto do curso temporal um efeito é facilitador, e quando a mesma variável pode ter um efeito inibidor. De todo modo, fica demonstrado tanto em âmbito de leitura quanto de escuta o impacto da variável densidade de vizinhança ortográfica/fonológica. Um parêntesis importante que se deve realizar é o achado de Marian *et al.* (2012) de que palavras com mais de 8 letras se desconectam da variável densidade de vizinhança ortográfica, já que tamanho maior da palavra significa maior complexidade ortográfica e uma vizinhança ortográfica pequena.

*Frequência cumulativa da raiz* - Outro estudo de MEG por Pylkkanen *et al.* (2004) com palavras ouvidas mostra o impacto de uma variável de natureza morfológica: frequência acumulada da raiz. Palavras como "batata-barata" ocupam uma mesma vizinhança fonológica mas sua relação está apenas na forma. Já palavras como "feliz, felizmente, felicidade" compartilham de uma mesma raiz e, por isso, diz-se que pertencem à mesma família morfológica. Assim, cada ocorrência de um membro de uma família morfológica acresce para a medida de frequência da raiz comum à família, por isso diz-se que uma família morfológica que ocorre frequentemente provoca um efeito de alta frequência cumulativa da raiz. Semelhantemente à densidade de vizinhança fonológica, foi mostrado por Pylkkanen *et al.* (2004) que o tamanho da família morfológica afeta o curso do processamento lexical, com uma diferença significativa: o efeito facilitatório da alta frequência cumulativa da raiz faz com que o estágio de competição ocorra mais cedo no curso temporal do processamento em comparação com o estágio de competição por similaridade fonológica exibido no estudo de Pylkkanen, Stringfellow e Marantz (2002). Esse achado se apoia na noção de que palavras de morfologia complexa são segmentadas no curso do processamento e também provê evidência favorável ao modelo de decomposição automática, além de constituir mais uma evidência de que a frequência é uma variável preditora de medidas comportamentais.

*Tamanho da família morfológica* - Em uma série de quatro experimentos, Jong, Baayen e Schreuder (2000) reportam um efeito da variável tamanho da família morfológica. Os autores mostraram que há um efeito facilitatório de famílias morfológicas grandes (comparadas com

famílias pequenas) que possui natureza semântica. Este efeito é dissociado de um efeito de vizinhança ortográfica, já que vizinhança ortográfica grande pode ter um efeito inibitório (*i.e.*, dificultador) no acesso lexical em decorrência de competição entre candidatos. Este estudo também mostra que a morfologia é parte do processamento lexical (rememora-se que há propostas de arquitetura da linguagem na mente que não preveem um nível estritamente morfológico). O efeito facilitador da variável tamanho de família morfológica também é reportado em Ford, Davis e Marslen-Wilson (2010).

Outras variáveis têm importância para a discussão sobre a ocorrência de dupla rota no processamento de palavras de morfologia complexa. A premissa é de que se há duas rotas, então há alguma variável (ou, ainda, variáveis) que ativa a rota de decomposição em morfemas ou de processamento pela forma inteira da palavra (sem olhar para estrutura interna dela).

*Frequência superficial* - Para Schreuder e Baayen (1995) uma delas é a frequência superficial. Como explicado anteriormente, a frequência é relativa ao total de ocorrências de uma palavra ou parte dela em uma determinada língua. Quando se refere à ocorrência da palavra inteira, refere-se à frequência superficial. Para os autores, como também mencionado anteriormente, o que define a rota que uma palavra de morfologia complexa toma no processamento é frequência e a transparência semântica: palavras de alta frequência e semântica transparente são processadas por forma inteira, enquanto palavras derivadas e flexionadas de baixa frequência e as opacas são segmentadas em morfemas.

O efeito de frequência superficial foi reportado como preditor de resultados em alguns estudos. No estudo de Leminen *et al.* (2013), palavras derivadas de alta frequência provocavam os maiores picos de MMN na gravação de dados do EEG, o que os autores tomam como evidência de alto nível de ativação, um reflexo da sua pronta disposição a ser ativada. Em Baayen, Dijkstra e Schreuder (1997), substantivos, palavras regulares no plural de alta frequência em holandês, tinham tempos de resposta mais curtos comparados com palavras nas mesmas condições mas de baixa frequência. Estes estudos sugerem que a alta frequência nas respectivas condições controladas favoreceu a armazenagem em forma inteira e reafirmam a existência de um modelo de dupla rota de processamento de palavras de morfologia complexa.

*Frequência relativa e frequência de base* - Outros dois tipos de frequência são a frequência de base e frequência relativa. A frequência de base é a frequência com que a palavra base de palavras morfológicamente complexas ocorre. Já a frequência relativa é uma medida, como diz o nome, relativa que contrasta a frequência de base com a frequência superficial. Hay (2001) atribui à frequência relativa o efeito de armazenagem em forma inteira ou

decomposicionalidade de palavras derivadas: quando a frequência da palavra base é maior que a frequência da forma inteira, segundo a autora, a palavra será segmentada em morfemas, independentemente da forma inteira da palavra ocorrer com alta ou baixa frequência em uma língua. No entanto, o experimento de Hay baseou-se na intuição dos participantes em determinarem qual palavra em um par é mais facilmente segmentada em morfemas. Barbosa (2017) testou a mesma hipótese em português em experimentos, mas utilizou também medida de tempo de resposta em tarefa lexical para apontar a mesma consideração de Hay (2001). Tal efeito é contestado em Taft (2004), que continua a defender seu modelo de decomposição precoce automática contra modelos de dupla rota.

*Regularidade* - A variável de regularidade da palavra nos processos de formação foi proposta como preditora para modelo de dupla rota por Pinker (1991) e testada por Ullman *et al.* (1997). Conforme citado anteriormente, palavras regulares são aquelas cuja formação se dá por uma regra, como o caso de palavras regulares no passado do inglês que utilizam "-ed" (ex.: *play-played/jogar-jogou*), e palavras irregulares são aquelas que perpassam alterações fonológicas/ortográficas que escapam a uma regra (ex.: *run-ran/correr-correu*). Segundo Pinker e Ullman e colaboradores, palavras regulares são segmentadas em morfemas, enquanto palavras irregulares são listadas na memória em forma inteira. Ressalta-se, no entanto, que não há consenso quanto a esta variável ser a única a atuar na preferência de via de processamento nos modelos de dupla rota nem quanto a não haver decomposicionalidade de palavras irregulares.

*Processo de formação da palavra* - Leminen *et al.* (2013), como citado anteriormente, manipularem a variável processo de formação da palavra (derivação *versus* flexão) em seu estudo. A sugestão apontada pelos autores foi que palavras de alta frequência superficial e derivadas se beneficiariam, em finlandês, da armazenagem em forma inteira já que a produtividade dos afixos derivacionais é baixa, enquanto a flexão é muito produtiva, de modo que para este último processo listar na memória a forma inteira de todas as palavras flexionadas de alta frequência seria muito custoso, sendo mais econômico manter constituintes que são conectados no curso do processamento.

O próximo capítulo se refere ao Experimento 1, de tarefa de decisão lexical sob o paradigma de *priming* encoberto. Nele, serão reportadas a metodologia utilizada (desenho experimental, materiais, participantes e procedimentos), a análise estatística e os resultados assim como será realizada uma discussão em detrimento dos resultados e dos achados reportados na literatura.

# 4 EXPERIMENTO 1

## 4.1 Metodologia

### 4.1.1 Hipótese e expectativas

Com base no estudo de Leminen *et al.* (2013), elaborou-se a hipótese de que palavras na condição de participio preferencialmente nominal (segundo tendência de leitura obtida a partir de pré-teste), por terem associação maior a classes nominais (substantivo e adjetivo), seriam casos de derivação e, por isso, quando fossem de alta frequência, seriam passíveis de armazenagem em forma inteira no léxico mental e processamento por esta via, conforme sugerido por Leminen *et al.* Já as palavras nas condições de preferencialmente nominal de alta frequência e preferencialmente verbal de alta e baixa frequência seriam processadas por segmentação morfológica automática precoce, preservando-se no léxico mental suas representações em constituintes. Esta hipótese está de acordo com modelos de dupla rota de processamento lexical ao considerar, de um lado, o impacto da variável de frequência superficial (SCHREUDER; BAAYEN, 1995) mas, por outro, a existência de vias paralelas ou por forma inteira ou por decomposição em morfemas (PINKER, 1991) que não interagem como proposto por Schreuder e Baayen.

A confirmação da hipótese se daria pela ocorrência de tempo de resposta (TR) médio correspondente às tarefas de palavras na condição preferencialmente nominal de alta frequência maior que o TR médio relativo à condição preferencialmente nominal de baixa frequência, já que estes se beneficiariam do *prime* encoberto de relação morfológica. Rememora-se que o paradigma de *priming* encoberto provoca facilitação significativa por relação morfo-ortográfica, mas não por relação somente ortográfica ou semântica (RASTLE; DAVIS; NEW, 2004; RASTLE; DAVIS, 2008) e que o *priming* encoberto é relacionado à ativação de etapas iniciais anteriores ao acesso semântico no curso do processamento lexical, chegando-se apenas à decomposição em morfemas com base em informações morfo-ortográficas. Se as palavras na condição preferencialmente nominal de alta frequência estivessem de fato armazenadas em forma inteira, a relação morfológica obtida pela segmentação em “raiz + morfema” do *prime* não deveria facilitar o tempo de processamento desta condição, já que a armazenagem em forma inteira e processamento nesta via não prevê informações de estrutura interna da palavra. Em

contrapartida, as palavras nas condições preferencialmente verbal de alta e baixa frequênciateriam TRs médios semelhantes ou próximos, já que o estudo de Leminen *et al.* (2013) sugere que casos de flexão verbal não estariam sujeitos à armazenagem em forma inteira independentemente de sua frequência superficial, o que possibilitaria que as palavras nestas condições se beneficiem de facilitação pela raiz já ativadas no *prime* encoberto de relação morfológica. Ademais, há duas expectativas com relação à comparação entre categorias separadas por tipo de *prime*. Quando o *prime* é morfológicamente relacionado, espera-se que os participios “preferencialmente” verbais de alta frequência teriam o TR médio mais rápido por se beneficiar tanto do efeito de alta frequência quanto do *prime* morfológico. Quando o *prime* é não-relacionado, espera-se que o participio “preferencialmente” nominal de alta frequência tenha o TR médio mais rápido, já que não requereria etapas a mais de processamento referentes ao processamento decomposicional e teria o efeito facilitador da alta frequência, enquanto o participio “preferencialmente” verbal de alta frequência teria facilitação pela alta frequência, mas com custo cognitivo por passar por etapas de processamento decomposicional<sup>6</sup>, sendo este mais lento que aquele, enquanto as condições com maior TR médio aproximados seriam os participios “preferencialmente” nominais e verbais de baixa frequência, já que ambos passariam por segmentação morfológica e não teriam o efeito facilitador da frequência superficial.

#### 4.1.2 Desenho experimental e materiais

O Experimento 1 utilizou o paradigma de *priming* encoberto com tarefa de decisão lexical cujo recorte é investigação sobre processamento lexical de palavras no participio passado de primeira conjugação ("-ar") do português brasileiro.

As variáveis independentes deste experimento foram: categoria (preferencialmente nominal ou verbal), frequência superficial (alta ou baixa) e tipo de relação entre *prime* e alvo (morfológicamente relacionados ou não-relacionados). A categoria diz respeito a uma tendência de leitura da palavra como pertencente a alguma classe, nominal (com propriedades potencialmente de adjetivos ou substantivos) ou verbal. Para que as palavras sejam atribuídas a uma dessas categorias, um pré-teste foi realizado, o qual será reportado mais adiante. A frequência superficial, que se refere à frequência de ocorrência da forma inteira de uma palavra em uma língua, foi dividida em duas condições: alta ou baixa. A classificação como alta ou baixa

---

<sup>6</sup> Esta expectativa tem por base considerações de um estudo em finlandês de Lehtonen *et al.* (2008) que detectou reconhecimento mais lento em palavras que perpassam segmentação morfológica.

frequência teve por base a Escala Zipf (VAN HEUVEN *et al.*, 2014), que também será detalha. Por fim, houve a variável de tipo de relação *prime*-alvo, que foi morfológicamente relacionado (ex.: expulsou-EXPULSADO) ou não-relacionado (ex.: reserva-EXPULSADO). Os *primes* relacionados morfológicamente com os alvos eram verbos de primeira conjugação flexionados na terceira pessoa do pretérito perfeito já que, por ser uma classe de conjugação regular e palavras de categoria verbal sem possíveis ambiguidades (como o caso do particípio passado, que pode gerar adjetivos), e se baseando nos estudos que apontam maior tendência à segmentação em flexões, esperamos que estes *primes* fossem segmentados e facilitassem a reativação da raiz no alvo, caso este fosse passível de segmentação morfo-ortográfica.

As condições das palavras alvo foram divididas conforme as variáveis independentes em: particípio preferencialmente nominal de alta frequência (PNAF, n=20), particípio preferencialmente nominal de baixa frequência (PNBF, n=20), particípio preferencialmente verbal de alta frequência (PVAF, n=20) e particípio preferencialmente verbal de baixa frequência (PVBF, n=20). Para cada palavra no particípio, houve um *prime* morfológicamente relacionado (ou seja, cuja raiz era a mesma da palavra alvo) conjugado no pretérito perfeito. Como a tarefa requer estímulos que não sejam palavras, foram adicionadas pseudopalavras no particípio passado (PP) de primeira conjugação (n=40) que também possuíam um *prime* com relação pseudomorfológica. Para estabelecer um *baseline* que permitisse a detecção do efeito de *priming*, cada palavra alvo e pseudopalavra também foi pareado com um *prime* não-relacionado. Duas listas de estímulos foram criadas de acordo com a relação *prime*-alvo que compuseram um formato de quadrado latino (ver Tabela 1): a lista 1 (L1) tinha a metade A dos estímulos com *primes* relacionados morfológicamente e a metade B com *primes* não-relacionados, e a lista 2 (L2) tinha a metade A com *primes* não-relacionados e a metade B com *primes* morfológicamente relacionados. Com o cruzamento das variáveis de tipo de particípio, tipo de frequência e tipo de *prime*, obteve-se um total de 8 condições nas palavras alvo.

	<b>Prime morfológicamente relacionado</b>	<b>Prime não relacionado</b>
<b>Lista 1</b>	ex.: colocou-COLOCADO	ex.: cadela-LARGADO
<b>Lista 2</b>	ex.: largou-LARGADO	ex.: casa-COLOCADO

Tabela 1 - distribuição em listas dos estímulos por tipo de *prime* no formato de quadrado latino do Experimento 1.

Também foram adicionadas palavras distratoras (n=40) e pseudopalavras distratoras (n=40). Optou-se por não utilizar a proporção de 50/50 entre palavras alvo e distratoras em decorrência de um potencial efeito de desengajamento pelos participantes no decorrer do experimento por motivos de tamanho deste. No total, cada participante via 240 pares de *prime*-alvo, dos quais 20 pares em cada condição (dividido entre par morfológicamente relacionado e não-relacionado), levando a um total de 80 pares experimentais, e mais 80 pares com alvo pseudopalavra (40 com pseudoprime “relacionado”, e 40 com pseudoprime não-relacionado), além de 80 distratores (metade pseudopalavras, metade palavras). A Tabela 2 mostra exemplos de estímulos alvos e distratores.

Condição	<i>prime</i> relacionado	ALVO	<i>prime</i> não-relacionado	ALVO	Frequência média
<i>PNAF</i>	ocupou	OCUPADO	papel	OCUPADO	4,34 (DP:0,24)
<i>PNBF</i>	alagou	ALAGADO	cadeira	ALAGADO	2,54 (DP:0,30)
<i>PVAF</i>	mostrou	MOSTRADO	espada	MOSTRADO	4,23 (DP:0,15)
<i>PVBF</i>	abdicou	ABDICADO	letra	ABDICADO	2,35 (DP:0,36)
<i>PP</i>	soilou	SOILADO	alface	SOILADO	<i>n.a.</i>
<i>Palavra distratora</i>	<i>n.a.</i>	<i>n.a.</i>	birra	BORRACHA	<i>n.a.</i>
<i>Pseudopalavra distratora</i>	<i>n.a.</i>	<i>n.a.</i>	certendo	CERTENDOR	<i>n.a.</i>

Tabela 2 - Exemplos de estímulos por condição e tipo de *prime* no Experimento 1. PN: particípio preferencialmente nominal. PV: particípio preferencialmente verbal. AF: alta frequência. BF: baixa frequência. PP: pseudopalavra no particípio. DP: desvio padrão

Os materiais e informações de frequência dos mesmos foram retirados do LexPorBr (ESTIVALET; MEURNIER, 2017). A divisão das palavras em particípio preferencialmente nominal ou preferencialmente verbal foi feita, em um primeiro momento, de acordo com a intuição do pesquisador. Para verificar a validade de tal divisão de forma menos arbitrária, dois pré-testes foram preparados em que 10 participantes foram instruídos a classificar as palavras apresentadas em caixa alta em uma escala de Likert de 1 a 5, em que 1 correspondia a preferencialmente verbal e 5, preferencialmente nominal (substantivo ou adjetivo). A instrução aos participantes para a realização da tarefa está retratada na Figura 7.



## Pre-teste final

Preencha as escalas abaixo conforme a opção que mais parece natural a você.

Por exemplo, em uma escala de 1 a 5, em que 1 é verbal e 5 é nominal(adjetivo/substantivo):

- Marque 1 se você acha que a palavra só faz sentido como verbo.
- Marque 2 se você acha que a palavra faz mais sentido (mas não comente) como verbo.
- Marque 3 se você acha que a palavra faz igualmente sentido como verbo e nome (adjetivo/substantivo).
- Marque 4 se você acha que a palavra faz mais sentido (mas não comente) como nome (adjetivo/substantivo).
- Marque 5 se você acha que a palavra só faz sentido como nome (adjetivo/substantivo).

---

PREPARADO

	1	2	3	4	5	
Verbo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Nome (adjetivo/substantivo)

ROUBADO

	1	2	3	4	5	
Verbo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Nome (adjetivo/substantivo)

Figura 7 - Screenshot de parte do pré-teste, realizado via Google Formulários.

Uma análise estatística de *kmeans* (HARTIGAN; WONG, 1979) foi feita para observar se as palavras se conformavam em uma divisão em dois *clusters*<sup>7</sup>, como previsto originalmente, e após apenas algumas palavras não recaírem sobre a classificação a elas atribuídas originalmente, estas foram retiradas, outras adicionadas em seu lugar e um novo pré-teste semelhante foi realizado a fim de que se completasse a lista de estímulos. O segundo pré-teste confirmou as últimas palavras adicionadas. As médias de julgamento de cada cluster foram de 3,43 (SD: 0,47) no "preferencialmente verbal" e de 2,12 (SD: 0,45) no "preferencialmente nominal".

<sup>7</sup> “Clusterização é o agrupamento automático de instâncias similares, uma classificação não-supervisionada dos dados (esse termo não é familiar? Veja nosso post sobre os três tipos de aprendizado de máquina!). Ou seja, um algoritmo que clusteriza dados os classifica em conjuntos de dados que ‘se assemelham’ de alguma forma - independentemente de classes predefinidas. Os grupos gerados por essa classificação são chamados clusters.” (HONDA, 2017)

## Divisão em 2 clusters

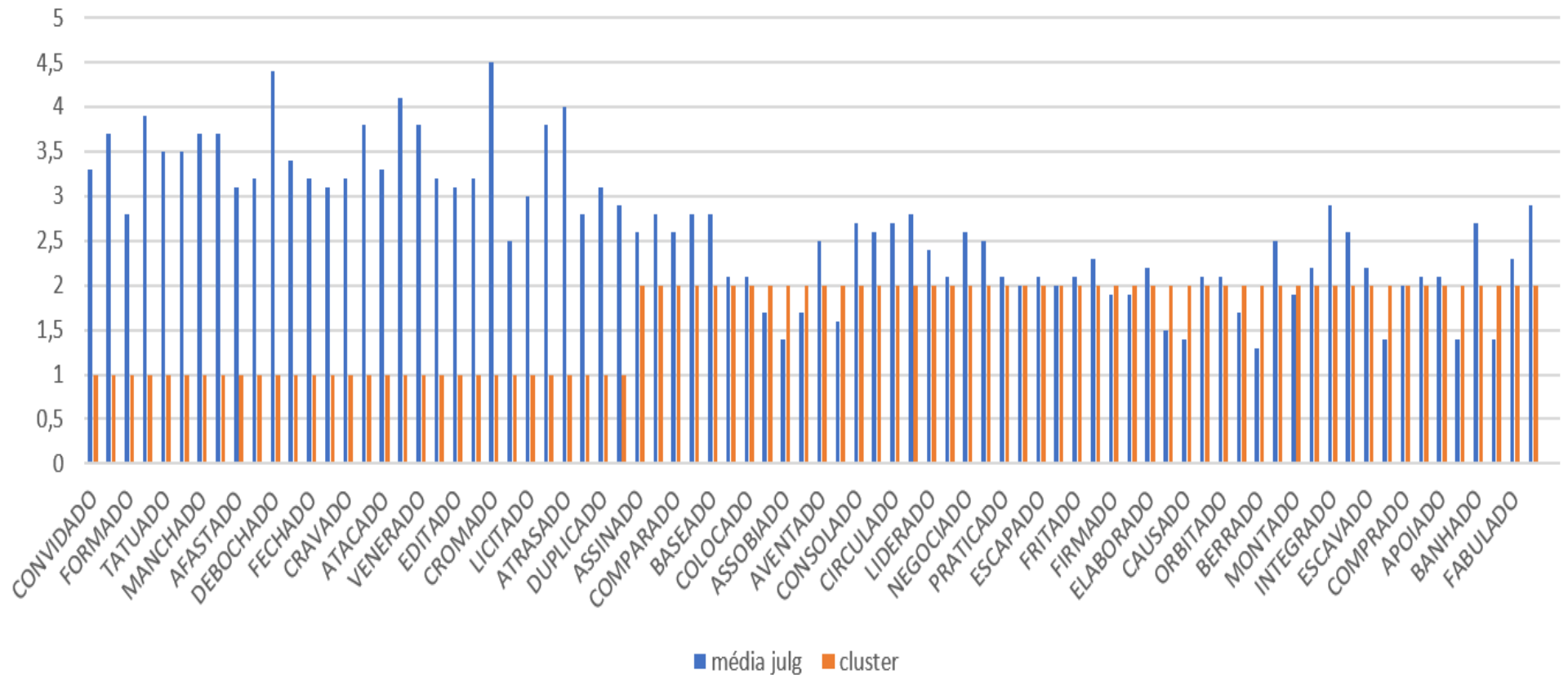


Gráfico 1 - resultado do agrupamento em clusters a partir da análise de *kmeans*. A média dos julgamentos do primeiro pré-teste no eixo Y. O *cluster* 1 equivale à categoria participio preferencialmente nominal e o 2, à categoria participio preferencialmente verbal.

As listas de alta ou baixa frequência foram divididas com base na Escala Zipf (VAN HEUVEN et al., 2014), em que um valor acima de 4 corresponde a alta frequência e abaixo de 3, a baixa frequência. Como algumas palavras no corpus possuíam mais de uma entrada, o que pode ocorrer quando ela está classificada simultaneamente como, por exemplo, verbo e adjetivo (ex.: roubado), também havia casos em que uma mesma palavra recebia duas ou mais contagens de ocorrência no corpus e dois valores na Escala Zipf. Por isso, foi necessário recalcular o valor atribuído a cada palavra com base na fórmula proposta por VAN HEUVEN et al. (2014) cujo resultado é o índice da escala:  $\log_{10}(\text{TOTAL DE OCORRÊNCIAS NO CORPUS} + 1/\text{TOTAL DE TOKENS POR MILHÃO} + \text{TOTAL DE TIPOS POR MILHÃO}) + 3$ , o que envolveu pesquisar uma mesma palavra no corpus e somar o total de frequências de todas as suas entradas, quando havia mais de uma. Não tinha diferença entre as palavras de alta frequência de condições diferentes ( $t=0,98$ ,  $p=0,76$ ), nem entre as de baixa frequência de condições diferentes ( $t=2,23$ ,  $p=0,13$ ); e, sim, entre os índices de Zipf de alta e baixa frequência para cada categoria (nominal:  $t=16,44$ ,  $p<0,0001$  e verbal:  $t=21,16$ ,  $p<0,0001$ ).

As palavras adicionadas ao experimento também foram escolhidas com base no comprimento e quantidade de sílabas. Optou-se pelo intervalo de palavras de entre 7 e 9 letras e que tivessem entre 3 e 4 sílabas. O intuito foi anular um possível efeito de tamanho da palavra na variável dependente, caso houvesse palavras com tamanhos muito diferentes.

A lista correspondente a cada grupo experimental também foi dividida em quatro listas pseudorandomizadas. O intuito relativo a essa medida foi evitar o risco de que um mesmo participante de um dado grupo experimental repetisse o experimento com as palavras aparecendo na mesma ordem. Ademais, a pseudorandomização permitiu que houvesse maior controle sobre a distribuição sequencial de tipos de estímulo exibidos e respostas à tarefa (por exemplo, foi possível evitar assim que um participante lesse dez vezes seguidas palavras cuja resposta à tarefa fosse a mesma).

#### 4.1.3 Participantes

Os participantes foram estudantes de graduação em Letras na faixa etária entre 18 e 29 anos. Houve um total de 63 participantes, dos quais 7 foram excluídos por taxa de acurácia abaixo do limiar de 80,5% ou por abandono do teste. A pesquisa foi aprovada pela Comissão de Ética da UFRJ sob o CAAE 50670421.3.0000.5699, parecer nº 5.185.244. Todos os

participantes confirmaram sua concordância digitalmente com o Termo de Consentimento Livre e Esclarecido.

#### 4.1.4 Procedimentos

O experimento tratou-se de um teste de leitura de palavras que apareciam em pares *prime*-ALVO, ambos exibidos visualmente. O teste foi realizado de forma remota *online* através da plataforma PCIBEX. Os participantes entraram através de um *link*, o qual levou à página do experimento. Após aceitar o Termo de Consentimento Livre e Esclarecido (TCLE), havia o preenchimento de algumas informações pessoais e o participante passava, então, a uma etapa de treinamento. Nesta etapa, o participante lia a palavra alvo no centro da tela e era requerido que pressionasse a tecla predeterminada correspondente à resposta da tarefa no teclado, sem que houvesse um limite de tempo: (A) para sim, ou (B) para não. O participante, então, era informado quanto a se sua resposta era condizente ou não com a expectativa. Após cinco rodadas de treinamento, passava-se a uma etapa em que o experimento era simulado como ocorreria nas rodadas principais, sem que houvesse pausas entre um estímulo e outro e com um *timeout* após o qual o experimento passava para o par *prime*-alvo seguinte, mesmo que o participante não tivesse respondido a tarefa. Surgia um aviso de entrada na parte principal do teste, e o participante era convidado a pressionar a barra de espaço do teclado para dar início ao experimento. Também orientou-se que a resposta fosse dada o mais rapidamente o possível. No teste principal, os procedimentos foram os seguintes: uma máscara de *hashtags* (#####) era exibida por 500ms, seguida do *prime* em letras minúsculas por 42ms e, então, a palavra alvo era exibida em letras maiúsculas por 500ms, sem intervalo entre ela e o *prime*. Havia um tempo de 2s para que o participante pressionasse a tecla correspondente à resposta para a tarefa após o qual não era mais possível realizá-la (*timeout*), adicionado de um intervalo de 1s entre uma rodada e a seguinte. A utilização da máscara e o tempo curto de exibição do *prime* estão de acordo com o paradigma de *priming* encoberto, para que não haja percepção consciente do *prime*. Para evitar efeito de cansaço por excesso de tempo realizando o experimento, o teste principal foi dividido em dois blocos com um intervalo no meio, em que o participante era instruído a realizar uma pausa, mas que não saísse da janela do experimento ou da frente do computador. Cada bloco incluía 120 *trials*, totalizando 240.

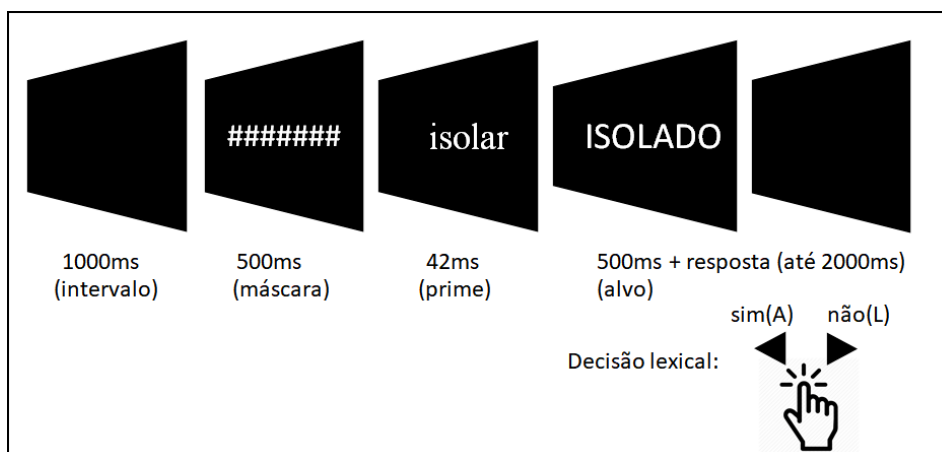


Figura 8 - Figura ilustrativa da etapa principal do experimento. Após um intervalo com a tela em branco, exibem-se a máscara, o *prime* e o alvo. Após o momento em que se exibe o alvo na tela, responde-se a tarefa de decisão lexical.

#### 4.2 Análise estatística

A análise estatística foi rodada no programa de código aberto *Rstudio* (versão 1.2.5033; R Core Team, 2013). O modelo adotado foi o modelo de regressão linear de efeitos mistos, cuja análise prevê comportamento aleatório em variáveis independentes, através da aplicação da função *lmer* do pacote *lme4* (BATES *et al.*, 2015). Houve duas variáveis dependentes no Experimento 1: a de tempo de resposta (TR), utilizada na discussão, e a de taxa de acurácia. A acurácia foi utilizada a fim de que se controlasse a atenção dos participantes e obteve-se dela um limiar abaixo do qual os participantes foram excluídos dos resultados (limite mínimo de 80,5%); por conta disso, apenas a variável TR foi incluída na análise estatística, baseando-se em um tempo mínimo de 200ms de TR para que se considerasse que a palavra fora realmente lida, critério este também utilizado na exclusão de dados. As variáveis independentes foram de categoria da palavra (PN ou PV), frequência superficial (alta ou baixa) e tipo de *prime* (relacionado ou não relacionado). E, por fim, as variáveis aleatórias foram participante e item (corresponde a cada palavra alvo). A identificação visual de *outliers* foi realizada pela função *deggplot* (WICKHAM, 2006); a partir dela, não houve remoção de dados. Apenas os TRs de respostas corretas foram incluídos na análise. Após rodar-se um teste de normalidade de Jarque-Bera (JARQUE; BERA, 1980), os dados foram transformados utilizando com uma função lambda através da função *BoxCox* do pacote *fpp* (HYNDMAN; KHANDAKAR, 2008), mas para fins de compreensão das médias, os TRs brutos foram mantidos nos gráficos. A significância dos modelos testados foi analisada através de comparações aninhadas, e os p-

valores foram calculados através do pacote lmerTest (KUZMETSOVA; BROCKHOFF; CHRISTENSEN, 2017). Testes de Tukey HSD foram aplicados para a realização de comparações de pares múltiplos *post-hoc* pelo pacote lsmeans (LENTH, 2015).

### 4.3 Resultados

#### 4.3.1 Acurácia

A taxa de acurácia média foi de 86,6%, com um desvio padrão de 8,1%. A partir de uma subtração da taxa média do desvio padrão, estabeleceu-se um limite mínimo de 80,5% de acerto nas tarefas para que os dados do participante não fossem excluídos dos resultados. A partir desse critério, os resultados de 7 participantes foram removidos do experimento, e 1 sujeito foi eliminado por problemas na gravação de dados do seu experimento, que não foi adequadamente finalizada. Dessa forma, 55 conjuntos de dados foram adicionados à análise.

Ao separar-se as condições, houve acerto médio de 92,7% (SD: 6,87%) nas tarefas relativas a palavras e de 85% (SD: 13,6%) no caso de pseudopalavras. Isto indica uma dificuldade maior nas respostas a tarefas com pseudopalavras, o que se justifica por uma semelhança fonológica entre estas e palavras reais do português em início de palavra, assim como pelo fato de que as pseudopalavras possuíam terminação em “-(a)do”, um afixo existente na língua.

#### 4.3.2 Tempos de resposta

O paradigma de *priming* encoberto demonstrou-se significativo na facilitação de palavras de morfologia complexa, efeito este que se obteve por um ajuste de fórmula em que as variáveis independentes eram o tipo de *prime* (morfologicamente relacionado ou não-relacionado) e o *status* lexical (palavra ou pseudopalavra) com participantes e itens como efeitos aleatórios (TR ~ tipo de *prime* + *status* lexical + (1|participante) + (1|item)). Uma comparação com um modelo nulo mostrou a significância estatística da fórmula apontada ( $X^2(3)=210,45$ ,  $p<0,000$ ). As palavras em que o *prime* possuía relação morfológica com o alvo foram processadas mais rapidamente que aquelas em que o *prime* era não-relacionado ( $\beta=-0,008$ ,  $SE=0,00018$ ,  $t=-4,59$ ,  $p<0,000$ ), como demonstrado no Gráfico 2. Esta diferença não se observou na comparação de pseudopalavras cujo *prime* era pseudo relacionado com as pseudopalavras em que o *prime* era não-relacionado ( $\beta=-0,00016$ ,  $SE=0,00026$ ,  $t=-0,621$ ,

p=0,93), o que soma evidências à constatação de que a facilitação pelo *priming* encoberto não decorre de relações puramente ortográficas e que o efeito de *priming* restringiu-se, neste experimento, às palavras.

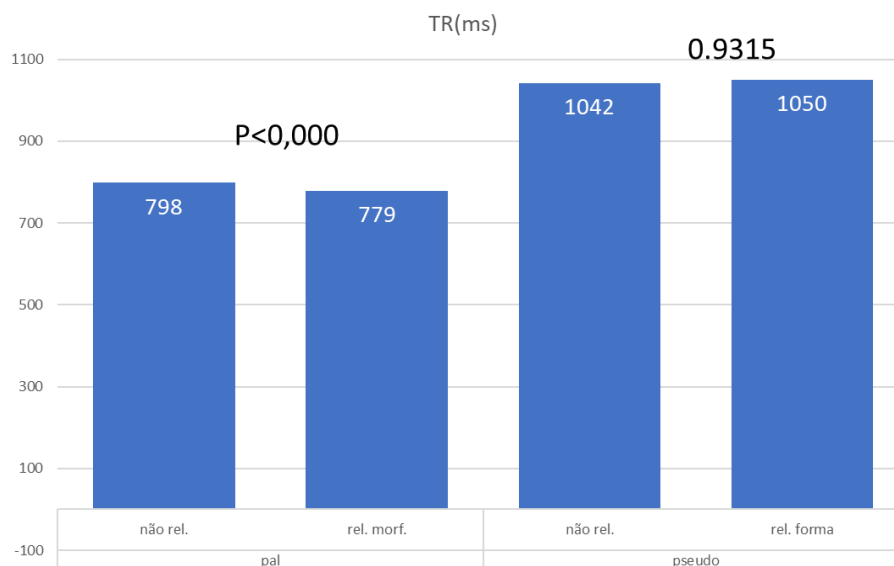


Gráfico 2 - comparação de efeito de *priming* em palavras e em pseudopalavras.

A análise do impacto da variável categoria da palavra (preferencialmente nominal ou preferencialmente verbal) foi realizada utilizando-se apenas o subconjunto de dados de palavras reais. A variável categoria da palavra não obteve significância estatística na análise. Para chegar a este resultado, realizou-se uma comparação aninhada do modelo mais complexo com interação ( $TR \sim \text{tipo de prime} * \text{categoria} * \text{frequência} + 1 | \text{participante} + 1 | \text{item}$ ) e com o modelo apenas com as variáveis tipo de *prime* e frequência com interação com o modelo nulo. O modelo mais complexo ( $X^2(3)=74,89$ ,  $p<0,000$ ) e o modelo que incluiu apenas as variáveis tipo de *prime* e frequência ( $X^2(2)=74,55$ ,  $p<0,000$ ) foram significativos comparados com o modelo nulo. Contudo uma comparação dos modelos com interação com os modelos sem interação não atingiu significância estatística ( $X^2(4)=4,91$ ,  $p=0,30$  e  $X^2(1)=0,12$ ,  $p=0,73$ , respectivamente). Por isso, utilizou-se o modelo mais simples ( $TR \sim \text{tipo de prime} + \text{frequência} + (1 | \text{participante}) + (1 | \text{item})$ ).

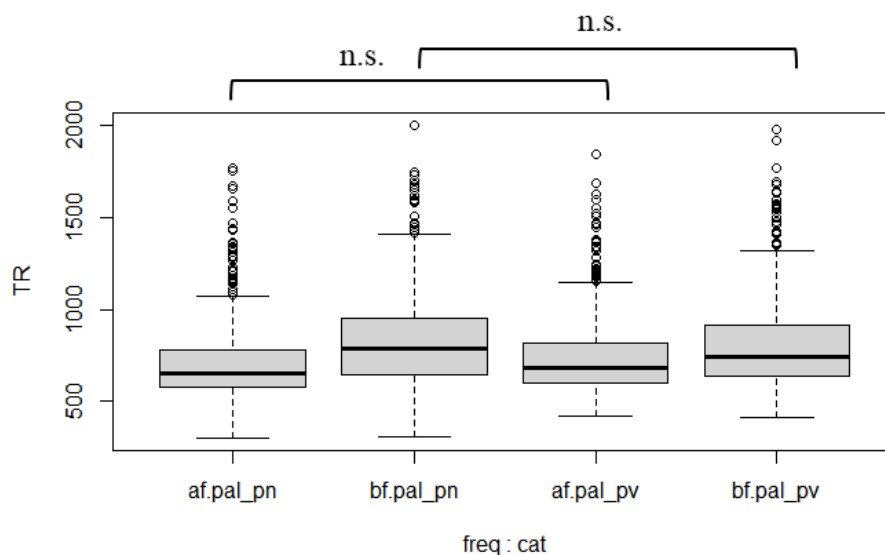


Gráfico 3 - distinção entre médias de Tempo de Resposta (em milissegundos) de categorias. Não se obteve significância estatística no Experimento 1.

No Gráfico 4, aponta-se a diferença significativa entre as condições com *prime* morfologicamente relacionado e com *prime* não relacionado. Via de regra, as palavras com *prime* relacionado tiveram TR médio mais rápido independentemente de sua categoria e de sua frequência ( $\beta=-0.0008$ ,  $SE=0.00018$ ,  $t=-4,46$ ,  $p<0,000$ ). Destaca-se que apesar de uma aparente diferença de TR entre categorias sob o mesmo tipo de *prime* e frequência (e.g., PNAF com *prime* relacionado ter TR médio menor que PVAF) o efeito de categoria não atingiu significância nas análises, como citado anteriormente. Essas diferenças podem ter encontrado alto grau de variabilidade em sujeitos e itens e, por isso, não terem sido significantes.



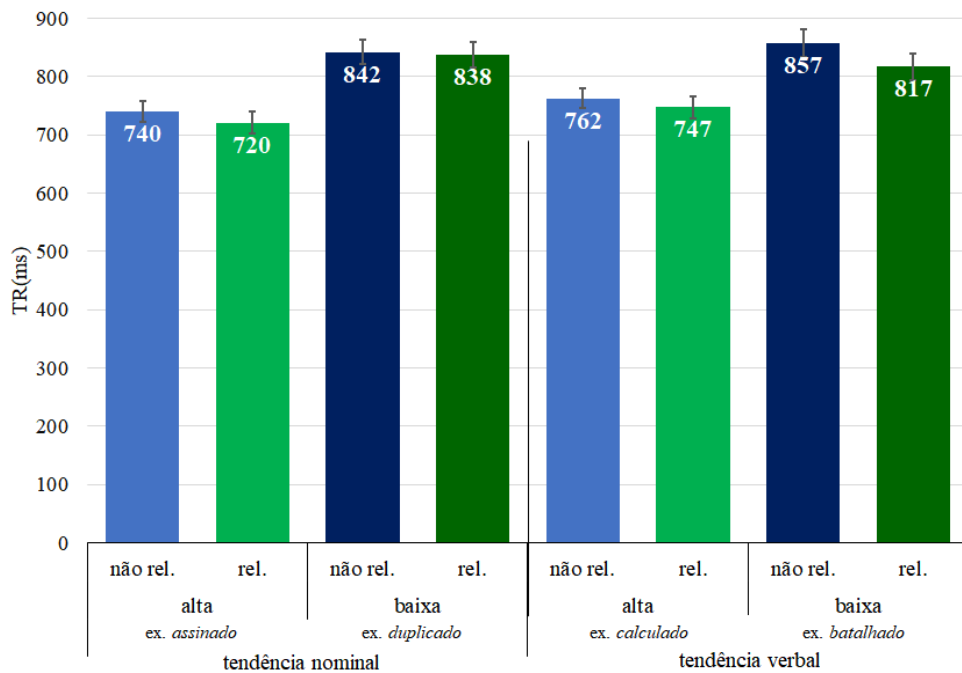


Gráfico 4 - TRs médios, em milissegundos, por tipo de *prime* (morfologicamente relacionado ou não-relacionado), frequência (alta ou baixa) e categoria (tendência nominal ou tendência verbal).

Outro efeito que se observa no Gráfico 5 diz respeito à frequência. Nota-se que palavras de alta frequência são processadas mais rapidamente que palavras de baixa frequência, independentemente do tipo de *prime*. Este dado embasa as informações de frequência superficial registradas no corpus LexPorBr. Uma análise *post-hoc* aponta que palavras alvo de baixa frequência na condição de relação morfológica são respondidas mais rapidamente que as de baixa frequência sem relação com o *prime* ( $\beta=-0.00081$ ,  $SE= 0.0002$ ,  $t=-4,46$ ,  $p<0,000$ ). Também aponta-se que palavras alvo de alta frequência tem TR médio menor na condição de par com relação morfológica que no par não-relacionado ( $\beta=-0.0008$ ,  $SE= 0.00018$ ,  $t=-4,46$ ,  $p<0,000$ ), o que demonstra a independência dos efeitos de tipo de *prime* e frequência.

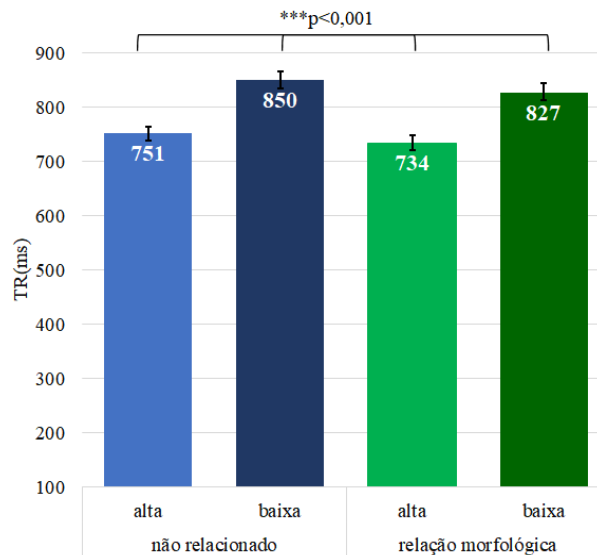


Gráfico 5 - TRs médios, em milissegundos, apenas de palavras reais por frequência (alta ou baixa) e tipo de *prime* (não relacionado ou com relação morfológica).

#### 4.4 Discussão

O Experimento 1 foi um estudo de *priming* encoberto com uma tarefa de decisão lexical. O *priming* encoberto é um paradigma conhecido por ser significativamente facilitador do processamento do alvo apenas quando é morfológicamente relacionado com o *prime* (RASTLE; DAVIS; NEW, 2004; RASTLE; DAVIS, 2008), sendo por este motivo utilizado como forma de evidenciar a ocorrência de processamento morfo-ortográfico. A hipótese foi que palavras no participio preferencialmente nominal de alta frequência não se beneficiariam do *prime* pois estariam listadas na memória em forma inteira, sem acesso à estrutura interna de constituinte, enquanto os participios preferencialmente nominais de baixa frequência e os preferencialmente verbais de baixa e alta frequência seriam segmentados no curso do processamento lexical e, desse modo, beneficiar-se-iam do *prime* morfológicamente relacionado, já que a raiz compartilhada já teria sido ativada pelo *prime*. Tal efeito facilitador do *prime* seria observável no tempo de resposta à tarefa reduzido em comparação com as mesmas condições mas exibidas com um *prime* não relacionado. A hipótese tem por base o sugerido por Leminen *et al.* (2013) quanto ao processamento de palavras flexionadas e derivadas, em que as primeiras são sempre segmentadas e as últimas, apenas quando são de baixa frequência. No presente trabalho, participios que tenham uma leitura preferencialmente nominal seriam um caso de derivação, enquanto os preferencialmente verbais seriam um caso de flexão. Uma evidência favorável à

hipótese seria, também, favorável a um modelo de dupla rota de processamento de palavras morfológicamente complexas.

Algumas considerações podem ser feitas quanto aos resultados. Primeiramente, as expectativas iniciais não se concretizaram, já que todas as condições principais se beneficiaram do *prime* morfológicamente relacionado e não houve interação com a variável categoria. Isso soma evidências à já extensa literatura que aponta a existência de processamento a nível morfológico quando processamos palavras de morfologia complexa, mas que ainda enfrenta resistências de algumas teorias, como a já discutida anteriormente sobre *Naïve Discriminative Learning* (Baayen *et al.*, 2011), por exemplo, ainda que a literatura sobre o processamento lexical em português brasileiro seja bastante menor. Por outro lado, apresenta uma contestação a modelos de dupla rota como de Schreuder e Baayen (1995) em que a alta frequência de palavras semanticamente transparentes de morfologia complexa seria um fator que as levaria ao processamento por forma inteira. Em nível de organização do léxico mental, os dados aqui apresentados podem ser interpretados à luz da Morfologia Distribuída no sentido de que (1) processamento decomposicional foi evidenciado, o que só é possível se o léxico mental armazena as palavras em morfemas e (2) mesmo uma aparente tendência de leitura nominal derivacional das palavras não implicaria em armazenagem em forma inteira como sugerido por Leminen *et al.* (2013), o que dialoga com a proposta da MD de que tanto derivação quanto flexão são processos que ocorrem na derivação sintática. Os dados aqui apresentados também seriam previstos pelo modelo proposto por Pinker (1991) e Ullman *et al.* (1994), segundo o qual palavras regulares (como as aqui utilizadas) são processadas via sistema de regras, por segmentação morfológica, e previstos pelo modelo de decomposição precoce automática (TAFT; FORSTER, 1975; STOCKALL, MARANTZ, 2006) que preveem a ocorrência de decomposição em morfemas sempre. Por fim, os resultados aqui apresentados também dialogam com os achados de Justino e Mota (2019), que apontam para segmentação morfológica de verbos no português brasileiro em primeira conjugação (-AR). Mas algumas ressalvas são apresentadas a seguir que justificam reticência quanto às considerações aqui apresentadas.

Há algumas questões a se levar em conta quanto aos resultados apresentados que indicam cautela quanto às considerações do parágrafo anterior. No preparo do experimento, havia a ciência de que a divisão em grupos de palavras no participio poderia não surtir efeito, já que a medida retirada por pré-testes poderia refletir algo subjetivo dos participantes que o realizaram. Ademais, para que não houvesse um possível efeito de gênero gramatical,

utilizamos apenas palavras terminadas em "-o", o que pode influenciar a leitura das palavras para suas versões verbais, que utilizam o "-o" como parte da desinência modo-temporal, impedindo o acesso a uma potencial entrada lexical em que o particípio é registrado como um adjetivo.

Mesmo se assumir-se que há, de fato, entradas lexicais em que o particípio passado de uma palavra está indexada como adjetivo, não há garantia de que ao ler a palavra isolada é esta a entrada lexical que será resultado do momento de acesso. Por isso, estes resultados ainda não podem ser interpretados à luz do modelo de decomposição precoce automática (TAFT; FORSTER, 1975; STOCKALL; MARANTZ, 2006), já que um potencial efeito de processamento por forma inteira pode ter ficado mascarado. Para assumir essa possibilidade, seria necessário também imaginar que um termo como "circulado" tem duas entradas que representam um caso de polissemia, já que em uma dessas entradas lexicais na memória a palavra figuraria como um verbo flexionado, como quando ocorre em tempo composto (ex.: "tinha circulado"), e na outra, como adjetivo, semelhante a quando ocorre em um sintagma nominal (ex.: "o número circulado"). Isso dialogaria com os achados de Matchin *et al.* (2019) no inglês, ainda que o experimento dos autores tenha se concentrado em detectar efeitos sintático e semântico no córtex quando compararam três condições: uma em que as palavras estavam em lista, uma em que a palavra aparecia como um particípio adjetivo em um sintagma nominal e uma em que a mesma palavra era exibida como a cabeça de uma oração.

Mas se existe mais de uma representação para palavras que, em uma leitura superficial apenas da forma ortográfica, parecem geradas pela mesma morfologia de particípio passado, então o que poderia ativar uma ou outra representação da palavra? Uma variável não testada neste experimento que utilizou leitura de palavras isoladas: o contexto sintático.

Por isso, foi elaborado um segundo experimento - neste trabalho chamado de Experimento 2 - de leitura automonitorada. Como será explicitado no próximo capítulo, a leitura automonitorada é uma técnica em que o participante lê sentenças palavra a palavra (ou bloco a bloco de palavras) ao apertar uma tecla no teclado para ler o segmento seguinte. O tempo de leitura de cada segmento é gravado (tempo entre exibição do segmento e próximo pressionamento de tecla) e constitui a variável dependente deste tipo de teste. Sob esse paradigma, foi possível comparar sentenças em que as palavras no particípio passado eram ora usadas como adjetivos em sintagmas nominais, ora como parte de um sintagma verbal de tempo composto. A ocorrência dentro destes sintagmas e as palavras precedentes restringiam a

possibilidade de categoria do segmento crítico, que é a palavra no particípio. O experimento e seus resultados serão detalhados no capítulo a seguir.

## 5 EXPERIMENTO 2

### 5.1 Metodologia

#### 5.1.1 Hipótese e expectativas

Com base nas considerações sobre os resultados do Experimento 1 que apontaram uma explicação alternativa para a ausência do efeito esperado, explicação esta que remete à potencial existência de mais de uma entrada no léxico mental para uma mesma palavra (uma em forma inteira para uma palavra de morfologia complexa de alta frequência, e uma em que há a representação segmentada desta palavra), elaborou-se a hipótese de que restrições distribucionais do contexto sintático indicariam qual entrada lexical referente a uma mesma forma ortográfica seria utilizada: a decomposicional quando a palavra é resultado de um processo de flexão verbal (*e.g.*, “*tinha convidado*”) ou via forma inteira quando a palavra é resultado de um processo de derivação e sua frequência superficial é alta (*e.g.*, “*o rapaz convidado*”). Tal hipótese está mais em linha com a proposta de dupla rota de Schreuder e Baayen (1995), que propõe a existência das duas vias no curso do processamento lexical da palavra. No entanto, diferentemente do que é proposto pelos autores, acrescentamos a variável “contexto sintático” como preditora da via preferível para o acesso lexical da palavra. Na realidade, as palavras apenas são as mesmas ao considerar-se um olhar superficial sobre sua forma completa, já que de acordo com a presente hipótese, se a palavra recai em um contexto que a restringe a um adjetivo (*e.g.*, “*o rapaz convidado*”) o processamento buscaria pela entrada referente a “*convidado*” em que ela está indexada como um adjetivo, uma qualidade atribuída a um ser animado, e que seria armazenada em forma inteira apenas quando a uma palavra for de alta frequência. Já no caso de “*tinha convidado*”, por exemplo, por se tratar de uma flexão verbal, conforme a sugestão de Leminen *et al.* (2013), a frequência não afetaria a forma como a palavra é armazenada, possuindo desta forma uma entrada lexical em que sua representação se dá em constituintes (uma raiz que carrega o significado referente a convidar, “*convid-*” mais o afixo que atribui o valor modo temporal relativo ao resultado de um evento, “*-(a)do*”).

A expectativa foi de que palavras no participio teriam tempo de leitura diferentes nas sentenças (tempo maior em palavras de baixa frequência<sup>8</sup>) em que ocorrem como parte da

---

<sup>8</sup> Um estudo com três experimentos de Lau, Rozanova e Phillips (2007) replicou o efeito de frequência superficial, encontrado em tarefas de decisão lexical, em experimentos de leitura automonitorada. No entanto, o efeito foi mais significativo no primeiro experimento, em que o contexto não permitia prever o traço de número do segmento crítico (variável independente investigada). Como os sintagmas e sentenças do presente trabalho foram montadas para reduzir efeitos de processamento *top-down*, esperava-se encontrar efeito de frequência superficial também no presente experimento.

perífrase verbal de tempo composto, nas quais são um caso de particípio verbal, enquanto as mesmas palavras nas sentenças em que seriam particípios adjetivos teriam tempos de leitura médios diferentes devido a frequência, mas com intervalo maior. Esta previsão parte da suposição de que a existência de um processamento decomposicional envolve o acréscimo de etapas no curso do processamento lexical, o que refletiria mais tempo de processamento (*e.g.*, Virpioja *et al.*, 2011 encontraram maior tempo de processamento em estudos comportamentais associado a palavras flexionadas em finlandês que para palavras monomorfêmicas). Assim, particípios adjetivos de alta frequência seriam mais rápidos que os de baixa frequência, já que não seriam decompostos em morfemas e teriam o benefício da alta frequência, com uma diferença mais ampla que na comparação de particípios verbais de alta com os de baixa frequência, já que os particípios verbais sempre seriam decompostos e os de alta frequência teriam o custo cognitivo (refletido em tempo de leitura) pela segmentação, mas com o efeito facilitador da alta frequência. Caso não haja efeito de categoria e/ou de frequência, uma explicação alternativa é que casos de particípio passado de primeira conjugação são processados sempre por via única e decomposicional.

### 5.1.2 Desenho experimental e materiais

O Experimento 2 utilizou o paradigma de leitura automonitorada com tarefas de compreensão acerca da sentença, cujo recorte também é a investigação sobre processamento lexical de palavras no particípio passado de primeira conjugação ("-ar") do português brasileiro. O participante leu as sentenças palavra a palavra no decorrer do experimento, passando para a palavra seguinte ao pressionar a barra de espaço do teclado. Ressalta-se que há outros testes em que a leitura é feita de bloco em bloco de palavras, diferentemente deste teste em que as palavras apareciam isoladamente uma após a outra. As tarefas de compreensão aplicadas tiveram por objetivo apenas a manutenção da atenção dos participantes.

Diferentemente do Experimento 1, apenas palavras da condição PN foram utilizadas. Como o intuito foi observar se uma mesma palavra no particípio passado poderia ter uma entrada no léxico mental em que está representada como um adjetivo e uma outra como verbo, apenas palavras desta condição possuem igual potencial para ambas as categorias. Ou seja, palavras classificadas como preferencialmente verbais não necessariamente tem o mesmo potencial para ocorrer naturalmente como adjetivo que as palavras classificadas como preferencialmente nominais possuem para ocorrer naturalmente como verbos, sendo o potencial mais alto para este último caso.

Dessa forma, os segmentos críticos das sentenças apresentadas foram as palavras listadas nas condições PNAF e PNBF, com manutenção dos mesmos controles de frequência superficial e tamanho da palavra (em número de segmentos ortográficos e sílabas). Uma mesma palavra aparecia em um contexto sintático em que é passível de interpretação como adjetivo (ex.: “o rapaz **atrasado**”, n=20) e em um contexto que restringia sua leitura a classe verbal em uma perífrase de tempo composto (ex.: “**tinha atrasado**”, n=20). Para que houvesse um mínimo de interferência contextual das palavras que precediam o segmento crítico por predição do falante (efeito *top-down*), o início das sentenças utilizou palavras genéricas que não gerassem muita expectativa sobre o que apareceria a seguir (exemplos na Tabela 3). Nas sentenças em que as palavras eram consideradas adjetivos de palavras animadas, o sintagma nominal tinha como cabeça do sintagma nomes genéricos animados (“garoto”, “cara” ou “rapaz”, como no exemplo “o rapaz atrasado...”, n=20), enquanto nos sintagmas com adjetivos de palavras inanimadas (n=20) foram utilizados termos genéricos inanimados como “negócio” (ex.: “o negócio duplicado...”), “troço” (ex.: “o troço roubado...”) ou “lugar” (ex.: “o lugar fechado...”). Para que não houvesse efeito de familiarização com a tarefa, houve sentenças distratoras que possuíam sintagmas nominais semelhantes aos das sentenças alvo (n=20) e que utilizaram os verbos auxiliares das perífrases de tempo composto com um argumento interno não-verbal (n=20), atingindo-se a proporção 50%/50% de distribuição de sentenças alvo (n=40) e distratoras (n=40).

Dois grupos experimentais foram criados para que nenhum participante visse a mesma palavra mais de uma vez. No grupo 1 (G1), a palavra *x* aparecia como parte de um sintagma nominal enquanto adjetivo na sentença, enquanto no grupo 2 (G2) a palavra *x* aparecia como parte de uma perífrase verbal na sentença.

<b>Lista</b>	<b>Sentença com palavra como adjetivo (segmento crítico <u>sublinhado</u>)</b>	<b>Sentença com palavra como verbo (segmento crítico <u>sublinhado</u>)</b>
<i>PNAF</i>	O rapaz <u>atrasado</u> estava ali fora.	Eu tinha <u>atrasado</u> ela ontem.
<i>PNBF</i>	O cara <u>tatuado</u> está ali.	Ela tinha <u>tatuado</u> isso tem tempo.
<i>Sentença distratora</i>	O troço <u>branco</u> estava aqui.	Eu tenho <u>calafrios</u> o tempo todo.

Tabela 3 - Exemplos de estímulos por condição do segmento crítico e por classe da palavra do Experimento 2. PNAF: participípio preferencialmente nominal de alta frequência. PNBF: participípio preferencialmente nominal de baixa frequência.



Os materiais correspondentes aos segmentos críticos são os mesmos utilizados no Experimento 1, nas condições de PNAF e PNBF. A função de randomização da sequência das sentenças exibidas da Plataforma PCIBEX foi utilizada para para que cada participante realizasse o experimento com ordem de sentenças diferentes.

As variáveis dependentes deste experimento são o tempo de leitura (TL) e a acurácia. Cientes de que há um efeito chamado *spill-over* em testes de leitura automonitorada, em que o tempo de processamento do segmento crítico é refletido no tempo de leitura da palavra seguinte, nenhuma sentença teve o segmento crítico como palavra final e a segunda palavra era controlada. Para a condição de sentença com particípio adjetivo, a palavra seguinte sempre foi o verbo de estar na 3ª pessoa singular, no presente ou passado imperfeito. Nas condições de sentenças com particípio verbal, a palavra seguinte sempre foi uma palavra funcional, como um pronome pessoal ou demonstrativo, ou preposição.

### 5.1.3 Participantes

Os participantes foram estudantes de graduação em Letras ou de pós-graduações na área, com faixa etária entre 18 e 29 anos. Houve um total de 42 participantes. Apesar da intenção de manter a uniformidade dos grupos de participantes a fim de garantir a comparabilidade dos resultados dos experimentos, a medida de incluir estudantes de pós-graduação foi necessária para que o experimento pudesse ser finalizado a tempo para a produção desta dissertação. A pesquisa foi aprovada pela Comissão de Ética da UFRJ sob o CAAE 50670421.3.0000.5699, parecer nº 5.185.244.

### 5.1.4 Procedimentos

O Experimento 2 tratou-se de um teste de leitura automonitorada de sentenças seguida de uma tarefa de compreensão. Assim como o Experimento 1, o teste foi realizado de forma remota através da plataforma PCIBEX cujo acesso se deu por um *link*. Após aceitar o Termo de Consentimento Livre e Esclarecido (TCLE), havia o preenchimento de algumas informações pessoais e o participante passava, então, a uma tela que continha um texto explicando o funcionamento do teste e instruções de como realizá-lo. O participante era instruído a pressionar a barra de espaço para que as palavras da sentença fossem exibidas no centro da tela do computador. Para cada palavra da sentença, havia um espaço sublinhado correspondente a ela (“\_\_\_\_\_”) que era substituído pela palavra quando a tecla era pressionada. Após a tecla

ser pressionada novamente, esta palavra era retirada da tela, substituída pelo espaço sublinhado e a palavra seguinte era exibida, procedimento que decorria até o final da sentença. Após o término da leitura da sentença, surgia uma pergunta de compreensão sobre a sentença lida, em que o participante deveria responder pressionando a tecla “A” se a resposta fosse sim, ou “L” se a resposta fosse não, levando-se em consideração o conteúdo da sentença. Por exemplo, se o participante leu a sentença “O negócio manchado esteve ali”, responderia a pergunta “O negócio está limpo?”, cuja resposta esperada é NÃO (tecla “L”). Após a tela que continha as instruções, passava-se a uma etapa de treinamento que continha 4 *trials*. Em seguida, uma tela era exibida informando o início do teste principal, que o participante começaria após pressionar a barra de espaço. Para que se reduzisse o efeito de fadiga pelo tempo de realização do experimento, ele foi dividido em dois blocos com uma pausa entre eles. Cada bloco consistiu de 20 *trials*, o que totalizou 40 *trials* por participante.

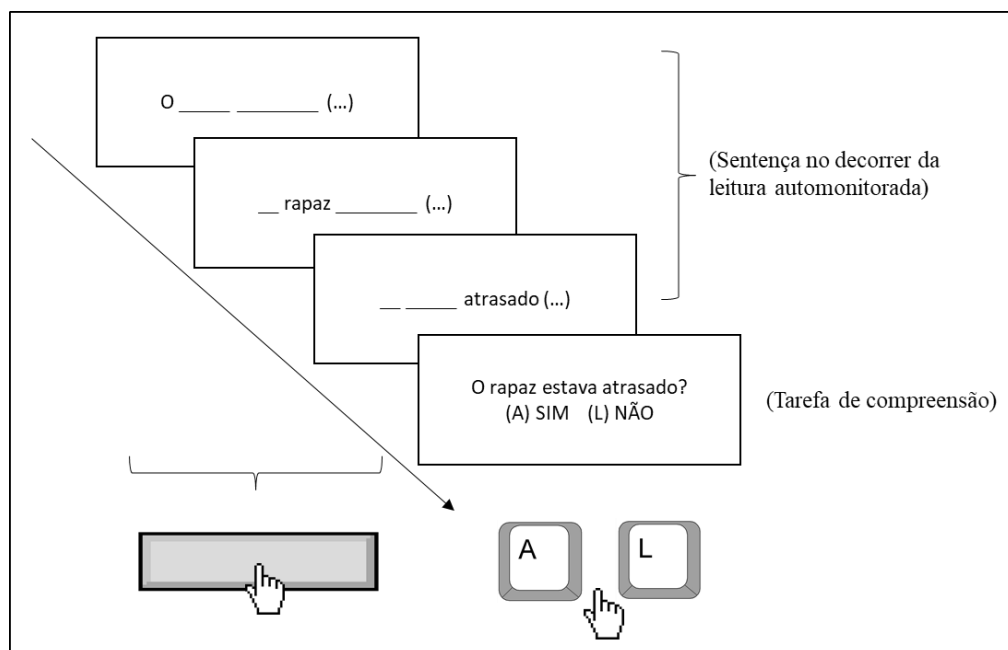


Figura 9 - Figura ilustrativa do teste de leitura automonitorada seguido da tarefa de compreensão.

## 5.2 Análise estatística

Da mesma forma que o Experimento 1, a análise estatística foi rodada no Rstudio (versão 1.2.5033; R Core Team, 2013). O modelo adotado também foi o modelo de regressão linear de efeitos mistos, que inclui em sua análise comportamento aleatório dos dados em variáveis independentes, pela aplicação da função lmer do pacote lme4 (BATES *et al.*, 2015).

O Experimento 2 teve duas variáveis dependentes: o tempo de leitura (TL) de cada segmento da sentença e a taxa de acurácia da tarefa de compreensão das sentenças. A taxa de acurácia foi usada para estabelecer um limiar abaixo do qual os participantes seriam excluídos do conjunto de dados, por uma potencial falta de atenção ao teste. O TL foi também utilizado na exclusão de dados, já que tempos muito curtos ou muito longos poderiam indicar desatenção ao teste, mas ao contrário da taxa de acurácia, a medida dependente de TL foi o dado principal analisado neste experimento. As variáveis independentes foram de categoria da palavra (PN ou PV) e frequência superficial (alta ou baixa); a depender do contexto sentencial, um mesmo item aparecia como adjetivo ou como verbo. Como o segmento posterior ao segmento crítico das sentenças com verbo e com adjetivo não são exatamente comparáveis entre si, rodamos primeiro modelos separados para participios adjetivos e participios verbais, e no final, apenas para o segmento crítico, um modelo com todas as variáveis. E, por fim, as variáveis aleatórias foram participante e item. Como há a possibilidade de um efeito de *spillover*, o TL tanto do segmento crítico quanto do primeiro segmento pós-crítico foram incluídos na análise, em modelos separados. *Outliers* foram removidos a partir de um cálculo que considerou como limite máximo a média do tempo de leitura somada de três vezes o desvio padrão. Foram feitas duas análises: uma incluindo apenas os *trials* com respostas corretas e uma incluindo todos os *trials*, incluindo respostas corretas e erradas. Consideramos que a pergunta de compreensão, mesmo que uma medida importante de verificação de atenção, não garante que uma resposta errada implique em uma não leitura da sentença anterior, justamente por serem consecutivos. Portanto, julgamos relevantes mostrar ambas as análises. Ademais, excluímos 31 perguntas de compreensão da análise de acurácia, de um total de 80 itens experimentais, que após análise *post-hoc* concluímos serem ambíguas em relação a sua resposta. Esses itens foram incluídos na análise junto com os *trials* de resposta correta. Após rodar-se um teste de normalidade de Jarque-Bera, os dados foram transformados utilizando com uma função lambda de transformação logarítmica pela função BoxCox do pacote *fpp* (HYNDMAN; KHANDAKAR, 2008), mas, assim como no Experimento 1, os TLs não transformados foram mantidos nos gráficos para fins de compreensão das médias. A significância dos modelos testados foi calculada através de comparações aninhadas, e os p-valores foram calculados através do pacote *lmerTest* (KUZMETSOVA; BROCKHOFF; CHRISTENSEN, 2017).

## 5.3 Resultados

### 5.3.1 Taxa de acurácia

A taxa média de acurácia das sentenças alvo (isto é, subtraídas as sentenças distratoras e aquelas cujas perguntas julgamos terem duas respostas possíveis) foi de 90,25% (DP: 4,96). Já a média geral, com os distratores incluídos, foi de 87,27% (DP: 5,54%). Estipulando uma margem de média menos 2\*desvio padrão (76,19%), nenhum dos participantes foi removido por baixa acurácia. Consideramos uma tolerância maior na acurácia pelo fato de nenhum participante estar com desempenho de chance, e pelo aparente grau de dificuldade ou ambiguidade possível da tarefa, sendo que o maior número possível de inclusão de participantes contribui para a robustez do modelo estatístico.

### 5.3.2 Tempo de leitura

Abaixo apresentam-se os dados referentes ao tempo de leitura (TL) dos segmentos analisados neste experimento. Além das remoções mencionadas anteriormente, houve também filtragem para que *outliers* não compusessem os conjuntos de dados analisados, tendo-se em conta um limite máximo para os TLs a soma das médias de TL dos respectivos segmentos (segmento crítico ou primeiro segmento pós-crítico) e de três vezes o desvio padrão de cada média. A opção por este cálculo se deu no intuito de preservar o máximo possível de dados, dando ampla margem no limite máximo através do triplo do desvio padrão, ao mesmo tempo corrigindo o *skewness* para direita a fim de aproximar uma distribuição normal dos dados.

Uma série de normalizações foram rodadas sobre os conjuntos de dados que entraram nos modelos Adjetivo e Verbo. Esses conjuntos foram analisados de duas formas, incluindo todos os trials e incluindo apenas os trials com acertos. O conjunto de *apenas acertos* se refere aos dados restritos a sentenças cuja resposta para a tarefa correspondente estava correta. Os resultados do teste de Jarque-Bera apontaram que a normalização foi atingida apenas nos subconjuntos de dados de segmento crítico transformados por função lambda, com todas as respostas (para adjetivos,  $X^2=3,5903$  e  $p=0,1661$ , para verbos,  $X^2=2,035$  e  $p=0,3615$ ) e com apenas acertos (para adjetivos,  $X^2=3,1423$  e  $p=0,2078$ , para verbos,  $X^2=2,1478$  e  $p=0,3417$ ). As mesmas análises foram rodadas no primeiro segmento pós-crítico devido a um possível efeito de *spillover*, mas estes não atingiram o limiar necessário de normalização mesmo através

da transformação lambda pela função BoxCox, apenas se aproximaram mais da normalidade (para adjetivos,  $X^2=36,99$ ,  $p<0,001$ , para verbos  $27,263$ ,  $p<0,001$ ).

Uma comparação aninhada mostrou que não há diferença significativa entre modelos nulos (TL.lambda  $\sim 1 + (1|participante) + (1|item)$ ) e modelos em que a frequência é inclusa na fórmula como um efeito fixo (TL.lambda  $\sim$  frequência + (1|participante) + (1|item)) para o segmento crítico. Os resultados estão resumidos na Tabela 4. Para fins de comparação, estão incluídos também os resultados de modelos que incluem todas as respostas (inclusive aquelas em que os participantes erraram a pergunta de compreensão). Da mesma forma que os modelos com respostas apenas corretas, os dados nesses modelos atingiram normalidade com TLs transformados para o segmento crítico, mas não para o segmento da posição posterior.

<b>Categoria</b>	<b>Segmento</b>	<b>Conjuntos</b>	<b>X<sup>2</sup></b>	<b>p-valor</b>
<b>Adjetivo</b>	Segmento crítico	Todas as respostas	$X^2=1,2075$	$p=0,2718$
		Apenas acertos	$X^2=1,2171$	$p=0,2699$
	Primeiro segmento pós-crítico	Todas as respostas	$X^2=0,0765$	$p=0,7821$
		Apenas acertos	$X^2=0,019$	$p=0,8905$
<b>Verbo</b>	Segmento crítico	Todas as respostas	$X^2=1,3586$	$p=0,2438$
		Apenas acertos	$X^2=2,6548$	$p=0,1032$
	Primeiro segmento pós-crítico	Todas as respostas	$X^2=0,003$	$p=0,9869$
		Apenas acertos	$X^2=0,4128$	$p=0,5206$

Tabela 4 - comparações aninhadas entre modelo nulo e modelo mais complexo de efeitos mistos para TLs para o segmento crítico. Para que haja significância estatística na comparação entre os modelos, é necessário que o p-valor esteja abaixo de 0,05 ( $p<0,05$ ).

Desse modo, a comparação aninhada mostra que a frequência não foi um efeito determinante na análise dos resultados se considerarmos contextos de adjetivos e verbos separados em nenhum dos conjuntos de dados, nem com todas as respostas (para a categoria adjetivo,  $X^2=1,2075$  e  $p=0,2718$ , para a categoria verbo,  $X^2=1,3586$  e  $p=0,2438$ ) nem para o subconjunto apenas com acertos (para a categoria adjetivo,  $X^2=1,2171$  e  $p=0,2699$ , para a categoria verbo,  $X^2=2,6548$  e  $p=0,1032$ ).

Para detectar efeitos de categoria, um conjunto de dados com todos os TLs de ambas as categorias testadas foi testado para normalização. Novamente, atingiu-se normalidade com a transformação lambda para o segmento crítico ( $X^2=4,9881$ ,  $p=0,08257$ ). Consideramos apenas o segmento crítico, pois entre as sentenças com verbos e adjetivos, esse segmento permite comparabilidade. Os resultados das comparações aninhados não acusaram nenhuma significância para os modelos considerados: frequência\*categoria ( $X^2=3,9292$ ,  $p=0,2692$ ), frequência+categoria ( $X^2=3,8811$ ,  $p=0,1436$ ), frequência ( $X^2=3,4013$ ,  $p=0,06515$ ) e categoria ( $X^2=0,5234$ ,  $p=0,4694$ ). Porém, notamos que o modelo com frequência aproxima o limiar de significância.

Isso significa que não há uma diferença significativa entre o TL médio de leitura do segmento crítico quando este é de baixa frequência (adjetivos, média=554,31 e DP=274,58ms, para verbos, média=567,83 e DP=278,66ms) com o TL médio do mesmo tipo de segmento quando este é de alta frequência (adjetivos, média=518,78ms e DP=228,65ms, para verbos, média=515,45 e DP=210,17ms). Na Tabela 5, apresentam-se as médias por segmento, categoria e frequência. Observando as médias, pode ser vista uma tendência de TLs mais rápidos para frequências mais altas, embora essa tendência não se confirme na análise estatística. Análises com tempos brutos, acusam efeitos de frequência (marginal, para adjetivos:  $X^2=2,8863$ ,  $p=0,08934$ , e para verbos:  $X^2=6,1956$ ,  $p=0,01281$ , e para o modelo juntando as duas categorias  $X^2=8,2249$ ,  $p=0,004$ ), lembrando que essas análises carecem de confiabilidade já que os TLs brutos não atingiram a normalização. Análise de tempos brutos para categoria acusam uma ausência de efeito de categoria ( $X^2=0,1078$ ,  $p=0,7427$ ) ou de interação entre frequência e categoria ( $X^2=0,1529$ ,  $p=0,6958$ ).

<b>Categoria</b>	<b>Frequência</b>	<b>Segmento</b>	<b>Média (em ms)</b>	<b>Desvio padrão (em ms)</b>
<b>Adjetivo</b>	AF	Segmento crítico	518,7802	228,6563
	BF	Segmento crítico	554,3130	274,5897
	AF	Segmento pós-crítico	485,5995	163,5062
	BF	Segmento pós-crítico	485,9296	167,5920
<b>Verbo</b>	AF	Segmento crítico	515,4587	210,1756
	BF	Segmento crítico	567,8372	278,6619
	AF	Segmento pós-crítico	482,7167	170,3303
	BF	Segmento pós-crítico	489,6589	154,7412

Tabela 5 - Médias de tempo leitura de sentenças cujas tarefas foram respondidas corretamente, com remoção de *outliers*.

Há um contraste entre o resultado apresentado aqui com aquele exibido no Experimento 1 e também reportado na literatura, que apontam para um efeito significativo de frequência superficial da palavra no dado comportamental que se obtém a partir de tais palavras.

Uma comparação visual dos TLs médios brutos nos gráficos 6 e 7 mostram que enquanto há uma aparente tendência de efeito de frequência tanto nos TLs de participios adjetivos quanto de participios verbais (um pouco mais expressivo), mesmo que não apontado na análise estatística, os TLs médios do primeiro segmento pós-crítico estão em um nível muito semelhante. Também podemos observar que não há efeito *spillover*, o primeiro segmento pós-crítico não tendo entre os TLs referentes às condições experimentais significância estatística.

Por fim, reportamos o resultado de uma análise adicional de um possível efeito dos nomes animados e inanimados para a condição de nome. Essa variação controlada nos estímulos não impactou o processamento (TR.lamba~animacidade vs. modelo nulo:  $X^2=0,0109$ ,  $p=0,9168$ , TR.lamba~animacidade\*frequência vs. modelo nulo:  $X^2=3,7825$ ,  $p=0,2859$  (segmento crítico)).

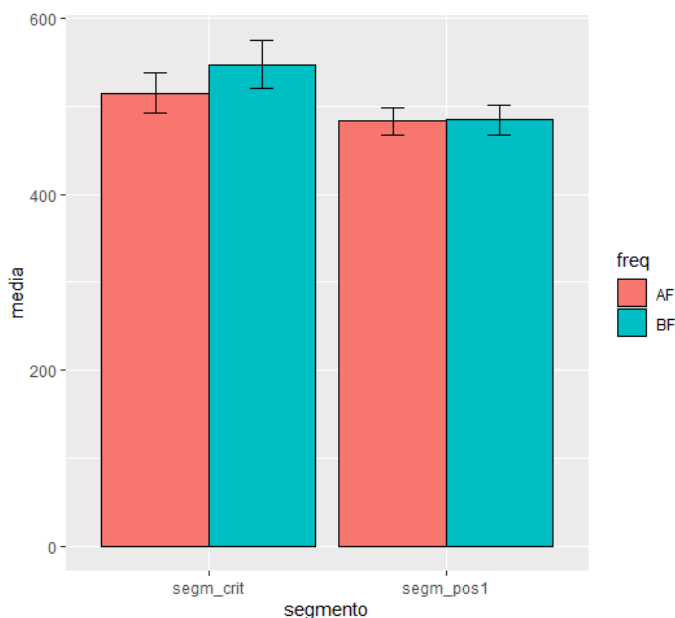


Gráfico 6 - médias de tempo de leitura (eixo y) do segmento crítico (segm\_crit) e do primeiro segmento pós-crítico (segm\_pos1) em milissegundos no subconjunto de dados de sentenças com participio adjetivo, cujas tarefas de compreensão foram respondidas corretamente.

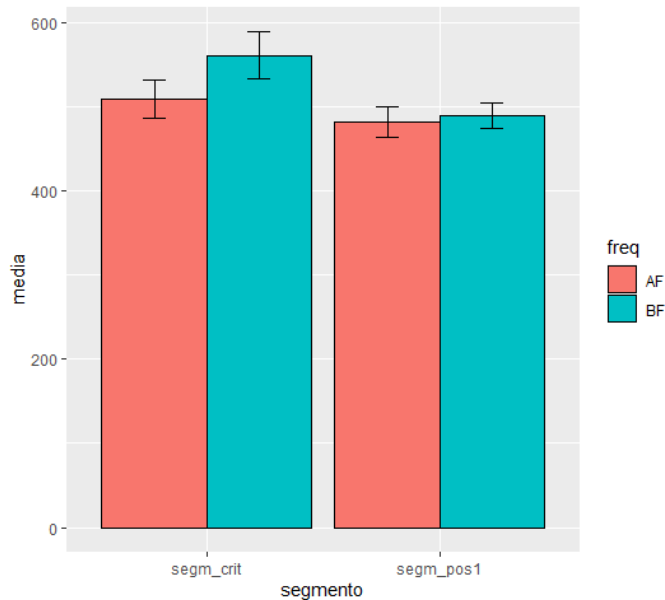


Gráfico 7 - médias de tempo de leitura (eixo y) do segmento crítico (segm\_crit) e do primeiro segmento pós-crítico (segm\_pos1) em milissegundos no subconjunto de dados de sentenças com particípio verbal, cujas tarefas de compreensão foram respondidas corretamente.

#### 5.4 Discussão

No Experimento 2, testou-se a hipótese de que o contexto sintático é determinante na ativação de duas entradas para palavras morfologicamente ambíguas no particípio passado, que apesar de compartilharem a mesma ortografia, possuem classes diferentes, uma sendo adjetivo e a outra, verbo. Parte da hipótese é que adjetivos participiais de alta frequência seriam processados via forma inteira e os de baixa frequência por decomposição em morfemas, enquanto verbos no particípio passado seriam segmentados em morfemas independentemente de sua frequência. Tal previsão parte de uma sugestão de Leminen *et al.* (2013) sobre o finlandês que palavras derivadas podem ser armazenadas em forma inteira (neste experimento, representados pelos particípios adjetivos) enquanto palavras flexionadas (neste experimento, representadas pelos particípios verbais) seriam sempre segmentados em morfema. A confirmação da hipótese ocorreria caso o TL de particípios adjetivos de alta frequência fosse menor que os de baixa frequência, já que a segmentação implica em mais estágios de processamento, enquanto os particípios verbais teriam tempos de leitura com intervalo com diferenças menos expressiva entre cada condição de frequência, já que haveria benefício pelo efeito de alta frequência, mas também haveria custo em ambas as condições devido aos estágios de processamento decomposicional. As sentenças com particípios verbais e as com particípios



adjetivos utilizaram a mesma lista de palavras, que no Experimento 1 compôs a lista “particípio preferencialmente nominal”. A interpretação do particípio como adjetivo ou verbo foi garantida incluindo-se a palavra em um sintagma nominal (ex.: “o rapaz *atrasado*”) quando a intenção era que ela fosse lida como um adjetivo, e em uma perífrase verbal de tempo composto (ex.: “*tinha atrasado*”) quando intencionava-se que ela fosse lida como verbo. Para que um mesmo participante não lesse a mesma palavra duas vezes, houve dois grupos experimentais, um em que um item aparecia como parte de um sintagma nominal, e outro em que o mesmo item aparecia como parte de uma perífrase verbal.

Enquanto o Experimento 1 obteve um efeito de frequência superficial com as mesmas palavras sobre os TRs, o Experimento 2 não exibiu qualquer efeito desta variável na análise estatística. Já similarmente ao Experimento 1, não houve efeito de interação com categoria ou comparação entre categorias. Há duas formas de se interpretar os resultados, que serão apresentadas a seguir, a primeira menos provável de representar o produto do experimento aqui reportado que a segunda. Ressalta-se que não houve comprovação da hipótese. Parte-se, então, para as explicações alternativas.

Primeiramente, pode-se interpretar os resultados como uma evidência a favor do modelo de via única de decomposição automática (TAFT; FORSTER, 1975; STOCKALL; MARANTZ, 2006). Como este modelo não prevê diferenças nas etapas de processamento que sejam dependentes da frequência superficial da palavra, não é contraditório que não haja diferença de efeito de frequência entre categorias, sendo neste experimento a similaridade pela ausência de efeito significativo. Seria o caso de afirmar que os resultados do teste de leitura automonitorada também podem ser interpretados sob a abordagem de modelos não-decomposicionais de processamento lexical, se o Experimento 1 já não tivesse exibido efeitos de *priming* encoberto que, ao nosso ver e com base na literatura, comprova a decomposição morfológica. Ademais, em um modelo de dupla rota como o de Pinker (1991) também não se esperaria armazenagem em forma inteira de palavras formadas por processo flexional regularmente, se considera-se que tanto o particípio adjetivo quanto o verbal são oriundos de uma flexão.

A segunda possível explicação para a diferença entre o resultado deste experimento com o anterior, de decisão lexical, quanto ao impacto da variável frequência superficial refere-se a fatores não controlados. Um deles é que o benefício facilitador de reconhecimento de palavras de alta frequência seja mascarado quando em contexto sentencial, já que há necessidade de concatenações sintáticas e semânticas com as palavras que ocorrem anteriormente. Pyllkkänen

(2019) reportou um efeito do tipo comparando sintagmas como “*red boat*”/”barco vermelho” e “*red canoe*”/”canoa vermelha”, em que “barco”, por ser mais genérico, elicitava um engajamento cognitivo elevado na composição semântica com o adjetivo, que implica em uma maior especificidade do barco entre os diversos tipos de embarcação, enquanto “canoa”, por si só já específico, não elicitava o mesmo engajamento ao se combinar com o adjetivo “vermelho”. Como os sintagmas nominais na condição de participio adjetivo usaram apenas palavras genéricas, um efeito como o reportado por Pykkänen (2020) pode ter compensado o efeito facilitador da alta frequência. Além disso, a tarefa de decisão lexical envolve principalmente o reconhecimento da palavra sem que, obrigatoriamente, seu significado seja computado até o momento de realização da tarefa, ou mesmo demais propriedades das palavras, como o fato de verbos requererem seleção argumental. Logo, o TR oriundo de tal tarefa pode refletir apenas estágios iniciais e os demais estágios do processamento não estarem refletidos nos TRs, ainda que ocorram. A ocorrência de facilitação por *priming* encoberto, por exemplo, sugeriria que os TRs de decisão lexical refletem o processamento ao menos até o estágio de decomposição em morfemas. Isto ressalta o impacto que a natureza da tarefa pode ter sobre o resultado e a interpretação dos dados e, se for uma das explicações corretas (ou a única) sobre as tarefas aqui utilizadas, também se encaixa com a interpretação apresentada no parágrafo anterior e com o fato de o efeito de frequência ser mais significativo no Experimento 1, em que o reconhecimento de uma palavra como uma palavra é muito facilitado por sua alta frequência. Por fim, um estranhamento pode também ter provocado a ausência de efeitos de frequência na condição de participios adjetivos, já que metade dos estímulos eram considerados como adjetivos inanimados e foram colocados em sintagmas nominais com substantivos genéricos “negócio” e “troço”. Assim, torna-se viável um efeito de estranhamento pelo substantivo que encabeçava o sintagma nominal que pode não ocorrer, por exemplo, com as palavras usadas para os adjetivos de nomes animados (“rapaz”, “garoto” e “cara”).

Mas uma interpretação qualitativa e visual dos gráficos levanta uma outra possibilidade, que também se desencontra com a hipótese original. Mesmo que a frequência como variável preditora da variável dependente de TL não tenha alcançado significância estatística nas análises rodadas, há uma aparente tendência de efeito de frequência. A justificativa para que não se tenha atingido significância estatística pode residir na necessidade de mais dados, o que implica a necessidade de mais participantes, e não na inexistência de um efeito de frequência em contexto sentencial. Um ponto que parece favorecer esta suposição é o fato de o modelo de análise de efeitos mistos com frequência como variável independente sem que os TLs tenham

sido transformados tenha atingido significância estatística mesmo que os mesmos dados não tenham atingido a normalização necessária para que fossem confiáveis, ainda que isso tenha ocorrido apenas para a categoria de verbos, e no modelo juntando as duas categorias, como apresentado na seção de resultados, no subconjunto de dados apenas com TLs de sentenças respondidas corretamente. Seja como for, os resultados aqui presentes contestam a hipótese que surgiu como explicação alternativa para os resultados do Experimento 1. Mesmo se considerarmos como verdadeira a tendência de efeito de frequência, ela ocorreria em ambas as categorias (particípio adjetivo e particípio verbal), o que contradiz a expectativa original. Também contraditório à expectativa original é o efeito de frequência aparentar ser mais expressivo na categoria de verbos, quando a expectativa é que verbos sempre seriam segmentados em morfemas. A comparabilidade entre categorias de palavra também acabaria prejudicada, já que um segundo olhar mostra que o contexto sintático (adjunto adnominal em um, verbo em outro) pode afetar o tempo de processamento da palavra.

Um outro olhar deve ser direcionado para aquilo que não foi testado neste experimento e que poderia demonstrar a validade do modelo de dupla-rota no português brasileiro. Como a regularidade da formação de palavras (PINKER, 1991; ULLMAN *et al.*, 1997) foi apontada como variável que determina qual via de processamento será ativada no curso do processamento lexical de palavras de morfologia complexa no inglês, não há como dizer com base nos resultados dos experimentos desta dissertação que tal modelo de dupla rota foi refutado. Já o modelo de dupla rota de Schreuder e Baayen (1995) prevê que mesmo palavras regulares, quando em alta frequência, podem ser armazenadas em forma inteira, o que seria divergente dos dados aqui apresentados. Porém os mesmos autores reportaram, em um outro estudo, os efeitos previstos em seu modelo apenas em casos de *substantivos* regulares flexionados em holandês (BAAYEN; DIJKSTRA; SCHREUDER, 1997), quando os experimentos desta dissertação utilizaram verbos e adjetivos, mesmo que esta última categoria esteja aberta ao debate se são, de fato, adjetivos. Dessa forma, ainda que os resultados dos Experimentos 1 e 2 dialoguem com o modelo de decomposição precoce automática, eles não se contrapõem com modelos de dupla rota.

E quanto ao que dizer a respeito da organização do léxico mental? A hipótese de que há mais de uma entrada lexical para uma mesma palavra não se sustenta frente aos resultados aqui apresentados. Tanto no Experimento 1 quanto no Experimento 2 não há como constatar uma diferença entre o processamento de particípios nominais e verbais: no Experimento 1, a comparação entre categorias "preferencialmente nominal" e "preferencialmente verbal"

(segundo o pré-teste) não alcançou significância estatística; e no Experimento 2, houve ausência de efeito de frequência na análise tanto nas sentenças em que os "particípios preferencialmente nominais" eram usados como adjetivos quanto nas que eles eram usados como verbos. Se é seguro afirmar que há decomposição morfológica tanto em uma categoria quanto em outra, averiguar a existência de entradas lexicais na memória diferentes para cada classe de uma palavra formalmente semelhante (ex.: uma para "roubado" como adjetivo, e uma para "roubado" como verbo) não foi possível pelas medidas aqui utilizadas de TR e TL, de forma que a única inferência viável sobre o léxico mental com os dados dos experimentos é a existência de representação de constituintes na memória referente às raízes e afixos das palavras que compuseram a lista de estímulos dos testes.

Sobre a organização do léxico mental, tal previsão é consonante com a proposta da Morfologia Distribuída (HALLE; MARANTZ, 1992) que postula um léxico mental com representação em constituintes e uma relação entre flexão e derivação que, apesar de não ser de sinonímia, é de similaridade por ambas serem parte da derivação sintática. Mas ressalta-se que é necessária cautela na interpretação dos resultados reportados à luz deste modelo, já que só a primeira conjugação foi testada e apenas palavras regulares e o estudo de Justino e Mota (2019) atestou efeitos de classe de conjugação que podem indicar armazenagem em forma inteira de palavras da segunda e terceira conjugações, o que contraria a proposta da MD, além das demais questões apontadas nos parágrafos anteriores.

Retoma-se a pergunta que intitula esta dissertação, sobre a existência de uma fronteira entre a flexão e a derivação. Os resultados deste estudo não permitem oferecer uma resposta definitiva, mas pode-se afirmar que com base nos dados dos experimentos desta pesquisa, não é possível afirmar a existência de uma fronteira ou mesmo de um *continuum*. Os dados do Experimento 1 e do Experimento 2 apontam para uma similaridade no processamento do particípio nominal e do verbal do PB, de forma que ou uma fronteira existe e ambos são formados pelo mesmo processo flexional, o que não permite a constatação de fronteira ou mesmo de *continuum*, ou tal fronteira não existe e os processos derivacionais e flexionais estão no mesmo campo, o que dialogaria com a MD.

O próximo capítulo apresenta a conclusão deste trabalho. Nele, retomam-se os principais pontos discutidos e indicam-se encaminhamentos.

## 6 CONCLUSÃO

Os testes realizados e reportados nesta dissertação tiveram por objetivo testar a hipótese de que o particípio passado se divide em duas categoriais, uma com leitura preferencialmente nominal que representa um caso de derivação e uma verbal que representa uma flexão, e que o processamento dos particípios preferencialmente nominais de alta frequência distinguir-se-iam das demais condições por ocorrerem via forma inteira da palavra. Esta hipótese partiu de uma observação de Leminen *et al.* (2013) de um estudo em finlandês que aplicou a técnica de extração de ERP para registrar uma resposta pré-atencional - o MMN - e aponto a possibilidade de que palavras derivadas podem ser armazenadas em forma inteira quando de alta frequência, ao contrário de palavras flexionadas cujo processo de formação é altamente produtivo. Se a hipótese fosse comprovada, não somente seria uma evidência favorável a modelos de dupla rota de processamento, como mostraria, no contexto do presente estudo, que há um processamento diferenciado nesta morfologia que ocupa a fronteira de dois processos morfológicos que, em outros casos, são mais bem distinguidos.

Os principais modelos de processamento lexical de palavras de morfologia complexa foram apresentados no capítulo 4 e alguns dos estudos que evidenciam um ou outro modelo foram apontados. Contrastaram-se os modelos de decomposição precoce automática (TAFT; FORSTER, 1975; TAFT, 2004; STOCKALL; MARANTZ, 2006) e de dupla rota (PINKER, 1991; SCHREUDER; BAAYEN, 1995), que representam os modelos mais influentes nos estudos de processamento morfológico e que pautaram o preparo dos experimentos e as discussões sobre os seus resultados. Enquanto o modelo de decomposição precoce automática prevê que palavras de morfologia complexa são sempre segmentadas nos estágios iniciais de processamento lexical, os de dupla rota preveem a existência de duas vias de processamento para as palavras de morfologia complexa, uma por decomposição em morfemas e uma pela forma inteira da palavra, a depender de determinadas variáveis como regularidade (PINKER, 1991) ou frequência superficial (SCHREUDER; BAAYEN, 1995).

Também no capítulo 4, uma revisão sobre variáveis que impactam o processamento lexical (de forma geral) e o processamento morfológico (mais especificamente) apontaram alguns dos principais efeitos debatidos na literatura. Dentre os mais relevantes para este trabalho, estavam a frequência superficial, que se refere à ocorrência de uma palavra em uma dada língua, variável esta que foi controlada nos experimentos e divididas em dois grupos: alta frequência ou baixa frequência. Outras variáveis com relevância para este trabalho são aquelas

que afetam o processamento visual de palavras, mas que foram mantidas constantes: tamanho da palavra (BAAYEN, 2005), densidade de vizinhança ortográfica (ESTIVALET, 2020), classe de conjugação (JUSTINO; MOTA, 2019) e regularidade do processo de formação das palavras (LEMNEN *et al.*, 2013).

O Experimento 1, reportado no capítulo 5, foi uma tarefa de decisão lexical sob o paradigma de *priming* encoberto, em modalidade visual. A tarefa se refere à decisão sobre se uma sequência de segmentos ortográficos é ou não uma palavra na língua alvo, neste caso o português brasileiro. Os estímulos foram divididos por frequência (alta ou baixa) e tipo de particípio (preferencialmente nominal ou preferencialmente verbal). Ademais, os estímulos foram divididos em duas listas de forma que cada participante viu uma metade das palavras alvo com *prime* morfologicamente relacionado e a outra metade com *prime* não-relacionado, em um formato de quadrado latino. Os resultados apontaram um efeito significativo de *priming* morfológico em todas as condições, o que contradiz a hipótese de que particípios com uma leitura preferencialmente nominal, que equivaleriam a casos de derivação, seriam armazenados em forma inteira quando fossem de alta frequência. A suposição é que palavras armazenadas na memória em sua forma inteira, sem informação de constituintes, não podem se beneficiar da pré-ativação da raiz da palavra compartilhada com o *prime* em um design com *priming* encoberto. No entanto, os resultados mostraram que há segmentação morfológica em todas as condições assim como um efeito facilitador da frequência superficial no reconhecimento visual das palavras, efeito este observável no tempo de resposta à tarefa médio reduzido quando há *priming* (comparado com a condição não primada) e quando a frequência superficial é alta (comparado com a condição de baixa frequência).

Mas todas as palavras no particípio "preferencialmente" nominal são, também, verbos na língua e a exibição da palavra fora de contexto sentencial não permite uma restrição de categoria. Para testar a hipótese de que haveria uma entrada lexical na memória separada para palavras no particípio passado de categoria adjetival, ainda que haja um compartilhamento de forma ortográfica e semântica com sua forma verbal, as palavras da lista de particípio preferencialmente nominal do Experimento 1 foram utilizadas no experimento de leitura automonitorada em dois tipos de sentença: uma em que ela aparecia como parte de um sintagma nominal (logo, como adjetivo) e uma em que ela era parte de uma perífrase de tempo composto (logo, como um verbo). Este teste - o Experimento 2 - foi reportado no capítulo 6. Os tempos de leitura médios romperam as expectativas e se mostraram não favoráveis (ainda que não desfavoráveis) à hipótese: não somente houve ausência de efeito de categoria da palavra, houve

também ausência de efeito de frequência significativa nas sentenças com os participios adjetivos e com os participios verbais. Várias motivações foram apresentadas para dar conta deste último fato, que contradiz os próprios resultados do Experimento 1, dentre elas: fatores não antecipados que podem ter mascarado o efeito de sequência, como o fato de as palavras estarem em contexto sintagmático e isso interferir no tempo de processamento das palavras; e o diálogo com o modelos de processamento lexical de via única, em específico o de decomposição automática, ao levar-se em consideração juntamente os resultados do Experimento 1.

De forma geral, tendo em mente os resultados dos testes de *priming* encoberto e de leitura automonitorada, a maior contribuição deste trabalho foi a demonstração de que há, sim, processamento a nível morfológico de decomposição precoce automática (TAFT; FORSTER, 1975; STOCKALL; MARANTZ, 2006), ao menos para palavras no participio passado de primeira conjugação do português brasileiro. Isto considera a possibilidade de que um modelo de dupla rota nos moldes propostos por Schreuder e Baayen (1995) e Pinker (1991) também seja viável em outras. Contudo, qualquer evidência que aponte para processamento morfológico decomposicional contraria modelos que não preveem a existência de representação de nível morfológico formal ou morfo-ortográfico (e.g., BYBEE, 1995; BAAYEN *et al.*, 2011), de forma que se pode considerar uma contestação destes modelos a partir do Experimento 1. Como grande parte dos estudos disponíveis se referem a outras línguas, experimentos originais e replicações de resultados em estudos sobre o Português Brasileiro são contribuições necessárias.

Alguns pontos ficaram em aberto, como abordado no final do capítulo 5, e que suscitam alguns encaminhamentos de pesquisa. Há efeitos de regularidade do processo de formação da palavra que já foram reportados na literatura com achados neurofisiológicos (ULLMAN *et al.*, 1997; MUNT, 1999; BAKKER *et al.*, 2013) e que parecem apoiar modelos de dupla rota como proposto por Pinker (1991), ainda que haja estudos com resultados na contramão desta proposta (e.g., STOCKALL; MARANTZ, 2006). Como os experimentos reportados nos capítulos 4 e 5 exibiram apenas palavras regulares como estímulos, não se pode descartar a existência de um modelo de dupla rota como o de Pinker (1991) com base nos dados aqui apresentados. Um primeiro encaminhamento é que se realize futuramente um teste sob o paradigma de *priming* encoberto, seja ele psicolinguístico ou com EEG, que compare o processamento de palavras no participio passado regulares com casos de participio abundantes (*i.e.*, que apresentam uma forma regular e uma irregular) e com aqueles casos de participio irregulares.

Outro ponto é que a medida de tempo de leitura, apesar de poder ser considerada *online*, não permite discernir com precisão de milissegundos a ocorrência de processos cognitivos já que o tempo de leitura de cada segmento possivelmente reflete o total do acúmulo de estágios do processamento. Desta forma, um experimento com leitura de sentenças utilizando o mesmo recorte e condições do Experimento 2 também seria proveitoso com técnicas como rastreamento ocular e EEG, já que estes permitem uma avaliação com maior granularidade de dados oriundos de processos cognitivos pela manipulação das condições experimentais e como elas afetam os picos de atividade elétrica neuronal que correspondem a eventos com precisão de milissegundos. Observar-se-ia a ocorrência de diferenças no processamento do segmento crítico de cada condição em dois componentes relacionados ao processamento de morfologia: o N250 (MORRIS; STOCKALL, 2012) e o N400 (LAVRIC *et al.*, 2010).

Um último encaminhamento é para um teste (ou testes) que comparem o processamento de casos não-ambíguos de derivação (por exemplo, "global") e flexão (por exemplo, "escreveu") no português brasileiro. O estudo de Leminen *et al.* (2013) utilizou o componente MMN, que é pré-atencional, para averiguar a diferença no processamento das categorias mencionadas, controladas quanto à variável frequência superficial. Como este componente é detectado no EEG, uma investigação que objetive replicar os resultados de Leminen *et al.* é viável em contextos de laboratórios brasileiros com acesso a esta tecnologia. Enquanto o atual teste explorou palavras no particípio cujo processo de formação não é claro (os particípios com propriedades nominais), um teste com palavras derivadas e flexionadas sem ambiguidade é importante para mapear as várias instâncias no processamento morfológico.



## REFERÊNCIAS

- ANDERSON, S. R. Where's morphology. **Linguistic Inquiry**, vol. 13, no. 4, 1982, p. 571–612.
- ARONOFF, M. Language: Between Words and Grammar. In: JAREMA, G.; LIBBEN, G. (org.). **The mental lexicon: Core Perspectives**. Amsterdam: Elsevier, 2008. Cap. 4, p. 55-80.
- ARONOFF, M. **Word formation in generative grammar**. Cambridge, MA: MIT Press, 1976.
- ARAD, M. Locality Constraints on the Interpretation of Roots: the Case of Hebrew Denominal Verbs. **Natural Language & Linguistic Theory**, v. 21, n. 4, 2003, p. 737-778.
- BAAYEN, R. H. Data Mining at the Intersection of Psychology and Linguistics. In: CUTLER, A. (org.). **Twenty-first century psycholinguistics: Four cornerstones**. Lawrence Erlbaum Associates Publishers, 2005. cap. 5, p. 69–83.
- BAAYEN, R. H. DIJKSTRA, T.; SCHREUDER, R. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. **Journal of Memory and Language**, v. 37, n. 1, 1997, p. 94–117.
- BAAYEN, R. H.; MILIN, P.; FILIPOVIC DURDJEVIC, D.; HENDRIX, P.; MARELLI, M. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. **Psychological Review**, 2011.
- BADDELEY, A. D.; THOMSON, Neil; BUCHANAN, Mary. Word Length and the Structure of Short-Term Memory. **Journal of Verbal Learning and Verbal Behavior**, v. 14, 1975, p. 575-589.
- BAKKER, Iske; MACGREGOR, Lucy J.; PULVERMÜLLER, Friedemann; SHYROV, Yury. Past tense in the brain's time: neurophysiological evidence for dual-route processing of past-tense verbs. **NeuroImage**, v. 71, 2013, p. 187-195.
- BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. Fitting Linear Mixed-Effects Models Using lme4. **Journal of Statistical Software**, v. 67, n. 1, 2015, p. 1–48.
- BARBOSA, M. F. M. **Processamento e Representação de Palavras Derivadas por sufixação no Português Brasileiro**. Tese (Doutrado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.
- BECHARA, E. **Moderna gramática portuguesa**. Rio de Janeiro: Nova Fronteira, 39. ed., 2019.
- BEYERSMANN, E.; CASTLES, A.; COLTHEART, M. Morphological processing during visual word recognition in developing readers: evidence from masked priming. **Experimental Psychology**, v. 65, n. 7, 2013, p. 1306-1326.
- BOOIJ, G. Inflection and Derivation. In: BROWN, K. (org.). **Encyclopedia of Language & Linguistics**. Second Edition, v. 5. Oxford: Elsevier, 2006, p. 654-661.

BYBEE, J. Regular morphology and the lexicon. **Language and Cognitive Processes**, v. 10, 1995, p. 425-455.

CÂMARA JR., Joaquim M. **Estrutura da Língua Portuguesa**. Petrópolis: Vozes, 47. ed., 2015.

CHOMSKY, N. Remarks on Nominalization. In: JACOBS, R.; ROSENBAUM, P. (org.). **Readings in English Transformational Grammar**. MA: Ginn, Waltham, 1970, p. 184-221.

CHOMSKY, N. **The Minimalist Program**. Cambridge Mass: MIT Press, 1995.

COCH, D.; BARES, J.; LANDERS, A. ERPs and morphological processing: the N400 and semantic composition. **Cogn Affect Behav Neurosci**, v. 13, 2012, p. 355-370.

COLTHEART, M.; DAVELAAR, E.; JONASSON, J. F.; BESNER, D. Access to the internal lexicon. In: DORNIG, S. (org.). **Attention & Performance VI**. Hillsdale, NJ: Erlbaum, 1977, p. 535-555.

CREEMERS, A.; DAVIES, A. G.; WILDER, R. J.; TAMMINGA, M.; EMBICK, D. Opacity, transparency, and morphological priming: A study of prefixed verbs in Dutch. **Journal of Memory and Language**, v. 110, 2020, 104055.

CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. Rio de Janeiro: Lexikon, 7. ed., 2014.

DE BELDER, M. M.; FAUST, N.; LAMPITELLI, N. On an inflectional and a derivational diminutive. In: **NORTH EAST LINGUISTIC SOCIETY (NELS 40)**, 40, Cambridge, Massachusetts (MIT), November 13-15, 2009.

EMMOREY, K.; FROMKIN, V. The mental lexicon. In: NEWMeyer, F.J. (org.). **Linguistics: The Cambridge Survey**, Vol. III. Cambridge: Cambridge University Press, 1988, p. 124-149.

ESTIVALET, G. L. Variáveis lexicais e ortográficas no acesso lexical das palavras do português brasileiro. **Revista Linguística**, Rio de Janeiro, v. 16, n. 1, 2020, p. 264-277.

ESTIVALET, Gustavo Lopez; MEUNIER, Fanny. The Brazilian Portuguese Lexicon: Na Instrument for Psycholinguistic Research. **PLOS ONE**, v. 10, n. 12, 2015, p. e0144016.

FIorentino, R; POEPPel, D. Compound words and structure in the lexicon. **Language and Cognitive Processes**, 12, 2007, p. 953-1000.

FOLTRAN, M. J.; CRISÓSTIMO, G. Os adjetivos participiais no português. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 13, n. 1, 2005, p. 129-154.

FORD, M. A.; DAVIS, M. H.; MARSLen-WILSON, W. D. Derivational morphology and base morpheme frequency. **Journal of Memory and Language**, v. 63, n. 1, 2010, p. 117-130.

FORSTER, K. I.; DAVIS, C. Repetition priming and frequency attenuation in lexical access. **Journal of Experimental Psychology: Learning, Memory & Cognition**, v. 10, 1984, p. 680-698, 1984.

FRANÇA, A.; LEMLE, M.; GESUALDI, A.; CAGY, M.; INFANTOSI, A. A neurofisiologia do acesso lexical: palavras em português. **Revista Veredas**, v. 2, 2008. p. 34-49.

GARCIA, D. C. **Efeitos composicionais no reconhecimento visual de palavras compostas em inglês: um estudo com MEG**. Tese de Doutorado. Rio de Janeiro: UFRJ/Faculdade de Letras, 2013.

GARCIA, D. C. Processamento de Palavras. In: MAIA, M. (org.). **Psicolinguística, Psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015, cap. 4, p. 59-70.

GARCIA, D. C.; MAIA, M. A. R.; FRANÇA, A. I. The time course of word recognition: evidence from Brazilian Portuguese. **ReVEL**, v. 10, n. 18, 2012, p. 169-181.

GIRAUDO, H.; GRAINGER, J. Effects of prime word frequency and cumulative root frequency in masked morphological priming. **Language and Cognitive Processes**, v. 15, 2000, p. 421-444.

HALLE, M. Prolegomena to a Theory of Word Formation. **Linguistic Inquiry**, v. 4, n. 1, 1973, p. 3-16, 1973.

HALLE, M.; MARANTZ, A. Distributed Morphology and the Pieces of Inflection. In: HALE, K.; KEYSER, S. J. (org.). **The View from Building 20**. Cambridge: MIT Press, 1993, p. 111-176.

HAY, J. Lexical frequency in morphology: Is everything relative?. **Linguistics**, v. 39, n. 6, 2001, p. 1041–1070.

HONDA, H. Introdução básica à Clusterização. Disponível em: <[https://lamfo-unb.github.io/2017/10/05/Introducao\\_basica\\_a\\_clusterizacao/](https://lamfo-unb.github.io/2017/10/05/Introducao_basica_a_clusterizacao/)>. Acesso em: 20 de fev. de 2022.

JAREMA, G.; LIBBEN, G. Introduction: Matters of Definition and Core Perspectives. In: JAREMA, G.; LIBBEN, G. (org.). **The mental lexicon: Core Perspectives**. Amsterdam: Elsevier, cap. 1, 2008, p. 1-8.

JARQUE, C. M.; BERA, A. K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. **Economics Letters**, v. 6, n. 3, 1980, p. 255–259.

JONG, N. H.; SCHREUDER, R.; BAAYEN, R. H. The morphological family size effect and morphology. **Language and Cognitive Processes**, v. 15, n. 4-5, 2000, p. 329–365.

JUSTINO, J.; MOTA, M. Processamento da morfologia flexional verbal do português brasileiro: um estudo com rastreamento ocular. **Diacrítica**, v. 33, n. 2, 2019, p. 69-88.

LAU, E.; ROZANOVA, K.; PHILLIPS, C. Syntactic prediction and lexical surface frequency effects in sentence processing. **University of Maryland Working Papers in Linguistics**, v. 16, 2007, p. 163–200.

HARTIGAN, J.A.; WONG, M.A. Algorithm AS 136 A KMeans Clustering Algorithm. **Applied Statistics**, v. 28, n. 1, 1979, p. 100-108

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic Time Series Forecasting: The forecast Package for R. **Journal of Statistical Software**, v. 27, n. 3, 2008, p. 1–22.

KAZANINA, N.; DUKOVA-ZHELEVA, G.; GEBER, D.; KHARLAMOV, V.; TONCIULESCU, K. Decomposition into multiple morphemes during lexical access: A masked priming study of Russian nouns. *Language and Cognitive Processes*, v. 23, n. 6, p. 800-823, 2008.

KUTAS, M.; HILLYARD, S. A. Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, v. 207, p. 203-207, 1980.

KUTAS, M.; FEDERMEIER, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, v. 62, p. 621–647, 2011.

KUZMETSOVA, A.; BROCKHOFF, P. B; CHRISTENSEN, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. **Journal of Statistical Software**, v. 82, n. 13, 2017, p. 1–26.

LAVRIC, A.; RASTLE, K.; CLAPP, A. What do fully visible primes and brain potentials reveal about morphological decomposition?. **Psychophysiology**, v. 48, 2011, p. 676-686.

LEMENEN, A.; LEMENEN, M. M.; KRAUSE, C. M. Time course of the neural processing of spoken derived words: an event-related potential study. **NeuroReport**, v. 21, n. 14, 2010, p. 948-952.

LEMENEN, A.; LEMENEN, M.; KUJALA, T.; SHTYROV, Y. Neural dynamics of inflectional and derivational morphology processing in the human brain. **Cortex**, v. 49, 2013. p. 2758-2771.

LEHTONEN, M.; CUNILLERA, T.; RODRÍGUEZ-FORNELLS, A.; HULTÉN, A.; TUOMAINEN, J.; MATTI, L. Recognition of morphologically complex words in Finnish: Evidence from event-related potentials. *Brain Research*, v. 1148, 2007, p. 123-137.

LEHTONEN, M.; VOROBYEV, V.; HUGDAHL, K.; TUOKKOLA, T.; LAINE, M. Neural correlates of morphological decomposition in a morphologically rich language: an fMRI study. **Brain and Language**, v. 98, 2006, p. 182-193.

LENTH, R. lsmeans: Least-Squares Means. Rpackageversion 2.20-23, 2015 <<http://CRAN.Rproject.org/package=lsmeans>>

LOBATO, L. Sobre a forma do participio do Português e o estatuto dos Traços formais. **DeLTA**, v. 15, n. 1, 1999.

MACGREGOR, L. J.; SHYTYROV, Y. Multiple routes for compound word processing in the brain: evidence from EEG. **Brain and Language**, v. 126, n. 2, 2013, p. 217-29.

MAIA, M. Processamento de frases. In: MAIA, M. (org.). **Psicolinguística, Psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015, cap. 1, p. 59-70.

MAIA, M.; LEMLE, M.; FRANÇA, A. I. Efeito stroop e rastreamento ocular no processamento de palavras. **Ciências & Cognição**, v. 12, 2007, p. 2-17.

MARANTZ, A. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. **University of Pennsylvania Working Papers in Linguistics**, v. 4, n. 2, 1997, p. 201-225.

MARIAN, V.; BARTOLOTTI, J.; CHABAL, S.; SHOOK, A. CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. **PLoS ONE**, v. 7, n. 8, 2012, p. e43230.

MARSLEN-WILSON, W.; TYLER, L. K.; WAKSLER, R.; OLDER, L. Morphology and meaning in the English mental lexicon. **Psychological Review**, v. 101, n. 1, 1994, p. 3-33.

MATCHIN, W.; LIAO, C.; GASTON, P.; LAUB, E. Same words, different structures: An fMRI investigation of argument relations and the angular gyrus. **Neuropsychologia**, v. 125, 2019, p. 116-128.

MCCORMICK, S. F.; BRYLSBAERT, M.; RASTLE, K. Is morphological decomposition limited to low-frequency words?. **The Quarterly Journal of Experimental Psychology**, v. 62, n. 9, 2009, p. 1706-1715.

MEDEIROS, A. B. **Traços morfossintáticos e subespecificação morfológica na gramática do Português: um estudo das formas participiais**. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

MEDEIROS, J.; WEISSHEIMER, J.; FRANÇA, A. I.; RIBEIRO, S. Acesso Lexical: uma rota dupla para o português brasileiro. **Fórum Linguístico**, Florianópolis, v. 11, n. 3, 2014, p. 278-292.

MORRIS, J.; STOCKALL, L. Early, equivalent ERP masked priming effects for regular and irregular morphology. **Brain & Language**, v. 123, 2012, p. 81-93.

MÜNTE, T. F.; SAY, T.; CLAHSSEN, H.; SCHILTZ, K.; KUTAS, M. Decomposition of morphologically complex words in English: evidence from event-related brain potentials. **Cognitive Brain Research**, v. 7, n. 3, 1999, p. 241-253.

NAME, C. Psicolinguística da Aquisição da Linguagem. In: MAIA, M. (org.). **Psicolinguística, Psicolinguísticas: uma introdução**. São Paulo: Contexto, 2015, cap. 5, p. 71-84.

PERINI, M. A. **Gramática do português brasileiro**. São Paulo: Parábola Editorial, 2010.

PINKER, S. Rules of language. **Science**, v. 253, n. 5019, 1991, p. 530-535.

PULVERMULLER, F.; SHYTYROV, Y. Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. **Progress in Neurobiology**, v. 79, 2006, p. 49-71.

PYLKKANEN L. Neural basis of basic composition: what we have learned from the red-boat studies and their extensions. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 375, 2020, 20190299.

PYLKKÄNEN, L.; FEINTUCH, S.; HOPKINS, E.; MARANTZ, A. Neural correlates of the effects of morphological family frequency and family size: an MEG study. **Cognition**, v. 91, 2004, p. 35-45.

PYLKKÄNEN, L.; STRINGFELLOW, A.; MARANTZ, A. Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. **Brain and Language**, v. 81, n. 1-3, 2002, p. 666-678.

R CORE TEAM R. **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>, 2013.

RASTLE, K.; DAVIS, M. H. Morphological decomposition based on the analysis of orthography. **Language and Cognitive Processes**, v. 23, n. 7-8, 2008, p. 942-971.

RASTLE, K.; DAVIS, M. H. Reading morphologically complex words: Some thoughts from masked priming. In: KINOSHITA, S.; LUPKER, S. J. (org.). **Masked priming: State of the art**. New York: Psychology Press, cap. 10, 2003, p. 279-305.

RASTLE, K.; DAVIS, M.; NEW, B. The Broth in my Brother's Brothel: Morpho-Orthographic Segmentation in Visual Word Recognition. **Psychonomic Bulletin & Review**, v. 11, n. 6, 2004, p. 394-421.

RODDEN, F.; STEMMER, B. A Brief Introduction to Common Neuroimaging Techniques. In: STEMMER, B.; WHITAKER, H. A. (org.). **Handbook of The Neuroscience of Language**. Oxford: Elsevier, 2008. cap. 6, p. 57-67.

NEVINS, A.; RODRIGUES, C. Athematic participle choice is predicted by argument structure. **Revista da ABRALIN**, v. 13, n. 2, 2014.

SAVINOVA, E.; MALYUTINA, S. Evidence for dual-route morphological processing across the lifespan: data from Russian noun plurals. **Language, Cognition and Neuroscience**, v. 36, n. 6, 2021, p. 730-745.

SCHREUDER, R.; BAAYEN, R. Modeling morphological processing. In: Feldman, L. B. (ed), **Morphological Aspects of Language Processing**. Lawrence Erlbaum, New Jersey: Hillsdale, 1995, cap. 6, p. 131-154.

SILVA, R.; CLAHSEN, H. Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. **Bilingualism: Language and Cognition**, v. 11, 2008, p. 245-260.

SMOLKA, E., ZWITSERLOOD, P.; RÖSLER, F. Stem access in regular and irregular inflection: Evidence from German participles. **Journal of Memory and Language**, v. 57, n. 3, 2007, p. 325–347.

SOTO, M. **ERP and fMRI evidence of compositional differences between linguistic computations for words and sentences**. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

STOCKALL, L.; MARANTZ, A. A single route, full decomposition model of morphological complexity: MEG evidence. **The Mental Lexicon**, v. 1, 2006, p. 85–123.

STOCKALL, L.; MANOULLIDOU, C.; GWILLIAMS, L.; NEOPHYTOU, K.; MARANTZ, A. Prefix stripping re-re-revisited: MEG investigation of morphological decomposition and recomposition. **Frontiers in Psychology**, v. 10, 2019.

TAFT, M. Lexical Access, Cognitive Psychology of. **International Encyclopedia of the Social & Behavioral Sciences**, 2001, p. 8743-8748.

TAFT, M. Morphological decomposition and the reverse base frequency effect. **The Quarterly Journal of Experimental Psychology Section A**, v. 57, 2004, p. 745–765.

TAFT, M.; FORSTER, K. Lexical storage and retrieval of prefixed words. **Journal of Verbal Learning and Verbal Behavior**, v. 14, n. 6, 1980, p. 638-647.

TURGEON, Y.; MACOIR, Joël. Classical and Contemporary Assessment of Aphasia and Acquired Disorders of Language. In: STEMMER, B.; WHITAKER, H. A. (org.). **Handbook of The Neuroscience of Language**. Oxford: Elsevier, 2008. cap. 1, p. 3-11.

ULLMAN, M. T.; CORKIN, S.; COPPOLA, M.; HCKOK, G.; GROWDON, J. H.; KOROSHETZ, W. J.; PINKER, S. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and the grammatical rules are processed by the procedural system. **Journal of Cognitive Neuroscience**, v. 9, n. 2, 1997, p. 266–276.

VAN HEUVEN, W. *et al.* SUBTLEX-UK: A new and improved word frequency database for British English. **The Quarterly Journal of Experimental Psychology**, v. 67, n. 6, 2014, p. 1176–1190.

VIRPIOJA, S.; LEHTONEN, M.; HULTÉN, A.; SALMELIN, R.; LAGUS, K. Predicting Reaction Times in Word Recognition by Unsupervised Learning of Morphology. In: HONKELA, T.; DUCH, W.; GIROLAMI, M.; KASKI, S. (org.). **Artificial Neural Networks and Machine Learning – ICANN**, 2011, p. 275-282.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Nova Iorque: Springer-Verlag New York, 2016.

WHITING, C. M.; MARSLEN-WILSON, W. D.; SHYTYROV, Y. Neural dynamics of inflectional and derivational processing in spoken word comprehension: laterality and automaticity. **Frontiers in Human Neuroscience**, v. 7, 2013, p. 1-15.

ZAUNER, A.; GRUBER, W.; HIMMELSTOß, N. A.; LECHINGER, J.; KLIMESCH, W. Lexical access and evoked traveling alpha waves. **Neuroimage**, v. 91, 2014, p. 252-261.

ZWEIG, E.; PYLKKANEN, L. A visual M170 effect of morphological complexity. **Language and Cognitive Processes**, v. 24, n. 3, 2009, 412 - 439.



## APÊNDICES

APÊNDICE 1 – lista de estímulos alvo do Experimento 1 (tarefa de decisão lexical com *priming* encoberto) por: categoria, frequência e tipo de *prime*.

<b>Categoria</b>	<b>Frequência</b>	<b>Prime não-relacionado</b>	<b>Prime morfológico</b>	<b>Alvo</b>
<b>Particípio “preferencialmente” nominal</b>	Alta frequência	janela	acabou	ACABADO
		papel	aplicou	AFASTADO
		abril	assinou	ASSINADO
		lustre	atacou	ATACADO
		lilás	atrasou	ATRASADO
		uva	condenou	CONDENADO
		uva	convidou	CONVIDADO
		cama	dedicou	DEDICADO
		lustre	derrotou	DERROTADO
		abril	editou	EDITADO
		cobra	elevou	ELEVADO
		cama	fechou	FECHADO
		verde	formou	FORMADO
		rosa	gravou	GRAVADO
		verde	inspirou	INSPIRADO
		lilás	isolou	ISOLADO
		papel	ocupou	OCUPADO
		janela	preparou	PREPARADO
		cobra	privou	PRIVADO
	rosa	roubou	ROUBADO	
	Baixa frequência	boca	agoniou	AGONIADO
	cadeia	alagou	ALAGADO	
	ponta	besuntou	BESUNTADO	
	tecla	castrou	CASTRADO	
	controle	circulou	CIRCULADO	

		caneca	cromou	CROMADO
		tecla	debochou	DEBOCHADO
		controle	duplicou	DUPLICADO
		cadeia	escavou	ESCAVADO
		boca	ferrou	FERRADO
		vinho	frisou	FRISADO
		ponta	gratinou	GRATINADO
		caixa	invejou	INVEJADO
		cadela	lambuzou	LAMBUZADO
		cadela	largou	LARGADO
		caixa	maculou	MACULADO
		caneca	manchou	MANCHADO
		vinho	mesclou	MESCLADO
		azul	tatuou	TATUADO
		azul	venerou	VENERADO
<b>Particípio</b> <b>“preferencialmente”</b> <b>verbal</b>	Alta frequência	mesa	anunciou	ANUNCIADO
		chave	apoiou	APOIADO
		espada	avaliou	AVALIADO
		unha	baseou	BASEADO
		orelha	calculou	CALCULADO
		casa	causou	CAUSADO
		bateria	chegou	CHEGADO
		bateria	cobrou	COBRADO
		casa	colocou	COLOCADO
		filme	comparou	COMPARADO
		camisa	comprou	COMPRADO
		meia	deixou	DEIXADO
		mesa	destinou	DESTINADO
		unha	elaborou	ELABORADO
		filme	firmou	FIRMADO
		meia	liderou	LIDERADO
espada	mostrou	MOSTRADO		

		camisa	negociou	NEGOCIADO
		chave	praticou	PRATICADO
		orelha	voltou	VOLTADO
	Baixa frequência	letra	abdicou	ABDICADO
		bunda	assobiou	ASSOBIADO
		aventou	blusa	AVENTADO
		tapete	banhou	BANHADO
		bunda	batalhou	BATALHADO
		pata	berrou	BERRADO
		pata	bocejou	BOCEJADO
		pasta	cessou	CESSADO
		marca	cintilou	CINTILADO
		novela	clicou	CLICADO
		mosca	cravou	CRAVADO
		tapete	dançou	DANÇADO
		mosca	decolou	DECOLADO
		novela	escapou	ESCAPADO
		reserva	expulsou	EXPULSADO
		reserva	fabulou	FABULADO
		marca	fritou	FRITADO
		letra	igualou	IGUALADO
		pasta	orbitou	ORBITADO
		blusa	peitou	PEITADO
	<b>Pseudopalavras no particípio</b>	<i>Não se aplica</i>	rúcula	beseou
alface			cacolou	CACOLADO
perfume			chicou	CHICADO
arroba			colcalou	COLCALADO
inveja			dilserou	DILSERADO
rúcula			fernadou	FERNADO
dúvida			grobou	GROBADO
inveja			idetou	IDETADO
manga			mindou	MINDADO

		seta	morcou	MORCADO
		baleia	nitagrou	NITAGRADO
		manga	nogaciou	NOGACIADO
		perfume	ofostou	OFOSTADO
		cabeça	pancorou	PACONRADO
		arropa	roldetou	ROLDETADO
		alface	soilou	SOILADO
		seta	tisdenou	TISDENADO
		baleia	traticou	TRATICADO
		dúvida	tros mou	TROSMADO
		cabeça	vincodou	VINCODADO
		vacina	aguilou	AGUILADO
		dente	arbetou	ARBETADO
		bolha	ascovou	ASCOVADO
		dente	banluzou	BANLUZADO
		dose	benhou	BENHADO
		pimenta	bufalou	BUFALADO
		tomate	cansilou	CANSILADO
		tomate	cebojou	CEBOJADO
		garrafa	cilitou	CILITADO
		esfera	clocou	CLOCADO
		bata	codelou	CODELADO
		bata	cramou	CRAMADO
		esfera	diplecou	DIPLECADO
		baguete	dochebou	DOCHEBADO
		vacina	elvajou	ELVAJADO
		dose	frosou	FROSADO
		pimenta	ostecou	OSTECADO
		poitou	baguete	POITADO
		bolha	tincilou	TINCILADO
		garrafa	tutuou	TUTUADO

APÊNDICE 2 – lista de estímulos do Experimento 2 (leitura automonitorada)

<b>Frequência superficial</b>	<b>Segmento crítico</b>	<b>Sentença – participio adjetivo</b>	<b>Sentença – participio verbal</b>
<b>Alta frequência</b>	ACABADO	O troço <b>acabado</b> estava lá.	Eu tinha <b>acabado</b> aquilo há tempos.
	AFASTADO	O cara <b>afastado</b> está aqui.	Eu tinha <b>afastado</b> elas o tempo todo.
	ASSINADO	O negócio <b>assinado</b> está aqui.	Eu tinha <b>assinado</b> isso ontem.
	ATACADO	O garoto <b>atacado</b> estava longe.	Eu tenho <b>atacado</b> eles constantemente.
	ATRASADO	O rapaz <b>atrasado</b> estava ali fora.	Eu tinha <b>atrasado</b> ela ontem.
	CONDENADO	O cara <b>condenado</b> esteve por lá.	Ele tinha <b>condenado</b> a mim um tempo atrás.
	CONVIDADO	O cara <b>convidado</b> está por perto.	Eu tinha <b>convidado</b> eles há semanas.
	DEDICADO	O garoto <b>dedicado</b> esteve lá.	Ela tinha <b>dedicado</b> isso a mim um tempo atrás.
	DERROTADO	Ela tem <b>derrotado</b> elas o tempo todo.	O rapaz <b>derrotado</b> está perto daqui.
	EDITADO	O negócio <b>editado</b> estava por ali.	Ela tinha <b>editado</b> ele ontem.
	ELEVADO	O lugar <b>elevado</b> está por lá.	Ele tinha <b>elevado</b> aquilo um tempo atrás.
	FECHADO	O lugar <b>fechado</b> está aqui	Eu tinha <b>fechado</b> aquilo faz tempo.
	FORMADO	O garoto <b>formado</b> estava ali.	Ele tinha <b>formado</b> a mim há um tempo.
	GRAVADO	O troço <b>gravado</b> está por ali.	Eu tenho <b>gravado</b> isso constantemente.
	INSPIRADO	O rapaz <b>inspirado</b> estava lá.	Ela tinha <b>inspirado</b> a mim há meses.
	ISOLADO	O rapaz <b>isolado</b> esteve longe.	Ela tem <b>isolado</b> eles sempre.
OCUPADO	O cara <b>ocupado</b> esteve aqui.	Ele tem <b>ocupado</b> a mim todo dia.	

	PREPARADO	O troço <b>preparado</b> está lá fora.	Ele tem <b>preparado</b> ela o tempo todo.
	PRIVADO	O negócio <b>privado</b> está longe.	Ela tem <b>privado</b> a mim sempre.
	ROUBADO	O troço <b>roubado</b> estava aqui.	Ele tinha <b>roubado</b> isso há anos.
<b>Baixa frequência</b>	AGONIADO	O cara <b>agoniado</b> estava ali.	Ela tem <b>agoniado</b> a ele sempre.
	ALAGADO	O lugar <b>alagado</b> estava por ali.	Ontem tinha <b>alagado</b> muito aqui.
	BESUNTADO	O troço <b>besuntado</b> estava ali.	Eu tenho <b>besuntado</b> aquilo sempre.
	CASTRADO	O animal <b>castrado</b> esteve por ali.	Ela tinha <b>castrado</b> eles há meses.
	CIRCULADO	O lugar <b>circulado</b> está ali.	Ela tinha <b>circulado</b> isso ontem.
	CROMADO	O negócio <b>cromado</b> estava aqui.	Ela tinha <b>cromado</b> aquilo um tempo atrás.
	DEBOCHADO	O garoto <b>debochado</b> esteve lá dentro.	Ele tem <b>debochado</b> dela o tempo todo.
	DUPLICADO	O negócio <b>duplicado</b> está lá fora.	Ele tinha <b>duplicado</b> isso há tempos.
	ESCAVADO	O lugar <b>escavado</b> estava longe.	Eu tenho <b>escavado</b> aqui o tempo todo.
	FERRADO	O garoto <b>ferrado</b> está ali.	Eu tenho <b>ferrado</b> ele o tempo todo.
	FRISADO	O troço <b>frisado</b> está aqui.	Eu tinha <b>frisado</b> isso o tempo todo.
	GRATINADO	O troço <b>gratinado</b> estava lá dentro.	Ela tem <b>gratinado</b> aquilo sempre.
	INVEJADO	O garoto <b>invejado</b> esteve lá.	Ele tem <b>invejado</b> a ela sempre.
	LAMBUZADO	O animal <b>lambuzado</b> esteve aqui perto.	Eu tinha <b>lambuzado</b> aquilo há um tempo.
	LARGADO	O cara <b>largado</b> esteve por perto.	Ele tinha <b>largado</b> aquilo faz tempo.
MACULADO	O rapaz <b>maculado</b> estava longe.	Eu tinha <b>maculado</b> a ele faz anos.	

	MANCHADO	O negócio <b>manchado</b> esteve ali.	Eu tinha <b>manchado</b> isso há meses.
	MESCLADO	O troço <b>mesclado</b> estava aqui fora.	Ele tem <b>mesclado</b> isso com aquilo.
	TATUADO	O cara <b>tatuado</b> está ali.	Ela tinha <b>tatuado</b> isso tem tempo.
	VENERADO	O garoto <b>venerado</b> está lá fora.	Eu tinha <b>venerado</b> elas anos atrás.