

RELATO DE PESQUISA

# UM ESTRANHO NO NINHO: O PARENTESCO MORFOLÓGICO ENTRE PALAVRAS

Julia CATALDO  

Universidade Federal do Rio de Janeiro (UFRJ)

Aniela FRANÇA  

Universidade Federal do Rio de Janeiro (UFRJ)



## EDITORES

- Miguel Oliveira, Jr. (UFAL)
- René Almeida (UFS)

## AVALIADORES

- José Neto (UFPB)
- Francisca Carvalho (UFPA)
- Marcus Maia (UFRJ)

## SOBRE OS AUTORES

- Julia Cataldo  
Conceptualização, Curadoria de Dados, Análise Formal, Aquisição de Financiamento, Investigação, Metodologia, Administração do Projeto, Software, Visualização, Escrita - Rascunho Original, Escrita - Revisão e Edição.
- Aniela França  
Conceptualização, Curadoria de Dados, Análise Formal, Aquisição de Financiamento, Investigação, Metodologia, Administração do Projeto, Recursos, Software, Supervisão, Validação, Visualização, Escrita - Revisão e Edição.

## DATAS

- Recebido: 20/08/2021
- Aceito: 19/11/2021
- Publicado: 14/12/2021

## COMO CITAR

CATALDO, Julia; FRANÇA, Aniela (2021). Um Estranho no Ninho: o parentesco morfológico entre palavras. *Cadernos de Linguística*, v. 2, n. 4, e499.

## RESUMO

Uma das operações linguísticas inevitáveis é a do acesso lexical, que nos permite, com enorme facilidade e rapidez, entender e/ou produzir palavras soltas. Porém, por trás da aparente simplicidade da relação entre forma e significado no âmbito da palavra, há mecanismos que estão longe de serem completamente compreendidos e que vêm gerando muita discussão nos últimos anos. Neste estudo psicolinguístico de priming transmodal, buscamos entender como representamos e acessamos um alvo que, apesar de compartilhar a raiz com seu *prime*, é estranho à acepção semântica do mesmo. Trata-se de pares de palavras que possuem relacionamento morfológico e semanticamente diacrônico entre si, mas não semanticamente sincrônico, como *refrigerar* e *refrigerante*. Será que, quando acessamos *refrigerante*, enxergamos *refrigerar*? Além disso, como representamos *refrigerante* no nosso léxico mental? Por fim, será que a rota de processamento de *refrigerar-refrigerante* apresenta diferenças em relação a uma rota unicamente semântica, como em *refrigerar-esfriar*, e a uma rota composicional com semântica transparente, como em *refrigerar-refrigerador*? Os resultados apontam para a confirmação da hipótese de modelos decompositivos, sugerindo: i) novas entradas no léxico mental para palavras como *refrigerante*; e ii) uma relação de identidade entre palavras composicionalmente relacionadas, uma relação mais fraca de referencialidade entre palavras

opacamente relacionais e uma relação variável de uso conjunto entre palavras apenas semanticamente relacionadas.

#### ABSTRACT

One of the inevitable linguistic operations is that of lexical access, which easily allows us to understand and/or produce words. However, behind the apparent simplicity of the relationship between form and meaning in the context of the word, there are mechanisms that are far from being fully understood and that have been generating a lot of discussion in recent years. In this psycholinguistic crossmodal priming study, we seek to understand how we represent and access a target word that, despite sharing the root with the prime, is foreign to its semantic meaning. These are pairs of words that have a morphological and a semantically diachronic relationship with each other, but that do not have a semantically synchronic one. For example, do we see *refrigerar* (meaning “to cool” in Portuguese) in *refrigerante* (“soda” in Portuguese)? Also, how do we represent *refrigerante* (soda) in our mental lexicon? Finally, does the *refrigerar-refrigerante* (to cool/soda) processing route differ from a purely semantic route, as in *refrigerar-esfriar* (morphologically entirely different words but both meaning “to chill”), and a compositional route with transparent semantics, as in *refrigerar-refrigerador* (to chill-fridge)? The results point to the confirmation of the hypothesis of decompositional models, suggesting: i) new entries in the mental lexicon for words like *refrigerante* (soda); ii) an identity relationship between compositionally related words, a weaker referential relationship between opaque related words and a variable joint use relationship between semantically-only related words.

#### PALAVRAS-CHAVE

Acesso Lexical; Opacidade Semântica; Morfologia Distribuída;  
Teste Psicolinguístico de Priming.

#### KEYWORDS

Lexical Access; Semantic Opacity; Distributed Morphology;  
Psycholinguistic Priming Test.

## O PROCESSAMENTO E O ACESSO A PALAVRAS

Ao ler *Depois do surto, a maioria das lojas vai entrar em liquidação*, será que se percebe que existe o nome *líquido* aninhado dentro de *liquidação*? Será que *líquido* também é ativado no cérebro?

É claro que enquanto análise superficial e *off-line* da língua, a palavra *líquido* é bem visível em *liquidação*, que também carrega o morfema *-ção*, muito recorrente no português. No entanto, *líquido* e *liquidação* podem parecer não-relacionáveis semanticamente, especialmente porque, durante o uso, raramente nos damos conta de que se trata de um processo de idiomatização pelo qual passa o verbo *liquidar*, cuja negociação semântica é, em seguida, herdada pelo nome derivado *liquidação*. A despeito de tal análise *off-line*<sup>1</sup>, a questão fundamental aqui é outra: realmente nos átimos de segundo de que dispomos para fazer sentido das frases que produzimos<sup>2</sup> e que nos chegam, seja por via escrita, falada ou sinalizada, não computamos uma análise consciente do que percebemos, pensamos ou fazemos.

A idiomatização é uma leitura que se aparta da leitura composicional. A leitura composicional de uma estrutura morfológicamente complexa oferece previsibilidade a cada instante em que uma nova camada morfológica é inserida. Por exemplo, em *liquidífica-d-or*, temos o conceito básico negociado na primeira vez que a raiz *liquid-* se combina com o nominalizador  $\emptyset$  e com a vogal temática *a*. Depois desse ponto, as camadas subsequentes de verbalizador e adjetivador se compõem composicionalmente, assim como o nominalizador *-or*, que fecha o ciclo de derivação dessa palavra: liquidificador = aquele que torna alguma coisa líquida. Contudo, a rota morfológica composicional não impede que haja uma rota paralela não-composicional. Esse é o caso de *liquidação*, em que o nome *líquido* se combina com um verbalizador e se torna *liquidar*, que negocia seu significado não a partir de *líquido* com um todo, mas a partir de uma propriedade dos líquidos que é a propensão para a evasão. Assim, tornar *líquido* alguma coisa sólida é restringido para assumir o significado idiomático de se esvaír, terminar. Depois desse ponto, as camadas subsequentes herdaram tal significado idiomático de base e passam a compor composicionalmente os derivados. Existe uma ligeira coerção nesse

1 Em uma análise *offline*, o processamento linguístico é reflexivo, isto é, não se trata do que acontece inconscientemente em nossa mente quando recebemos um estímulo linguístico oral, visual ou escrito, mas sim de uma reflexão ativa e posterior sobre tal estímulo. Ele se opõe ao processamento *online*, que é reflexo e cujas medidas são aferidas durante o processamento inconsciente (KENNEDY, 2015).

2 Agradecemos a um dos revisores por nos fazer explicitar aqui, de início, o que entendemos ser acesso lexical. Pela abordagem formal da Gramática Gerativa, escolhida neste artigo, acesso lexical, na compreensão ou na produção, é engendrado pela mesma operação de análise pela síntese. Ela ativa uma representação discreta mental que relaciona forma e conteúdo de uma palavra solta ou em contexto sentencial, a qual ouvimos, lemos ou vamos falar (cf. MORRIS; FRANK, 1959; BEVER; POEPEL, 2010).

processo no contexto de venda, muito produtivo por sinal, em que *liquidação* assume o sentido de uma ação em que o estoque da loja está no processo de término e o ajuste pragmático nos faz entender que deva ser em relação à remarcação de preços. Note que também há um certo ajuste desse tipo com a própria palavra *liquidificador*, já que ele não necessariamente liquidifica alimentos (cf. LEMLE; FRANÇA, 2006 e PEDERNEIRA; LEMLE, 2015 para um maior detalhamento dos casos e do funcionamento morfológico da idiomatização no Português Brasileiro).

Tomando como base a teoria da Morfologia Distribuída (HALLE; MARANTZ, 1993), que propõe que exista sintaxe até dentro das palavras, podemos interpretar as proposições de Pylkkänen (2019) – um texto novo, porém já seminal – pensando na relação entre forma e conteúdo no estudo do processamento lexical. Nesse texto (PYLKKÄNEN, 2019, p. 65), a autora mostra que, apesar de tomarmos como verdade que a sintaxe opera independentemente da semântica – como Chomsky (1957, p.15) provou com seu célebre exemplo *Colorless green ideas sleep furiously* –, até hoje não conseguimos decifrar completamente seu funcionamento. Por exemplo, não sabemos como a sintaxe é computada em nosso cérebro, nem como ela atravessa as interfaces. Assim, dissociá-la da semântica, no acesso lexical e no processamento de sentença nas línguas naturais, continua sendo um grande desafio da pesquisa linguística.

Dessa forma, na área do processamento e representação lexical, concordamos que o falante-ouvinte não consegue ter uma percepção clara sobre a forma com a qual acessou uma palavra, isto é, se ela foi: i) fisgada inteira de dentro de um pequeno reino de palavras independentes no cérebro; ou ii) se suas partes foram escolhidas dentre várias já estocadas para serem, então, compostas em uma linha de montagem de palavras que estariam em vias de serem faladas ou interpretadas.

O presente trabalho enfoca o acesso a palavras formadas por relação semântica que se dá apenas diacronicamente. Melhor dizendo, nos interessamos por palavras derivadas que, apesar de não apresentarem mais relação semântica transparente com suas palavras primitivas, mantêm a estrutura morfológica das mesmas. É o caso da já apresentada *liquidação*, que vem de *líquido*, mas também de *refrigerante*, que vem de *refrigerar*, e de *professor*, que vem de *professar*. Ao pedirmos um copo de Coca-Cola ou de Guaraná, estamos pensando que, se seguissemos a semântica original de *refrigerar*, a função do *refrigerante* seria a de nos refrescar? Ao nos dirigirmos a um *professor*, temos consciência ativa de que ele *professa* um saber durante suas aulas?

Pretendemos responder tais perguntas ao observar as diferenças de processamento entre três grupos de palavras: i) palavras morfológicamente relacionadas cuja relação semântica é sincronicamente transparente (como *líquido-liquidificar*); ii) palavras morfológicamente relacionadas cuja relação semântica é sincronicamente opaca (como

*líquido-líquidação*); e iii) palavras que não apresentam relacionamento morfológico entre si, mas cuja relação semântica é transparente (como *líquido-aquoso*).

Assim, realizamos um experimento psicolinguístico, através do qual nos lançamos a entender: i) como se dá o processamento das palavras (se são acessadas inteiras ou por suas partes); ii) se o processamento de palavras que são morfológicamente relacionadas, mas semanticamente opacas às que as originaram é, de alguma maneira, diferente do de palavras com significado sincrônico transparente; e iii) se a rota morfológica de processamento difere da rota semântica, analisando de que maneira elas podem favorecer/acelerar ou desfavorecer/desacelerar o acesso lexical.

A todos esses três pontos subjazem mecanismos que todo falante realiza para associar forma e conteúdo em uma palavra. Dessa maneira, esperamos, com esse estudo, ter dado um passo adiante no estudo da sintaxe e da semântica dentro das palavras.

## 1. HIPÓTESES PSICOLINGÜÍSTICAS: AS TRÊS ABORDAGENS

As questões postas pelo acesso lexical vêm gerando há muitas décadas uma enorme gama de testes sofisticados de comportamento, de neurofisiologia e de hemodinâmica, os quais permitem um acompanhamento *online* preciso do processamento. Em geral, esses experimentos têm um desenho metodológico especialmente moldado para testar hipóteses formuladas a partir de questões legítimas, oriundas, simplificada, de três tipos diferentes de abordagens linguísticas.

A hipótese da primeira dessas abordagens é a de que o cérebro tem uma capacidade de armazenamento virtualmente ilimitada e, em decorrência desse estado de coisas, existiria acesso lexical direto às palavras, sem a necessidade de estoque e remontagem de suas partes. Assim, em tais modelos de listagem plena, a memória é essencial, porque não só o item é memorizado de maneira independente e em todas as suas possíveis formas, como também seriam memorizadas todas as situações de seu uso conjunto a outras palavras. Seria, então, formada uma rede de itens – menos ou mais reforçados pela frequência de cada um e pela frequência de suas famílias – que poderiam ser acessados diretamente em qualquer acepção. Portanto, o caminho tomado nessa abordagem começa pela semântica e vai na direção de cima para baixo (*top-down*), supostamente garantindo a eficiência necessária ao uso *online* imediato e plenamente contextualizado da linguagem (BUTTERWORTH, 1983; HAY; BAAYEN, 2005; JACKENDOFF, 1983).

Na segunda abordagem, a hipótese entretida é a decomposicional. O cérebro humano teria uma capacidade de decompor as palavras que processa e também de juntar morfemas para produzir novos compostos. Essa capacidade, determinada pela genética

da espécie, seria minimamente necessária à linguagem e, por isso, a operação básica de *merge* – ou concatenação, em português – de morfemas (CHOMSKY, 1995) não apresentaria custo para a computação e poderia ser eficientemente utilizada todo tempo: *it's syntax all the way down* (sintaxe todo tempo; MARANTZ, 1997). Portanto, a aposta desses modelos é na computação morfossintática plena. Eles preveem a decomposição completa das partes na direção de baixo para cima (*bottom-up*), indo da sintaxe à semântica através de regras computacionais explícitas. É costume nesses modelos decomposicionais acreditar-se que guardamos o menos possível de léxico na mente, submetendo os morfemas à atuação de algumas regras, que nos permitiriam formar diversas palavras com aqueles pareamentos de forma e significado já estocados. Os pesquisadores que reforçam esses modelos defendem que eles seriam menos custosos do que os não-decomposicionais (TAFT, FORSTER, 1975; 1976; FORSTER, 1979; TAFT, 1979; PYLKÄNNEN et al., 2000; STOCKALL, MARANTZ, 2006; EMBICK et al., 2001; PYLKÄNNEN et al., 2002; PYLKÄNNEN, MARANTZ, 2003; PYLKÄNNEN et al., 2003).

Também há pesquisadores que acreditam que, uma vez que existem no cérebro córtices para processamento e também áreas para armazenamento, um processo não necessariamente excluiria o outro. Esses perfazem, portanto, um terceiro tipo de abordagem teórica: a dos modelos mistos. Eles propõem que o cérebro esteja pronto tanto ao acesso *bottom-up* quanto ao *top-down*, de maneira que a forma utilizada seria sempre a mais eficiente, a depender da palavra que se quer acessar: palavras mais transparentes e/ou regulares seriam acessadas por decomposição e, palavras mais opacas e/ou irregulares, por acesso às redes (CARAMAZZA et al., 1988; MARSLÉN-WILSON et al., 1994; PINKER, 1991).

Nossas previsões para este trabalho, contudo, lidam com estímulos diferentes dos que foram relatados aqui. Acreditamos que haja decomposição *bottom-up* não só em palavras com vínculos vivos e sincrônicos, mas também em palavras como *refrigerante*, cujas fronteiras morfológicas parecem ter sido refeitas a ponto de não negociarem sentido sincrônico com *refrigerar*.

## 2. ESTUDOS MOTIVADORES

Diversos estudos tentam oferecer uma contribuição ao desempate entre as três propostas apresentadas anteriormente. Os resultados de quatro, em particular, motivaram as perguntas experimentais do presente trabalho.

Em primeiro lugar, Garcia (2009) realizou um teste de *priming*<sup>3</sup> monomodal<sup>4</sup> encoberto<sup>5</sup> com quatro condições de relacionamento de palavras: morfológico (pares como FILA-fileira), apenas semântico (pares como ORDEM-fileira), apenas fonológico/ortográfico (pares como FILÉ-fileira) e nulo (pares como MATO-fileira). Os tempos de resposta foram significativamente menores para os pares da condição morfológica e, para os da condição fonológica, foram até maiores que para os da condição de relacionamento nulo. Apesar de essa última diferença não ter sido estatisticamente relevante, a autora sugere a interpretação de que a semelhança fonológica entre *prime* e alvo possa inibir o processamento lexical. Ainda, os baixos tempos de resposta para os pares da condição morfológica estimularam Garcia a defender o modelo decomposicional (a segunda abordagem apresentada na sessão anterior), no qual o léxico mental seria estruturado morfológicamente, morfema a morfema, de maneira que fatores semânticos tardios não seriam relevantes para os momentos iniciais do processamento.

O segundo estudo que nos motivou foi o de França et al. (2008). Os autores se debruçaram sobre a hipótese da diferença derivacional entre o acesso semântico à parte arbitrária da palavra, isto é, à concatenação da raiz com o primeiro morfema categorizador, e à parte composicional da palavra. Para tanto, fizeram um teste de *priming* multimodal, dividindo seus estímulos em dois grupos, fonológico (GF) e morfológico (GM), cada um contendo 3 tamanhos diferentes de palavras: GF possuía estímulos com 3 quantidades diferentes de sílabas, enquanto GM possuía estímulos de 3 quantidades diferentes de camadas morfológicas. Eles acreditavam que essa diferença de tamanho os ajudaria a melhor entender a derivação lexical, uma vez que variava o ponto da arbitrariedade saussuriana em GF, mas não em GM.

Analisando os ERPs (*Event Related Potentials*)<sup>6</sup> produzidos pelos participantes, os autores constataram que, enquanto em GF o acesso a alvos menores era mais rápido que

3 O paradigma experimental de *priming* consiste em apresentar a uma pessoa duas palavras e, através de uma resposta comportamental (como apertar um botão), ou cerebral (ondas retiradas de um eletroencefalógrafo – EEG) que ela apresente, analisar se a presença da primeira palavra – chamada de *prime* – ajuda de alguma forma o reconhecimento e o processamento da segunda – chamada de alvo ou *target*, em inglês. Essa metodologia é muito usada na experimentação do acesso lexical porque ela permite que se teste pares de palavras com diferentes tipos de relação entre si, como morfológica (casa-caseiro), fonológica (batata-barata) e semântica (professor-escola), para analisar se e como esses fatores influenciam o processamento lexical.

4 O paradigma de *priming* pode ser monomodal, quando *prime* e alvo são apresentados através da mesma modalidade, (por exemplo, modalidade visual, em que ambos estariam escritos), ou pode ainda ser transmodal, quando o *prime* é apresentado em uma modalidade diferente da do alvo (por exemplo, *prime* visual e alvo auditivo).

5 *Priming* encoberto é uma técnica na qual o *prime* aparece escrito na tela por cerca de apenas 30ms (milissegundos), não sendo computado conscientemente pelo participante. O alvo, por sua vez, permanece na tela por cerca de 300ms. Mesmo com tal ínfimo tempo de exposição, o *prime* exibido encobertamente exerce muita influência na ativação do alvo (para uma revisão completa desse tipo de metodologia, cf. VAN DEN BUSSCHE et al., 2009).

6 ERPs (Potenciais Relacionados a Eventos, em português) é uma técnica utilizada para se analisar a atividade cerebral decorrente de um estímulo específico, no contexto de uma tarefa experimental. Tal atividade (elétrica pós-sináptica) é medida via eletroencefalografia (EEG). Ao se extrair os ERPs dos resultados da EEG, analisa-se: i) a

o a maiores, em GM o tamanho dos alvos não gerou nenhuma diferença estatisticamente relevante nos tempos de processamento. Além disso, os resultados gerais de GM foram mais rápidos que os de GF. Os autores sugeriram, então, a morfologia como um fator de aceleração do reconhecimento de palavras e analisaram seus resultados à luz da Morfologia Distribuída (HALLE; MARANTZ, 1993), uma abordagem decomposicional que propõe que a derivação lexical passe por três fases. A concatenação da raiz com o primeiro morfema categorizador se daria na terceira fase, a Enciclopédia, em que um significado seria pareado com a forma da palavra, sendo esse o momento do acesso lexical por excelência. Após essa primeira computação, as demais camadas morfológicas da palavra seriam processadas composicionalmente.

França et al. (2008) terminam seu estudo se perguntando, no entanto, o que aconteceria com pares de palavras que suscitam dúvidas nos falantes quanto à sua relação semântica sincrônica. Por exemplo, parece que não se reconhece mais *restaurar* dentro de *restaurante* e *refrigerar* dentro de *refrigerante*, ambos exemplos dados pelos próprios autores. Nesses casos, em que há relacionamento semântico diacrônico, mas não sincrônico entre palavras pertencentes às mesmas famílias morfológicas, será que a morfologia também será um fator de aceleração no reconhecimento lexical?

Em Maia, Lemle e França (2007), foram aplicados dois testes utilizando rastreamento ocular em uma plataforma para medir o efeito *stroop*<sup>7</sup>. O primeiro teste verificou o nível de conhecimento intuitivo dos morfemas nas palavras. Os resultados apontaram para a existência de um processamento morfológico consciente tanto para morfemas que foram computados composicionalmente quanto para aqueles que geravam leituras arbitrárias. Portanto, os resultados do primeiro teste corroboram a existência do processamento *bottom-up*. Os resultados do segundo teste, no entanto, indicaram uma diferença de fixações oculares para, de um lado, morfemas com leitura composicional e, de outro, morfemas com leitura arbitrária e pseudomorfemas. Isso, por sua vez, advoga pela existência (simultânea) de um processamento *top-down*. Os autores ressaltaram que as diferenças de metodologia e tarefa experimental podem ter sido os responsáveis por tais contradições e sugeriram que mais estudos fossem feitos.

Em 2015, Maia e Ribeiro trabalharam com o acesso lexical a pseudopalavras mono e polimorfêmicas (como *liboramima* e *norbalense*, respectivamente), comparando-o com o

latência das ondas cerebrais, que está diretamente relacionada ao instante da apresentação da palavra-alvo experimental; e ii) sua amplitude, que tem correlação com a intensidade da atividade cortical, permitindo, assim, que os esforços para a realização da integração sintática e/ou semântica sejam quantificados.

7 No paradigma do efeito *stroop*, escreve-se o nome de uma cor, como *vermelho* e muda-se a cor da fonte para uma outra que não vermelho: amarelo, por exemplo. Assim, o participante vê a palavra *vermelho* escrita em amarelo. Sua tarefa é nomear a cor em que a palavra *vermelho* está escrita, portanto, ele deve dizer *amarelo*. Um grupo controle em que a palavra escrita corresponde à cor que ela denota também é testado (*vermelho* escrito em vermelho). Para um melhor detalhamento da técnica, cf. Stroop, 1935.

de palavras também mono e polimorfêmicas (como *jabuticaba* e *jornaleiro*, respectivamente). O uso das pseudopalavras com sufixos comuns na língua estudada permite que se saiba qual curso segue o processamento linguístico de um participante, quando a ele é apresentada o que a princípio é uma nova palavra do português. Concluíram que o acesso a palavras e pseudopalavras polimorfêmicas segue um curso *bottom-up*, pela possibilidade de decomposição do lexema. Propuseram ser esse o mesmo motivo pelo qual as pseudopalavras e as palavras multimorfêmicas tiveram mais aceitação na tarefa de decisão lexical, quando comparadas às palavras e pseudopalavras monomorfêmicas. Não descartaram, no entanto, a possibilidade da existência de um processamento dupla rota.

Diante de um cenário ainda de muitas dúvidas sobre o curso do processamento lexical, o presente trabalho conta com um grupo seletivo de palavras estímulos que manipula morfologia sob os recortes da sincronia e da diacronia. Acreditamos que, com ele, poderemos avançar em alguma medida no entendimento desse campo fundamental da linguagem que é o acesso lexical.

### 3. OS EXPERIMENTOS COMPORTAMENTAIS

Elaboramos um teste comportamental *offline* e um *online*. O teste *offline* nos permitiu colher e validar estímulos que, posteriormente, foram utilizados no teste *online*, cujo objetivo era a observação do curso temporal do processamento lexical e, conseqüentemente, de seu custo.

#### 3.1. EXPERIMENTO COMPORTAMENTAL OFFLINE – PRÉ-TESTE

Uma vez escolhidos os *primes*, restava a árdua tarefa de elencar e validar seus alvos. O experimento comportamental *offline* teve, então, duas grandes razões de ser: i) colher possíveis alvos para as condições morfológicas do experimento *online*; e ii) validar a existência de relacionamento semântico entre *primes* e alvos que havíamos inicialmente selecionado para a condição apenas semântica do experimento *online*. Tal revalidação é importante porque, apesar de ambas as autoras do presente trabalho serem falantes nativas de português, o nível de relação semântica atribuído a um par de palavras pode variar de falante para falante, de acordo com as idiosincrasias e particularidades dos meios interacionais em que se encontram.

Assim, tal teste comportamental *offline* serviu como um pré-teste de validação dos materiais experimentais do teste *online*. Optamos por realizá-lo na plataforma *Google* Formulários.

### 3.1.1. PARTICIPANTES

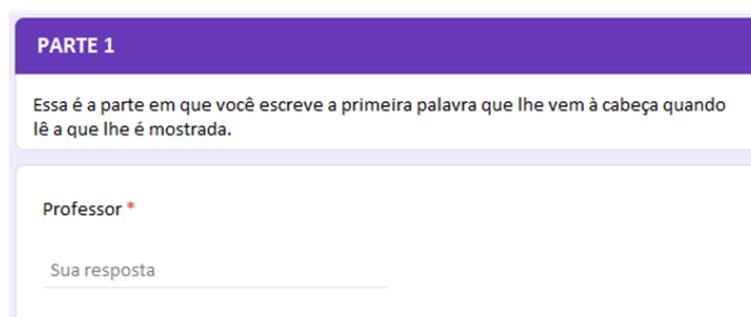
No total, 30 pessoas (diferentes das que posteriormente participaram do experimento comportamental *online*) registraram suas repostas. Suas idades variavam entre 18 e 35 anos e todas aceitaram participar voluntariamente do teste. No grupo, havia indivíduos com nível superior completo e indivíduos que cursavam sua graduação. As áreas do saber de suas atividades eram variadas, perpassando as ciências humanas, as exatas, as biológicas, as linguísticas, as literárias e as educacionais. Todos moravam no estado do Rio de Janeiro.

### 3.1.2. MATERIAIS E DESENHO EXPERIMENTAL

Dividimos o formulário em dois grandes momentos. No primeiro, cujo objetivo era colher possíveis alvos para as condições morfológicas, utilizamos uma adaptação da técnica de *cloze* aberto<sup>8</sup>. Processos cognitivos como a percepção e a memória estão envolvidos nessa técnica experimental, que se mostrou interessante para nosso estudo por validar que o emissor (o participante) e o receptor (as autoras do teste) compartilhassem um mesmo conjunto de mensagens, isto é, uma mesma opinião sobre a existência de um relacionamento entre duas palavras (para uma revisão mais completa dessa metodologia, cf. ABREU et al., 2017).

Fornecíamos um *prime* escrito e pedíamos aos participantes que escrevessem a primeira palavra que lhes vinha à cabeça quando o liam (Figura 1). Cada participante avaliou 24 *primes* experimentais – exatamente os mesmos que, após essa validação, foram utilizados no experimento principal – e 6 *primes* distratores. Cada participante podia dar apenas uma palavra como resposta. Cada estímulo, acompanhado do espaço para a reação do participante, aparecia em um quadro. Cada quadro era separado dos demais.

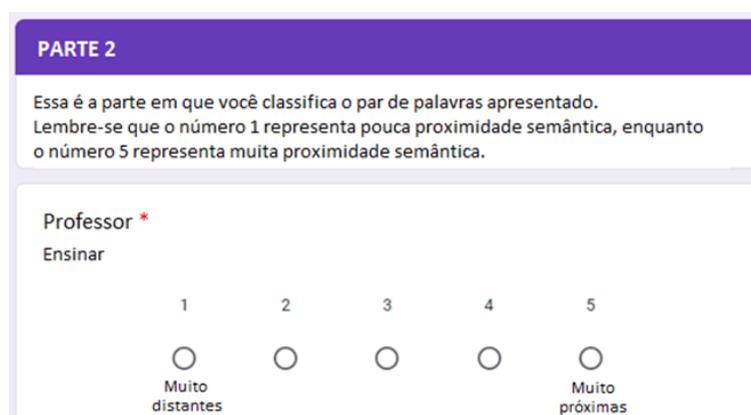
8 O teste de *cloze* consiste em suprimir (aleatória ou deliberadamente) partes de um texto e induzir a escrita do participante, que deve completar tais lacunas. Se elas puderem ser preenchidas com absolutamente qualquer palavra, dizemos que a resposta é aberta (*cloze* aberto – versão clássica do experimento). Se, no entanto, o voluntário possui uma lista de opções pré-delimitadas de resposta, dentre as quais escolherá a mais conveniente, dizemos se tratar de uma resposta fechada (*cloze* fechado).



**Figura 1.** Formulário *Google* com teste de *cloze* aberto adaptado, para colher possíveis novos alvos para as condições morfológicas. Fonte: elaborada pelas autoras.

Em uma segunda seção do mesmo formulário, formatamos a técnica de *likert scale* – ou escala *likert*, em português<sup>9</sup> (JOSHI et al., 2015), para validar a existência de relacionamento semântico entre 48 pares de palavras, que planejávamos futuramente usar como estímulos das condições unicamente semânticas (Sm3 e Sm4, cf. quadro 1) do experimento comportamental *online*.

Nesse momento, os participantes viam duas palavras escritas na tela – respectivamente os *primes* e alvos das duas condições apenas semânticas do teste de *priming* principal – e classificavam o relacionamento dessas duas palavras de 1 a 5, sendo 1 atribuído a um relacionamento muito distante entre ambas e 5 a um relacionamento muito próximo (Figura 2).



**Figura 2.** Formulário *Google* com escala *likert*, para avaliar o nível de relacionamento entre os pares das condições apenas semânticas. Fonte: elaborada pelas autoras.

<sup>9</sup> A escala *likert* é uma ferramenta psicométrica muito flexível, frequentemente utilizada em pesquisas em ciências sociais e educacionais. Ela corresponde a um conjunto de declarações sobre uma situação real ou hipotética que esteja em estudo. Solicita-se aos participantes que mostrem seus níveis de concordância com uma afirmação dada, indo, por exemplo, de *discordo totalmente* a *concordo totalmente*. Outras escalas métricas ou de opinião similares são válidas.

Cada participante avaliou 6 pares de distratores e 24 pares de palavras experimentais, os quais estavam distribuídos *Within-Subjects*: 12 eram provenientes da condição semântica com 3 sílabas (Sm3) e 12 eram provenientes da condição semântica com 4 sílabas (Sm4). Ambas serão melhor explicitadas na próxima seção, onde estão esquematizadas no quadro 1. Os distratores foram os mesmos em ambas as versões e as primeiras palavras de cada par eram as mesmas utilizadas previamente como estímulo no teste de *cloze*.

### 3.1.3. HIPÓTESES

Como respostas à parte 1 do experimento, esperávamos encontrar palavras que fossem morfologicamente relacionadas às disponibilizadas. Contudo, uma vez que o comando passado era muito aberto, tínhamos consciência de que essa operação poderia não ser tão frutífera.

Para a parte 2, por sua vez, almejávamos classificações de alto nível de relacionamento semântico entre o par de palavras, i.e., pelo menos 70% de marcações para um par estarem concentradas nos números 3, 4 e 5 (sendo esse último o correspondente a “palavras muito próximas”). Os pares em que essa situação fosse observada seriam validados como estímulo da condição unicamente semântica do experimento comportamental *online*. Caso algum par não fosse validado e as respostas do teste de *cloze* nos permitissem o aproveitamento de estímulos, assim procederíamos.

### 3.1.4. RESULTADOS

Na tabela 1, é possível verificar as respostas dadas para 14 dos 30 estímulos apresentados. As respostas dadas aos outros 16 estímulos podem ser encontradas nos documentos suplementares. Todas as respostas registradas para cada um dos 14 estímulos foram transcritas na tabela, com o número de vezes que as constamos entre parênteses. Conforme dito anteriormente, cada participante podia dar apenas uma palavra como resposta ao estímulo experimental. Observamos que as repostas fornecidas foram apenas semanticamente relacionadas aos estímulos.

Estímulo apresentado	Respostas registradas
Professor	Aluno (10), aula (7), escola (6), estudante (3), ensino (2), educador (1), disciplina (1)
Animal	Cachorro (13), zoológico (4), irracional (3), selvagem (3), estimação (2), vivo (2), ave (1), cavalo (1) vida (1)
Líquido	Água (14), sólido (4), gasoso (4), fluido (2), molhado (2), bebida (2), leite (1), suco (1)
Militar	Uniforme (5), farda (4), soldado (4), arma (4), ditadura (4), rigidez (2), patente (1), firmeza (1), golpe (1), hierarquia (1), corrupto (1), coxinha (1), colégio (1)
Palavra	Letras (6), letra (5), escrever (4), texto (3), significado (2), frase (2), comunicação (2), cantada (1), escrita (1), expressão (1), alfabeto (1), verbo (1), conversa (1)
Concerto	Música (11), abstrato (5), firme (2), show (2), cinza (2), consistente (2), chão (1), cimento (1), concerto (1), teatro (1), muro (1), instrumentos (1)
Coleta	Lixo (18), seletiva (5), dados (3), coletor (1), entrega (1), morador de rua (1), recolher (1)
Título	Texto (9), redação (5), livro (4), filme (3), prêmio (2), tema (2), eleitor (1), ideia (1), introdução (1), autor (1), parágrafo (1)
Estado	Cidade (10), Rio de Janeiro (6), país (5), município (5), localidade (1), islâmico (1), lugar (1), local (1)
Carteira	Dinheiro (23), motorista (2), couro (1), documentos (1), investimento (1), trabalho (1), assalto (1)
Sequência	Ordem (7), linearidade (4), matemática (4), números (3), progressão (2), numérica (2), didática (2), série (2), padrão (1), lógica (1), continuar (1), Star Wars (1)
Concreto	Cimento (6), chão (5), firme (3), construção (2), armadura (2), obra (2), cinza (2), engenharia (1), certo (1), matéria (1), armado (1), composto (1), líquido (1), pedra (1), de novo? (1)
Óculos	Lentes (7), sol (5), visão (4), olhos (3), miopia (3), armação (2), leitura (1), haste (1), míope (1), ódio (1), grau (1), lente (1)
Certeza	Dúvida (9), absoluta (6), segurança (3), decisão (3), incerteza (3), tranquilidade (1), afirmação (1), verdade (1), teimosia (1), fato (1), firmeza (1)

**Tabela 1.** Respostas recebidas para cada um dos estímulos do teste de *cloze*, junto a sua frequência. Entre parênteses encontra-se a quantidade de vezes que a resposta foi registrada, sendo que cada participante dava apenas uma palavra como resposta para cada estímulo.

Já na tabela 2, é possível conferir a força do relacionamento semântico entre os pares de palavras apresentados aos voluntários, isto é, o quanto os sentidos veiculados por essas duas palavras estão ligados. Para cada par de palavras, fornecemos a porcentagem de classificações em 3, 4, 5 (sendo esse último o correspondente a “palavras muito próximas”). Assim, se a tabela apresenta 100% como porcentagem para algum par, isso significa que todos os voluntários marcaram ou 3, ou 4 ou 5 como classificação do nível de relacionamento entre o par. Por fim, na tabela 3, apresenta-se a força do relacionamento semântico entre os pares de palavras distratoras.

Pares de palavras críticas (condição Sm3)	Porcentagem de classificações como 3, 4 e 5	Pares de palavras críticas (condição Sm4)	Porcentagem de classificações como 3, 4 e 5
Professor – ensinar	100%	Professor – explicação	73,3%
Animal – bestial	73,3%	Animal – criatura	86,7%
Líquido – aquoso	100%	Líquido – derretido	80%
Conversa – pappear	100%	Conversa – diálogo	80%
Militar – soldado	100%	Militar – armamento	93,3%
Correto – acerto	100%	Correto – acertado	80%
Palavra – fala	100%	Palavra – falatório	73,3%
Concerto – musical	100%	Concerto – apresentação	80%
Árvore – planta	100%	Árvore – vegetação	73,3%
Revista – jornal	86,7%	Revista – notícia	73,3%
Coleta – recolha	100%	Coleta – arrecadação	73,3%
Título – diploma	100%	Título – certificado	86,7%
Estado – cidade	100%	Estado – governador	93,3%
Carteira – moedeira	93,3%	Carteira – documento	86,7%
Partícula – pedaço	86,7%	Partícula – componente	80%
Dívida – cobrança	100%	Dívida – pagamento	80%
Sequência – listagem	93,3%	Sequência – progressão	73,3%
Concreto – sólido	93,3%	Concreto – segurança	73,3%
Sereia – oceano	93,3%	Sereia – marítimo	80%
Cavalo – equino	100%	Cavalo – hipismo	80%
Óculos – miopia	100%	Óculos – cegueira	86,7%
Cinema – película	86,6%	Cinema – exibição	93,3%
Resíduo – sujeira	100%	Resíduo – impureza	86,7%
Certeza – firmeza	93,3%	Certeza – segurança	100%

**Tabela 2.** Nível de relacionamento semântico entre os pares críticos de palavras apresentados no teste de Escala Likert. Na primeira coluna, há os pares pensados inicialmente para a condição Sm3 do experimento online e, na terceira coluna, encontram-se os pares pensados para a condição Sm4 do mesmo experimento. A diferença entre eles está na quantidade de sílabas: três para o primeiro grupo e quatro para o segundo.

Pares de palavras distratoras	Porcentagem de classificações como 3, 4 e 5
Lâmpada – garrafa	6,7%
Pirâmide – biscoito	13,35%
Cobertor – chuveiro	6,7%
Pintura – geladeira	13,35%
Parede – teclado	6,7%
Estrada – comida	13,35%

**Tabela 3.** Nível de relacionamento semântico entre os pares distratores de palavras apresentados no teste de Escala Likert.

Verificamos que, para todos os pares, pelo menos 70% dos participantes atribuíram nota 3, 4 ou 5 para classificar a proximidade semântica entre os estímulos. Já para os distratores, essa média foi de no mínimo 6,7% e no máximo 13,35%. Por fim, como dividimos essa segunda parte do experimento em duas versões, apenas 15 pessoas classificaram os pares pensados para a condição Sm3 do experimento online e as 15 outras pessoas classificaram os pares pensados para a condição Sm4. Isso não aconteceu com o teste de *cloze*, em que as 30 pessoas responderam aos mesmos 30 estímulos.

### 3.1.5. DISCUSSÃO

Quanto ao teste de *cloze*, é possível observar que, ao contrário do esperado, as palavras coletadas como resposta foram sempre apenas semanticamente – e não morfológicamente – relacionadas às palavras-estímulo. Além disso, idiosincrasias influenciaram consideravelmente os resultados. Alguns exemplos foram: i) as palavras “corrupto”, “cozinha” e “colégio” como resposta para o estímulo “militar”; ii) as palavras “Star Wars” e “didática” para o estímulo “sequência”; iii) a palavra “odeio” para o estímulo “óculos”. Em todos esses exemplos, as vivências e opiniões pessoais dos voluntários induziram suas escolhas.

Observamos, também, uma confusão entre os estímulos “concerto” e “concreto” e vimos que, em diversos momentos, as respostas dadas se juntam ao estímulo fornecido em sintagmas nominais de alta frequência de uso. É o caso, por exemplo, de “palavra cantada” (grupo musical infantil), “colégio militar”, “coleta seletiva”, “título de eleitor”, “estado islâmico”, “carteira de motorista”, “carteira de trabalho”, “sequência numérica” e “certeza absoluta”.

Por conta de todos esses fatores, não pudemos aproveitar as respostas do teste de *cloze* como estímulos para as condições morfológicas. Dessa forma, as condições morfológicas do experimento online tiveram, posteriormente, seus estímulos integralmente selecionados pelas pesquisadoras. No entanto, as idiosincrasias e a formação dos sintagmas nominais foram dados importantes para que não utilizássemos tais estímulos, problemáticos durante o teste de *cloze*, nas condições apenas semânticas. Por causa da inadequação do teste de *cloze* para nossos propósitos principais, confeccionamos nossos estímulos e procuramos rodar um outro teste de estímulos para validá-los: a Escala *Likert*.

Finalmente, usando a Escala *Likert* pudemos observar que não apenas todos os pares de estímulos críticos obtiveram pelo menos 70% de classificações de um relacionamento semântico intrapalavras satisfatório (marcações dos números 3, 4 ou 5, sendo esse último o correspondente a “palavras muito próximas”), como também pouquíssimas pessoas classificaram os pares distratores como possuindo relacionamento forte. Essa última constatação nos fez validar a assertividade do método experimental e, junto aos 70% de classificações em 3, 4 e 5, nos fez validar integralmente tais pares críticos de palavras para uso como estímulos das condições unicamente semânticas do teste comportamental *online*.

### 3.2. EXPERIMENTO COMPORTAMENTAL *ONLINE* – TESTE PRINCIPAL

A realização do experimento comportamental *online* teve como grande objetivo a aferição do curso temporal e do custo do processamento lexical, uma vez que a natureza mais espontânea dessa técnica permite a análise dos processos mais reflexos.

Estruturamos um teste de *priming* transmodal que contava com *primes* auditivos e alvos visuais, para que pudéssemos eliminar a influência de vieses derivados da exposição à ortografia. Mesmo que diversos estudos pontuem a não interferência de fatores ortográficos nos efeitos de *priming* morfológico (RASTLE et al., 2000; DOMÍNGUEZ; VEJA; BARBER, 2004; MORRIS et al., 2007; LÁZARO; GARCIA; ILLERA, 2021), a utilização do *priming* multimodal é um cuidado a mais para se evitar o viés da ortografia. Tratou-se de uma questão importante para nós porque tínhamos quatro condições experimentais em que *primes* e alvos mantinham relacionamento morfológico entre si (sendo muito difícil que, no português, não apresentassem, como consequência, coincidência ortográfica) e duas condições em que *primes* e alvos mantinham apenas relação semântica entre si, sem apresentar qualquer relacionamento ortográfico e/ou morfológico. Dessa maneira, tomamos o cuidado de evitar quaisquer complicações que nos impossibilitassem comparações intercondições.

Por fim, mantivemos um *stimulus onset asynchrony* (SOA – diferença de tempo entre *prime* e alvo) controlado de 150ms e escolhemos como tarefa cognitiva a decisão lexical, por ela instigar o acesso lexical.

Todos os alvos de todas as condições partiam do mesmo *prime*, de maneira que o tempo de acesso a cada um desses alvos pudesse ser diretamente relacionado ao tipo de relacionamento que ele mantinha com seu *prime*. Dessa maneira, esperávamos conseguir observar o comportamento linguístico de participantes durante o acesso de palavras complexas com diversos tipos de relacionamentos entre si, assim como com diferentes níveis de transparência semântica em relação a suas raízes.

### 3.2.1. PARTICIPANTES

Participaram desse experimento 45 voluntários, falantes nativos de português brasileiro, destros<sup>10</sup>, com visão normal ou corrigida e diferentes dos que haviam anteriormente realizado o experimento comportamental *offline*. Suas idades variavam entre 18 e 35 anos e todos realizavam curso de graduação na UFRJ, em uma das seguintes unidades: Centro de Ciências da Saúde (CCS), Centro de Ciências Matemáticas e da Terra (CCMN), Centro de Tecnologia (CT), Escola de Belas Artes (EBA), Escola de Comunicação (ECO) e Faculdade de Letras (FL). Acreditamos ser importante a participação de pessoas de diferentes centros/faculdades, uma vez que, na Faculdade de Letras, os alunos podem já estar mais

10 Ser destro é importante porque há indícios de que a habilidade da linguagem esteja situada majoritariamente no hemisfério esquerdo do cérebro (WILLEMS et al., 2014), o responsável pela operação do lado direito do corpo. Assim, pessoas canhotas podem se valer de diferentes técnicas para acessar a parte esquerda do cérebro enquanto utilizam linguagem, o que poderia causar diferenças nos resultados do experimento.

sensíveis e familiarizados com os procedimentos de experimentação linguística, influenciando seus resultados.

A apresentação dos estímulos se deu no Laboratório ACESIN, na Faculdade de Letras da UFRJ. Os participantes não receberam remuneração financeira pelo teste e todos assinaram termo de livre consentimento e esclarecimento.

### 3.2.2. MATERIAIS E MÉTODOS

De acordo com Bozic et al. (2007, p. 1): “*the independence of form, meaning and morphological structure can be directly investigated using derivationally complex words, because derived words can share forms but need not share meaning*”. Dessa maneira, escolhemos analisar o processamento de itens lexicais complexos no momento de sua derivação.

Cada participante foi apresentado a 24 estímulos e 24 distratores (o grupo controle, do qual faziam parte apenas não-palavras). Os *primes* foram mantidos fixos e os alvos distribuídos em quadrado latino, de forma a que todos os participantes ouvissem os mesmos 24 *primes* dos estímulos, mas variassem no alvo que acompanhava cada um deles, recebendo sempre a mesma quantidade de estímulos de cada condição. Quanto aos distratores, *prime* e alvo foram os mesmos em todas as versões do experimento.

Foram cruzadas dois tipos de variáveis independentes: i) para observarmos as diferenças entre as rotas semântica e morfológica, desenhamos estímulos com três níveis de diferenças entre *prime* e alvo: relacionamento morfológico composicional, relacionamento morfológico opaco e relacionamento semântico; e ii) para investigarmos se acessamos palavras inteiras (JACKENDOFF, 1983) ou morfema a morfema (STOCKALL; MARANTZ, 2006) desenhamos estímulos com diferentes quantidade de sílabas/camadas morfológicas (*layers*) – em dois níveis: palavras com 3 sílabas e palavras com 4 sílabas. Cruzando os dois grupos de variáveis independentes construímos um experimento 3X2, ou seja, com 6 condições distribuídas pelo sistema *Within-Subject* para investigarmos se acessamos palavras inteiras. Tais condições podem ser verificadas no quadro abaixo, que traz exemplo de três células experimentais:

Célula	Condição		Prime Auditivo	Alvo Visual
1	Sm3	Relacionamento apenas Semântico com 3 sílabas	professor	ensino
	Sm4	Relacionamento apenas Semântico com 4 sílabas	professor	educação
	MC3 ( <i>baseline</i> )	Relacionamento Morfológico Composicional com 1 camada morfológica (aprox. 3 sílabas)	professor	professor
	MC4	Relacionamento Morfológico Composicional com 2 camadas morfológicas (aprox. 4 sílabas)	professor	professoril
	MO3	Relacionamento Morfológico Opaco com 1 camada morfológica (aprox. 3 sílabas)	professor	professar
	MO4	Relacionamento Morfológico Opaco com 2 camadas morfológicas (aprox. 4 sílabas)	professor	professado

2	Sm3	Relacionamento apenas Semântico com 3 sílabas	líquido	aquoso
	Sm4	Relacionamento apenas Semântico com 4 sílabas	líquido	derretido
	MC3 ( <i>baseline</i> )	Relacionamento Morfológico Composicional com 1 camada morfológica (aprox. 3 sílabas)	líquido	líquido
	MC4	Relacionamento Morfológico Composicional com 2 camadas morfológicas (aprox. 4 sílabas)	líquido	liquidificar
	MO3	Relacionamento Morfológico Opaco com 1 camada morfológica (aprox. 3 sílabas)	líquido	liquidar
	MO4	Relacionamento Morfológico Opaco com 2 camadas morfológicas (aprox. 4 sílabas)	líquido	liquidação
3	Sm3	Relacionamento apenas Semântico com 3 sílabas	militar	soldado
	Sm4	Relacionamento apenas Semântico com 4 sílabas	militar	armamento
	MC3 ( <i>baseline</i> )	Relacionamento Morfológico Composicional com 1 camada morfológica (aprox. 3 sílabas)	militar	militar
	MC4	Relacionamento Morfológico Composicional com 2 camadas morfológicas (aprox. 4 sílabas)	militar	militarista
	MO3	Relacionamento Morfológico Opaco com 1 camada morfológica (aprox. 3 sílabas)	militar	milícia
	MO4	Relacionamento Morfológico Opaco com 2 camadas morfológicas (aprox. 4 sílabas)	militar	militante

**Quadro 1.** Condições do experimento e três exemplos de pares de palavras, nas seis condições experimentais.

Primeiro, então, os participantes ouviam um *prime* auditivo (por exemplo, *professor*) e, depois, media-se o tempo de acesso lexical a um alvo visual (escrito) moldado a uma das seis condições programadas: (i) alvo com 3 sílabas que mantinha apenas relacionamento semântico com o *prime* (*professor-ensino*); (ii) alvo com 4 sílabas que mantinha apenas relacionamento semântico com o *prime* (*professor-educação*); (iii) alvo com 1 camada morfológica (aproximadamente 3 sílabas) que mantinha relacionamento morfológico composicional (isso é, de semântica sincronicamente transparente) com o *prime* (*professor-professor*); (iv) alvo com 2 camadas morfológicas (aproximadamente 4 sílabas) que mantinha relacionamento morfológico composicional (isso é, de semântica sincronicamente transparente) com o *prime* (*professor-professoril*); (v) alvo com 1 camada morfológica (aproximadamente 3 sílabas) que mantinha relacionamento morfológico opaco (isso é, de semântica sincronicamente opaca) com o *prime* (*professor-professar*); (vi) alvo com 2 camadas morfológicas (aproximadamente 4 sílabas) que mantinha relacionamento morfológico opaco (isso é, de semântica sincronicamente opaca) com o *prime* (*professor-professado*).

Note-se que a condição (iii) MC3 foi usada como linha de base (*baseline*), pois traz a relação máxima de identidade entre *prime* e alvo: são a mesma palavra. Temos, assim, uma progressão bastante sutil entre as seis condições do experimento, começando pela plena identidade entre o *prime professor* e o alvo *professor* e progredindo para situações em que o conteúdo das palavras do par é menos relacionado, como em *professor* e *professado*

(mesmo que, nesse caso, *prime* e alvo exibam relacionamento morfológico regular e seja possível, numa análise offline posterior, reconhecer *professor* aninhado em *professado*).

Também se faz importante salientar que, apesar de o nosso *template* de estímulos ter sido muito controlado para suprir o desenho pretendido, precisamos enfrentar adversidades no campo da frequência das palavras. *Professor*, por exemplo, é uma palavra de alta frequência, enquanto *professado* é menos usado. Dessa maneira, nem sempre tendo conseguido chegar a frequências comparáveis, tivemos vieses cujas consequências pudemos perceber e apontar no presente estudo. Acreditamos, contudo, que tal dificuldade não invalide a contribuição do trabalho.

Por fim, a tarefa experimental dos participantes foi a decisão lexical, ou seja, precisavam julgar se o alvo era uma palavra ou uma não-palavra. O índice de acerto e erro dessa decisão, junto com o tempo de resposta (*response time* – RT) de cada participante eram as variáveis dependentes do experimento. Como variável independente, havia a relação entre *prime* e alvo.

### 3.2.3. DESENHO EXPERIMENTAL

Codificamos e rodamos o teste na plataforma *OpenSesame*. Trata-se de um construtor gráfico de experimentos voltado para as ciências sociais. O *OpenSesame* é gratuito, de código aberto e multiplataforma (para uma completa referência, cf. MATHÔT; SCHREIJ; THEEUWES, 2012).

Através desse programa, mostramos ao participante uma série de eventos cronologicamente ajustados ao experimento. Na figura 3, é possível ver uma linha do tempo do mesmo. Nela, tanto a cruz de fixação quanto o pré-alvo tinham como função fixar o olhar do voluntário e, conseqüentemente, sua atenção. Em seguida a elas apareciam, respectivamente, o *prime* auditivo e o alvo visual. Assim que o alvo desaparecia da tela, o participante precisava dar sua resposta, apertando um botão verde (localizado na tecla S do teclado) caso o alvo fosse uma palavra do português e um botão vermelho (localizado na tecla K do teclado) caso não o fosse. Se a resposta do participante demorasse mais de 1500ms, ela não era computada (*timeout*). No entanto, o experimento só iniciava uma nova seqüência (*trial*) quando o participante desse uma resposta à seqüência anterior.

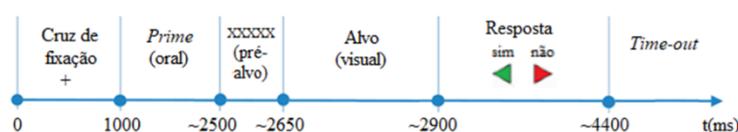


Figura 3. Cronologia do experimento. Fonte: elaborada pelas autoras.

Cada participante passou por uma fase de treino, que seguia uma linha do tempo exatamente igual à descrita na figura acima, mas que se usava de palavras não-experimentais. O proveito de se fazer um treino foi familiarizar o participante com a tarefa de decisão lexical e com o manuseio dos botões, além de checar a boa altura do som e sanar possíveis dúvidas do voluntário. Essa fase podia ser repetida quantas vezes fossem necessárias e seus resultados não foram considerados na análise.

### 3.2.4. HIPÓTESES

Em primeiro lugar, acreditamos que o acesso lexical apresentará cursos temporais distintos entre as condições morfológicas e a apenas semântica, i.e., que os RT dos voluntários apresentarão diferença estatisticamente relevante entre essas condições. Um controle valioso será também a diferença estatisticamente relevante entre o  $\Delta t$  (tempo) das condições apenas semânticas de 3 e 4 sílabas (Sm3 e Sm4) e o  $\Delta t$  das condições morfológicas de 1 e 2 camadas morfológicas, tanto as transparentes (MC3 e MC4) quanto as opacas (MO3 e MO4). Se isso for confirmado, concluiremos que há diferença entre as rotas de processamento semântica e morfológica.

Em segundo lugar, esperamos que o acesso lexical nos pares de relacionamento morfológico seja mais veloz que o dos pares de relacionamento apenas semântico. De acordo com a abordagem decomposicional do acesso lexical, isso se daria porque, uma vez que as palavras seriam montadas morfema a morfema, elas passariam por diferentes estágios derivacionais, de maneira a que, antes de chegar a *liquidificar*, a derivação atingisse a representação de *líquido*. Portanto, o alvo *liquidificar* teria a vantagem de ter recebido de herança a ativação de *líquido*. Caso isso aconteça, compreenderemos que o processamento das palavras tome um curso decomposicional, isto é, que elas sejam acessadas por suas partes e não inteiras. De acordo com as reflexões que realizamos no âmbito da Morfologia Distribuída, é o que apostamos que acontecerá. No entanto, caso essa diferença não se verifique, teremos indícios de que o acesso lexical se deu com a palavra ainda completamente montada.

Em terceiro lugar, esperamos detectar resultados positivos de priming na condição de relacionamento morfológico opaco (O). Isso aconteceria caso a média de RT dessa condição fosse estatisticamente irrelevante quando comparada à da condição morfológica composicional. Dessa maneira, poderíamos concluir que acessamos *refrigerante* por sua raiz *refrig-*, compartilhada com *refrigerar*, o que significaria que, no português brasileiro, o acesso a palavras semanticamente opacas não se diferenciaria do a palavras semanticamente transparentes. Por outro lado, caso obtenhamos resultados negativos de *priming*, acreditaremos que a lexicalização de *refrigerante* faz com que essa palavra

possua uma nova entrada em nosso léxico mental, a qual seria distinta da de *refrigerar* e apresentaria como raiz *-refrigerant*.

### 3.2.5. RESULTADOS

Apesar de os dados-resposta do experimento terem sido colhidos junto a uma população adequada à tarefa e escolhida de acordo com características homogêneas (mesma faixa etária, lado motor dominante e nível de escolaridade), existem muitas variações individuais (idiosincrasias) que devem ser levadas em consideração.

As taxas de acerto e erro, por exemplo, foram satisfatórias apenas para 42 dos originalmente 45 voluntários. Interpretamos erros demasiados cometidos por esses três voluntários como falta de atenção e/ou incompreensão da tarefa experimental e excluímos toda e qualquer resposta dessas pessoas, considerando apenas as 42 restantes para os resultados que aqui apresentaremos.

Quanto aos tempos de resposta (RT), a detecção de *outliers* é uma rotina importante para limpar os dados que se desviam excessivamente da média. Nesse sentido, tratamos os RT muito díspares através da aplicação da *Outlier Labeling Rule* (HOAGLIN; IGLEWICZ; TURKEY, 1986) aos resultados brutos. Um total de 2,34% da amostra foi classificado como *outlier* e retirado da mesma, seguindo a técnica de *trimming* (ALLEN, 2017). Em seguida, a análise estatística dos resultados sem os *outliers* foi processada pelo programa *EzAnova*, modelo *2-Within Subject Factors*.

Como podemos detectar nas sentenças de abertura do teste ANOVA, na figura 4 abaixo, há um efeito principal robusto do tipo de relacionamento (A):  $F(2,80)=18,3$ ,  $p<0,000001$ , da quantidade de sílabas/camadas morfológicas (*layers*) (B):  $F(1,40)=84,1$ ,  $p<0,000001$  e também entre os dois fatores (AxB):  $F(2,80)=25,5$ ,  $p<0,000001$ . Os tempos médios de resposta por condição podem ser verificados no gráfico 1, com seus erros padrões disponíveis na tabela 4. As comparações par a par – que demarcam significância estatística – decorrentes da amostragem estão disponíveis na tabela 5.

```
ANOVA: Design 2 Within Subject Factors
A F(2,80) = 18,3 p<0,000001 SS=350032,34 MSE=9582,04
Greenhouse-Geisser{0,6967} p<0,0000109 Huynh-Feldt{0,714} p<0,0000088
B F(1,40) = 84,1 p<0,000001 SS=536324,61 MSE=6380,41
A*B F(2,80) = 25,5 p<0,000001 SS=428844,48 MSE=8405,36
```

Figura 4. Efeitos Principais da ANOVA.

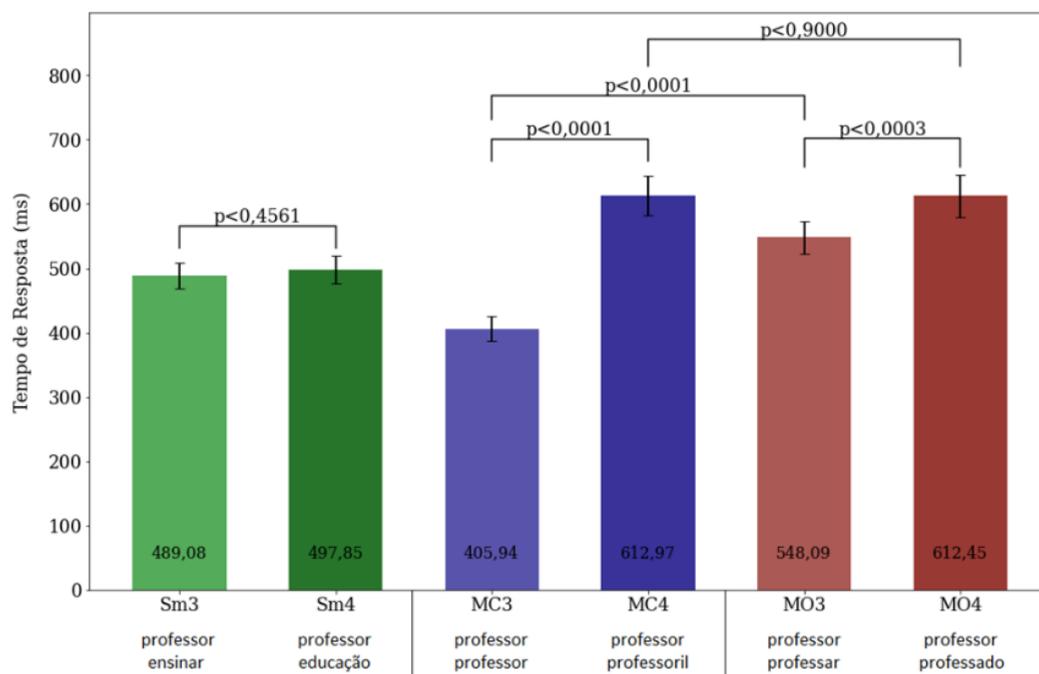


Gráfico 1. Tempos de resposta (RT) por condição.

#	Sm3	Sm4	MC3	MC4	MO3	MO4
Erro Padrão	19,95	21,04	19,06	30,65	25,64	33,54

Tabela 4. Valores de erro padrão para cada uma das condições.

Comparações par a par	
[MC_3]vs[MC_4] t(38)=9,28 p< 0,0001	[MC_4]vs[Sm_4] t(38)=5,91 p< 0,0001
[MC_3]vs[MO_3] t(38)=8,86 p< 0,0001	[MO_3]vs[MO_4] t(38)=3,99 p< 0,0003
[MC_3]vs[MO_4] t(38)=9,43 p< 0,0001	[MO_3]vs[Sm_3] t(38)=4,14 p< 0,0002
[MC_3]vs[Sm_3] t(38)=5,51 p< 0,0001	[MO_3]vs[Sm_4] t(38)=4,08 p< 0,0002
[MC_3]vs[Sm_4] t(38)=7,03 p< 0,0001	[MO_4]vs[Sm_3] t(38)=5,91 p< 0,0001
[MC_4]vs[MO_3] t(38)=3,05 p< 0,0041	[MO_4]vs[Sm_4] t(38)=6,28 p< 0,0001
[MC_4]vs[MO_4] t(38)=0,13 p< 0,9000	[Sm_3]vs[Sm_4] t(38)=0,75 p< 0,4561
[MC_4]vs[Sm_3] t(38)=7,59 p< 0,0001	

Tabela 5. Comparações par a par geradas pelo programa *EzAnova*, valores significativos apresentam p<0,05.

É possível perceber que as condições semânticas (Sm) apresentaram um tempo médio de resposta baixo, mas com uma diferença estatisticamente irrelevante entre Sm3 e Sm4. No entanto, ambas foram maiores que a condição de relacionamento morfológico composicional com 1 camada morfológica/3 sílabas (MC3). Já a condição de relacionamento morfológico composicional de 2 camadas morfológicas/4 sílabas (MC4) foi muito mais alta que MC3, apresentando também diferença estatisticamente irrelevante à condição de relacionamento morfológico opaco de 2 camadas morfológicas/4 sílabas (MO4). Além disso, a condição de relacionamento morfológico opaco de 1 camada morfológica/3 sílabas (MO3) foi menor que

MO4, através de uma diferença estatisticamente relevante, e maior que MC3, também de maneira estatisticamente relevante.

Por fim, a dispersão dos resultados, por tipo de relacionamento e por quantidade de sílabas/camadas morfológicas, está apresentada nos gráficos 2 e 3, respectivamente. É possível perceber que, quanto ao tipo de relacionamento, MO foi o que obteve os resultados mais díspares, enquanto Sm obteve os menos dispersos. Já quanto à quantidade de sílabas/camadas morfológicas, as condições com 4 sílabas/2 camadas morfológicas apresentaram dispersão maior que as de 3 sílabas/1 camada morfológica.



Gráfico 2. Dispersão por tipo de relacionamento.

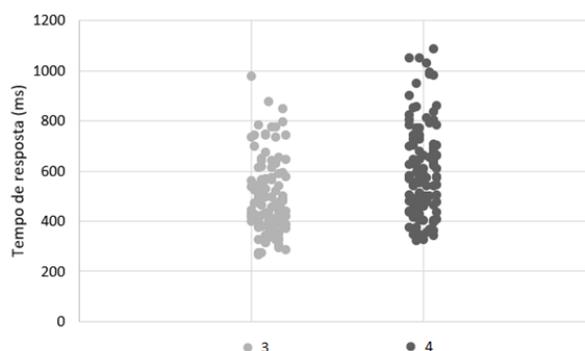


Gráfico 3. Dispersão por quantidade de sílabas/camadas morfológicas.

Essas dispersões podem indicar processamento não reflexo (aquele que não é instintivo e espontâneo, ou seja, que sofre influência da reflexão ativa) e uma maior dificuldade sentida pelo participante durante a realização da tarefa lexical.

### 3.2.6. DISCUSSÃO

Em primeiro lugar, é válido ressaltar que o tempo médio de resposta de MC3 (405,94ms), que tinha como alvo exatamente a mesma palavra do *prime*, se aproxima do tempo médio normal de processamento neurolinguístico, o N400, conforme registrado em diversos

estudos com EEG (para um exemplo, *cf.* GOMES, 2009). Isso nos indica, além do sucesso do *prime* no acesso ao alvo (isto é, que ter acessado tal *prime* antes de tal alvo ajudou no processamento do último), a existência de uma relação de identidade entre *primes* e alvos dessa condição. Trata-se, na realidade, da identidade mais forte que é possível encontrar entre duas palavras: elas serem exatamente iguais. Apesar de alguns estudos, como Bozic et al. (2007), não encontrarem ativação de *priming* para condições com relacionamento de identidade, em nosso estudo os RT dessa condição foram um valioso controle para a tomarmos como linha de base de comparação com das demais

Nessa linha de raciocínio, ao compararmos os RT de MC3 (405,94ms) com os de MO3 (548,09ms), verificamos que os primeiros são consideravelmente menores que os segundos, configurando uma diferença estatisticamente relevante ( $p < 0,0001$ ). Em primeiro lugar, isso nos sugere que exista uma relação de referencialidade entre *primes* e alvos de MO3 (*cf.* FRANÇA et al., 2008; PYLKKÄNEN et al., 2002; 2003; 2004; PYLKKÄNEN; MARANTZ, 2003; RASTLE et al, 2000). Em segundo lugar, essa diferença nos sugere que, para que se chegue a *professar* a partir de *professor*, passe-se por uma renegociação do sentido da raiz, já que a primeira palavra é morfológica e diacronicamente derivada da segunda, mas sincronicamente apartada semanticamente da mesma (STOCKAL; MARANTZ, 2006).

Essa interpretação sustentaria ainda que tal renegociação seria apenas encontrada nos pares de relacionamento morfológico opaco (MO), os únicos a apresentarem RT tão altos na condição com apenas uma camada morfológica (MC3 vs MO3 vs Sm3). Isso, por sua vez, ratificaria pressupostos decomposicionais não lexicalistas de teorias como a Morfologia Distribuída (HALLE; MARANTZ, 1993), segundo a qual poderíamos propor que tal renegociação de sentido se daria na primeira camada morfológica (constituída de raiz + morfema categorizador), que é onde ocorreria a arbitrariedade saussuriana (FRANÇA et al., 2008). Assim, o *prime* falharia, i.e., sua herança seria negativa e ele não ajudaria o processamento do alvo. Isso faria com que o voluntário precisasse desativar esse *prime* e voltar à Enciclopédia para renegociar o sentido da primeira camada morfológica de *profess-*, o que aumentaria consideravelmente o tempo de processamento dessa palavra. Tais reflexões nos permitem analisar nossa terceira hipótese experimental, que foi refutada: como o *priming* foi negativo, entendemos que o processo de lexicalização de uma palavra faz com que uma nova entrada na Enciclopédia seja criada para a mesma. No entanto, ao contrário do que havíamos pensado anteriormente, não acreditamos que isso signifique que o processamento de palavras semanticamente opacas seja diferente do de semanticamente transparentes.

Ainda segundo a MD, depois de construída a primeira base sobre a qual discursamos no parágrafo anterior, o processamento das outras camadas morfológicas herdaria o sentido especial da raiz e tomaria um curso composicional, sem que se precisasse acessar

novamente a Enciclopédia. Analisando os RT de MO4 (612,45ms) à luz dessa teoria, poderíamos propor que, nas condições de relacionamento morfológico opaco, depois que a difícil negociação entre a raiz e o primeiro morfema categorizador (de *profess+or* para *profess+ar*) é feita, restaria apenas uma adição simples e composicional de um segundo morfema (de *profess+ar* para *profess+a+do*), em MO4, o que parece ser a razão de haver um aumento muito pequeno nos RT de MO3 para MO4. Esse aumento, no entanto, é estatisticamente relevante ( $p < 0,0003$ ), comprovando uma diferença de processamento entre tais condições. França et al. (2008) também encontraram, proporcionalmente, aumentos semelhantes de RT para operações morfológicas composicionais. Com essa análise, podemos validar a nossa terceira hipótese experimental, propondo que as palavras semanticamente opacas, tal qual as transparentes, sejam acessadas por suas partes, isto é, que seu processamento siga um curso decomposicional assim que a nova entrada da Enciclopédia é recuperada. Achamos essa interpretação válida, mesmo constatando que os RT de MO3 e MO4 foram consideravelmente maiores que os de Sm3 (489,08ms) e Sm4 (497,85), que, em nossas hipóteses, esperávamos serem maiores.

Todavia, para validar completamente tal hipótese, é preciso comparar as condições de relacionamento morfológico às de relacionamento apenas semântico. Para tanto, não só a comparação de RT entre tipos diferentes de relacionamento será importante, como também o  $\Delta t$  entre as diferentes quantidades de sílabas/camadas morfológicas dos diferentes relacionamentos será mais uma vez muito relevante. Debruçando-nos sobre tais resultados, pudemos ver que as condições semânticas Sm3 e Sm4 foram um valioso controle. Em primeiro lugar, elas apresentaram RT muito parecidos entre si, a ponto de a diferença entre eles ser estatisticamente irrelevante ( $p < 0,4561$ ). Isso nos faz concluir que, para os pares da condição semântica, o aumento na quantidade de sílabas não produziu diferença no processamento dos alvos. Tal afirmativa, por sua vez, nos remete à nossa primeira hipótese experimental, sugerindo a existência de tipos diferentes de processamento nas rotas morfológica e semântica. Enquanto na primeira o acesso lexical se daria por um algoritmo linguístico de formação de palavra, a segunda teria como suporte um mecanismo mnemônico, contando com memória de uso conjunto e redes de associações, motivo pelo qual o acréscimo de uma sílaba não geraria mudança significativa.

Além disso, o fato de, apesar de muito baixos, os RT de Sm3 (489,08ms) serem maiores do que os de MC3 (405,94ms), sendo a diferença entre ambos estatisticamente relevante ( $p < 0,0001$ ), corrobora nossa proposta de que MC3 apresenta relação máxima de identidade entre *prime* e alvo: mesmo uma computação linguística rápida como a mnemônica foi mais custosa que a de MC3. A comparação de Sm3 (489,08ms) com MO3 (548,09ms) ( $p < 0,0002$ ) também contribui para a interpretação que fizemos dos resultados da condição morfológicamente opaca (MO), já que a ativação de uma rede de associação parece ser menos custosa que a renegociação do sentido de uma palavra na Enciclopédia.

Por fim, também corrobora a assumpção de que palavras morfológicamente complexas sejam decompostas em suas partes antes do acesso lexical.

No entanto, ao tentarmos comparar todas as condições de 4 sílabas/2 camadas morfológicas entre si, isso é, comparar Sm4 (497,85ms) com MC4 (612,97ms) e MO4 (612,45ms), encontramos um problema – o mesmo enfrentado ao se tentar comparar MC3 (405,94ms) com MC4: os altíssimos RT de MC4, que foram surpreendentemente demorados para uma computação composicional de semântica transparente. Esperávamos uma disparidade entre MC3 e MC4 por conta do acréscimo composicional de uma camada morfológica no segundo, mas não uma diferença tão marcada. Os RTs de MC4 quase se igualam aos de MO4 (não apresentam diferença estatisticamente relevante em relação a eles), que, pela nossa interpretação, passariam por um processo muito mais custoso e demorado. Acreditamos ter sido a baixa frequência das palavras-alvo de MC4 a responsável por essa situação. Enquanto *professor* é uma palavra muito comum, *professoril* é, na realidade, muito incomum e pode até mesmo gerar dúvida quanto ao estatuto de palavra do português. Dessa forma, a diferença estatisticamente irrelevante ( $p < 0,9867$ ) entre MC4 e MO4 e a grande disparidade entre MC3 e MC4 ( $p < 0,0001$ ) foram interpretados como conseqüências do relaxamento no controle da frequência de MC4 e tentarão ser corrigidos em um futuro experimento.

#### 4. CONSIDERAÇÕES FINAIS

Com o trabalho apresentado, esperamos ter contribuído para o melhor entendimento do funcionamento do acesso lexical, um processo que é tão rápido quanto usual. A interpretação de nossos resultados apontou para a existência de um processamento decomposicional em nossos processos linguísticos mentais, o qual demandaria menos espaço de armazenamento e mais aplicações de regras. Além disso, através dos resultados inferimos a presença de duas rotas de processamento: uma semântica e uma morfológica. A primeira funcionaria através de memória de uso conjunto e de redes de associação, enquanto a segunda seria sustentada por algoritmos linguísticos de formação de palavras.

Por fim, respondendo à pergunta principal do estudo, obtivemos resultados negativos de *priming* para o processamento de palavras que perderam relação semântica sincrônica com suas raízes. No entanto, observamos que tal processamento é decomposicional. Sugerimos, então, que palavras idiomatizadas ganhem novas entradas em nosso léxico mental, a Enciclopédia, e que, após o acesso lexical à sua [raiz + morfema categorizador], isto é, à sua parte arbitrária, seu processamento também se dê morfema a morfema. Em

resumo, apontamos a existência de múltiplos acessos lexicais, originados de processos neurofisiologicamente diferentes.

Como próximos passos para nossa pesquisa, pretendemos replicar o experimento apresentado, aplicando-o com um aparelho de eletroencefalografia (EEG), na busca de analisar a contraparte cerebral dos processos mentais que aqui pudemos observar. A observação do tempo de processamento, sem que ele esteja influenciado por uma resposta comportamental, será um grande ganho para o estudo. Aproveitaremos para realizar melhorias em nossos estímulos e em nosso design experimental, variando o SOA, melhor controlando a frequência dos itens experimentais críticos e adicionando condições de palavras sem nenhum tipo de relacionamento entre si, a fim de funcionarem como controle. Essa condição controle constituiria uma *baseline* diametralmente oposta à de MC3, oportunizando o refinamento de nossas interpretações.

Acreditamos que, dessa maneira, poderemos analisar e descrever com ainda mais precisão as particularidades dos processos mentais envolvidos no acesso lexical, em suas vias morfológica e semântica (tanto opaca quanto transparente) de processamento.

## REFERÊNCIAS

ABREU, K.N.M; GARCIA, D.C.; HORA, K.F.P.N.A; SOUZA, C.R. O teste de Cloze como instrumento de medida da proficiência em leitura: fatores linguísticos e não linguísticos. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 25, n. 3, p. 1767-1799, 2017. DOI: <http://doi.org/10.17851/2237-2083.25.3.1767-1799>.

ALLEN, M. Data trimming. In: *The SAGE Encyclopedia of Communication Research Methods*, 2017. DOI: <https://doi.org/10.4135/9781483381411.n130>.

BEVER, T; POEPEL, D. Analysis by Synthesis: A (Re-)Emerging Program of Research for Language and Vision. *Biolinguistics*, v.4, p. 174-200, 2010.

BOZIC, M.; MARSLÉN-WILSON, W. D.; STAMATAKIS, E. A.; DAVIS, M. H.; TYLER, L. K. Differentiating Morphology, Form, and Meaning: Neural Correlates of Morphological Complexity. *Journal of Cognitive Neuroscience*, v. 19, n. 9, p. 1464-1475, 2007. DOI: <http://doi.org/10.1162/jocn.2007.19.9.1464>.

BUTTERWORTH, B. Lexical Representation. In: BUTTERWORTH, B. (Ed.) *Language Production*. London: Academic Press, v. 2, p. 257-294, 1983.

CARAMAZZA, A.; LAUDANNA, A. & ROMANI, C. Lexical Access and Inflectional Morphology. *Cognition*, v. 28, p. 297-332, 1988. DOI: [https://doi.org/10.1016/0010-0277\(88\)90017-0](https://doi.org/10.1016/0010-0277(88)90017-0).

CHOMSKY, N. *Syntactic structures*. The Hague/Paris, Mouton, 1957.

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

DOMÍNGUEZ, A.; VEGA, M.; BARBER, H. Event-related Brain Potentials Elicited by Morphological, Homographic, Orthographic, and Semantic Priming. *Journal of Cognitive Neuroscience*, v. 16, n. 4, p. 598-608, 2004. DOI: <http://doi.org/10.1162/089892904323057326>.

EMBICK, D.; HACKL M.; SCHAEFFER, J.; KELEPIR, M.; MARANTZ, A. A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research*, v. 10, n. 3, p. 345-8, 2001. DOI: [http://doi.org/10.1016/s0926-6410\(00\)00053-7](http://doi.org/10.1016/s0926-6410(00)00053-7).

- FRANÇA, A. I.; LEMLE, M.; GESUALDI, A. R.; CAGY, M.; INFANTOSI, A, F. C. A neurofisiologia do acesso lexical: palavras em português. *Revista Veredas*. Juiz de Fora, v. 2, p. 14-39, 2008.
- FORSTER, K. I. Levels of processing and the structure of the language processor. In: COOPER, W. E.; WALKER, E. C. T. (Eds.). *Sentence processing: psycholinguistics studies presented to Merrill Garrett*, p. 27-85. Hillsdale, NJ: Erlbaum, 1979.
- GARCIA, D. C. *Elementos estruturais do acesso lexical: o reconhecimento de palavras multimorfêmicas no português brasileiro*. 88 p. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.
- GOMES, J. N. *A Direcionalidade no Relacionamento Semântico: um Estudo de ERP*. 162 p. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.
- HALLE, M.; MARANTZ, A. Distributed Morphology and the pieces of inflection. In: HALE, K. L.; KEYSER, S. J. (Eds.). *The view from building 20*. Cambridge: MIT Press, p. 111-176, 1993.
- HALLE, M; STEVENS, K. Analysis by synthesis. In: WATHEN-DUNN, W.; WOODS, L. E. (Eds.). *Proceeding of the Seminar on Speech Compression and Processing*, Vol. II, paper D7, 1959.
- HAY, J. B.; BAAYEN, R. H. Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, v. 9, n. 7, p. 342-348, 2005. DOI: <http://doi.org/10.1016/j.tics.2005.04.002>.
- HOAGLIN, D.; IGLEWICZ, B.; TURKEY, J.W. Performance of some resistant rules for outlier labeling. *Journal of American Statistical Association*, v. 81, p. 991-999, 1986. DOI: <https://doi.org/10.2307/2289073>.
- JACKENDOFF, R. *Semantics and Cognition*. Cambridge. Mass: MIT Press, 1983.
- JOSHI, A.; KALE, S.; CHANDEL, S.; PAL, D. K. Likert Scale: Explored and Explained. *Current Journal of Applied Science and Technology*, v. 7, n. 4, p 396-403, 2015. DOI: <http://doi.org/10.9734/BJAST/2015/14975>.
- KENEDY, E. Psicolinguística na descrição gramatical. In: MAIA, M. (Org.). *Psicolinguística, psicolinguísticas: uma introdução*, p. 143-156. São Paulo: Contexto, 2015.
- LAZARO, L.; GARCÍA, L.; ILLERA, V. Morpho-orthographic segmentation of opaque and transparent derived words: New evidence for Spanish. *Quarterly Journal of Experimental Psychology*, v. 74, n. 5, p. 944-954, 2021. DOI: <http://doi.org/10.1177/1747021820977038>.
- LEMLE, M.; FRANÇA, A. I. Arbitrariedade Saussureana em Foco. *Revista Letras*. Curitiba, Paraná, v. 69, n. Maio/Ago, p. 269-288, 2006.
- MAIA, M.; LEMLE, M.; FRANÇA, A. Efeito stroop e rastreamento ocular no processamento de palavras. *Ciências & Cognição* (UFRJ), v. 12, p. 02-17, 2007.
- MAIA, M.; RIBEIRO, A. Jabuticaba liboramina lê mais fácil do que jornaleiro norbalense: um estudo de rastreamento ocular de palavras e pseudopalavras mono e polimorfêmicas. In: BUCHWEITZ, A.; MOTTA, M. (Org.). *Linguagem & Cognição*. 1ed. Porto Alegre: EDIPUC-RS, p. 143-154, 2015.
- MARANTZ, A. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In: *Proceedings of the 21st Annual Penn Linguistics Colloquium*. University of Pennsylvania Working Papers in Linguistics (PWPL), Pennsylvania: v. 4, n. 2, p. 201-225, 1997.
- MARSLÉN-WILSON, W.; TYLER, L. K.; WAKSLER, R.; OLDER, L. Morphology and Meaning in the English Mental Lexicon. *Psychological Review*, v. 101, n. 1, p. 3-33, 1994. DOI: <https://doi.org/10.1037/0033-295X.101.1.3>.
- MATHÔT, S.; SCHREIJ, D.; THEEUWES, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavioral Research Methods*, v. 44, p. 314-324, 2012. DOI: <https://doi.org/10.3758/s13428-011-0168-7>.
- MORRIS, J.; FRANK, T.; GRAINGER, J.; HOLCOMB, P. J. Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, v. 44, p. 506-521, 2007. DOI: <http://doi.org/10.1111/j.1469-8986.2007.00538.x>.
- PEDERNEIRA, I. L.; LEMLE, M. Word-internal syntactic complexity. *Revista da ABRALIN*, v. 14, p. 389-410, 2015.
- PINKER, S. Rules of Language. *Science*, v. 253, p. 530-535, 2 August 1991. DOI: <http://doi.org/10.1126/science.1857983>.

PYLKKÄNEN, L.; STRINGFELLOW, A.; FLAGG, E.; MARANTZ, A. A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. *Proc. 12th Int. Conf. Biomagnetism*, Helsinki University of Technology, p. 363–367, 2000.

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(1-3), p. 666–78, 2002. DOI: <https://doi.org/10.1006/brln.2001.2555>.

PYLKKÄNEN, L.; MARANTZ, A. Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, v. 7, p.187–189, 2003. DOI: [http://doi.org/10.1016/s1364-6613\(03\)00092-5](http://doi.org/10.1016/s1364-6613(03)00092-5).

PYLKKÄNEN, L.; GONNERMAN, L.; STRINGFELLOW, A.; MARANTZ, A. *Disambiguating the source of phonological inhibition effects in lexical decision: an MEG study*. (ms), 2003.

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, n. 3, p. B35-B45, 2004. DOI: <https://doi.org/10.1016/j.cognition.2003.09.008>.

PYLKKÄNEN, L. The neural basis of combinatory syntax and semantics. *Science*, v. 366, p. 62-66, 4 October, 2019. DOI: <https://doi.org/10.1126/science.aax0050>.

RASTLE, K.; DAVIS, M.; MARSLÉN-WILSON, W.; TYLER, L. K. Morphological and semantic effects in visual word recognition: a time course study. *Language and Cognitive Processes*, v. 15, n. 4-5, p. 507-538, 2000. DOI: <https://doi.org/10.1080/016909600050119689>.

STOCKALL, L.; MARANTZ, A. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, v. 1, n. 1, 2006. DOI: <https://doi.org/10.1075/ml.1.1.07sto>.

STROOP, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, v. 18, p. 643-662, 1935. <https://doi.org/10.1037/h0054651>.

TAFT, M.; FORSTER, K. I. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning & Verbal Behavior*, v. 14, n. 6, p. 638-647, 1975. DOI: [https://doi.org/10.1016/S0022-5371\(75\)80051-X](https://doi.org/10.1016/S0022-5371(75)80051-X).

TAFT, M.; FORSTER, K. I. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, v. 15, n. 6, p. 607-620, 1976. DOI: [http://doi.org/10.1016/0022-5371\(76\)90054-2](http://doi.org/10.1016/0022-5371(76)90054-2).

TAFT, M. Lexical access via an orthographic code: the basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, v. 18, n. 1, p. 21-39, 1979. DOI: [https://doi.org/10.1016/S0022-5371\(79\)90544-9](https://doi.org/10.1016/S0022-5371(79)90544-9).

VAN DEN BUSSCHE, E.; VAN DEN NOORTGATE, W.; REYNVOET, B. Mechanisms of Masked Priming: A Meta-Analysis. *Psychological bulletin*, v. 135, p. 452-77, 2009. DOI: <http://doi.org/10.1037/a0015329>.

WILLEMS, R.; VAN DER HAEGEN, L.; FISHER, S.; FRANCKS, C. (2014). On the other hand: Including left-handers in cognitive neuroscience and neurogenetics. *Nature reviews. Neuroscience*. DOI: <http://doi.org/10.1038/nrn3679>.